VI. Multilinear Algebra, 248-305

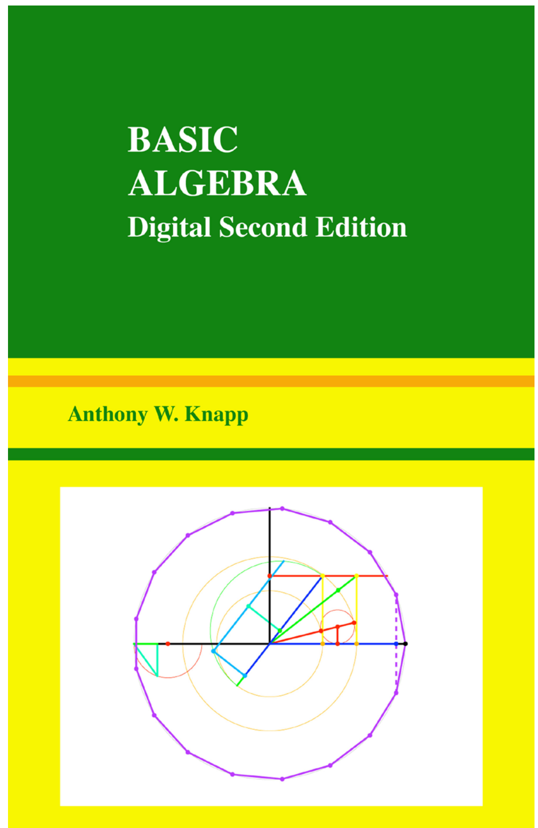
DOI: 10.3792/euclid/9781429799980-6

from

Basic Algebra

Digital Second Edition

Anthony W. Knapp



Full Book DOI: 10.3792/euclid/9781429799980
ISBN: 978-1-4297-9998-0

Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/∼aknapp

Title: Basic Algebra
Cover: Construction of a regular heptadecagon, the steps shown in color sequence; see page 505.

Mathematics Subject Classification (2010): 15–01, 20–01, 13–01, 12–01, 16–01, 08–01, 18A05, 68P30.

First Edition, ISBN-13 978-0-8176-3248-9

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the Author

CHAPTER VI

Multilinear Algebra

Abstract. This chapter studies, in the setting of vector spaces over a field, the basics concerning multilinear functions, tensor products, spaces of linear functions, and algebras related to tensor products.

Sections 1–5 concern special properties of bilinear forms, all vector spaces being assumed to be finite-dimensional. Section 1 associates a matrix to each bilinear form in the presence of an ordered basis, and the section shows the effect on the matrix of changing the ordered basis. It then addresses the extent to which the notion of "orthogonal complement" in the theory of inner-product spaces applies to nondegenerate bilinear forms. Sections 2–3 treat symmetric and alternating bilinear forms, producing bases for which the matrix of such a form is particularly simple. Section 4 treats a related subject, Hermitian forms when the field is the complex numbers. Section 5 discusses the groups that leave some particular bilinear and Hermitian forms invariant.

Section 6 introduces the tensor product of two vector spaces, working with it in a way that does not depend on a choice of basis. The tensor product has a universal mapping property—that bilinear functions on the product of the two vector spaces extend uniquely to linear functions on the tensor product. The tensor product turns out to be a vector space whose dual is the vector space of all bilinear forms. One particular application is that tensor products provide a basis-independent way of extending scalars for a vector space from a field to a larger field. The section includes a number of results about the vector space of linear mappings from one vector space to another that go hand in hand with results about tensor products. These have convenient formulations in the language of category theory as "natural isomorphisms."

Section 7 begins with the tensor product of three and then n vector spaces, carefully considering the universal mapping property and the question of associativity. The section defines an algebra over a field as a vector space with a bilinear multiplication, not necessarily associative. If E is a vector space, the tensor algebra $T(E)$ of E is the direct sum over $n \geq 0$ of the n-fold tensor product of E with itself. This is an associative algebra with a universal mapping property relative to any linear mapping of E into an associative algebra A with identity: the linear map extends to an algebra homomorphism of $T(E)$ into A carrying 1 into 1.

Sections 8–9 define the symmetric and exterior algebras of a vector space E. The symmetric algebra $S(E)$ is a quotient of $T(E)$ with the following universal mapping property: any linear mapping of E into a commutative associative algebra A with identity extends to an algebra homomorphism of $S(E)$ into A carrying 1 into 1. The symmetric algebra is commutative. Similarly the exterior algebra $\bigwedge(E)$ is a quotient of $T(E)$ with this universal mapping property: any linear mapping l of E into an associative algebra A with identity such that $l(v)^2 = 0$ for all $v \in E$ extends to an algebra homomorphism of $\bigwedge(E)$ into A carrying 1 into 1.

The problems at the end of the chapter introduce some other algebras that are of importance in applications, and the problems relate some of these algebras to tensor, symmetric, and exterior algebras. Among the objects studied are Lie algebras, universal enveloping algebras, Clifford algebras, Weyl algebras, Jordan algebras, and the division algebra of octonions.

1. Bilinear Forms and Matrices

This chapter will work with vector spaces over a common field of "scalars," which will be called $\mathbb{K}$. In Section 6 a field containing $\mathbb{K}$ as a subfield will briefly play a role, and that will be called $\mathbb{L}$.

If V is a vector space over $\mathbb{K}$, a **bilinear form** on V is a function from $V \times V$ into $\mathbb{K}$ that is linear in each variable when the other variable is held fixed.

EXAMPLES.

(1) For general $\mathbb{K}$, take $V = \mathbb{K}^n$. Any matrix A in $M_n(\mathbb{K})$ determines a bilinear form by the rule $\langle v, w\rangle = v^t A w$.

(2) For $\mathbb{K} = \mathbb{R}$, let V be an inner-product space, in the sense of Chapter III, with inner product $(\,\cdot\,,\,\cdot\,)$. Then $(\,\cdot\,,\,\cdot\,)$ is a bilinear form on V.

Multilinear functionals on a vector space of row vectors, also called k-linear functionals or k-multilinear functionals, were defined in the course of working with determinants in Section II.7, and that definition transparently extends to general vector spaces. A bilinear form on a general vector space is then just a 2-linear functional. From the point of view of definitions, the words "functional" and "form" are interchangeable here, but the word "form" is more common in the bilinear case because of a certain homogeneity that it suggests and that comes closer to the surface in Corollary 6.12 and in Section 7.

For the remainder of this section, all vector spaces will be finite-dimensional.

Bilinear forms, i.e., 2-linear functionals, are of special interest relative to k-linear functionals for general k because of their relationships with matrices and linear mappings. To begin with, each bilinear form, in the presence of an ordered basis, is given by a matrix. In more detail let V be a finite-dimensional vector space, and let $\langle\,\cdot\,,\,\cdot\,\rangle$ be a bilinear form on V. If an ordered basis $\Gamma = (v_1, \dots, v_n)$ of V is specified, then the bilinear form determines the matrix B with entries $B_{ij} = \langle v_i, v_j\rangle$. Conversely we can recover the bilinear form from B as follows: Write $v = \sum_i a_i v_i$ and $w = \sum_j b_j v_j$. Then

$$\langle v, w\rangle = \big\langle \textstyle\sum_i a_i v_i,\ \sum_j b_j v_j\big\rangle = \sum_{i,j} a_i \langle v_i, v_j\rangle b_j .$$

In other words, $\langle v, w\rangle = a^t B b$, where $a = \begin{pmatrix} v \\ \Gamma \end{pmatrix}$ and $b = \begin{pmatrix} w \\ \Gamma \end{pmatrix}$ in the notation of Section II.3. Therefore

$$\langle v, w\rangle = \begin{pmatrix} v \\ \Gamma \end{pmatrix}^t B \begin{pmatrix} w \\ \Gamma \end{pmatrix}.$$

Consequently we see that all bilinear forms on a finite-dimensional vector space reduce to Example 1 above—once we choose an ordered basis.

Let us examine the effect of a change of ordered basis. Suppose that $\Gamma = (v_1, \ldots, v_m)$ and $\Delta = (w_1, \ldots, w_n)$, and let B and C be the matrices of the bilinear form in these two ordered bases: $B_{ij} = \langle v_i, v_j \rangle$ and $C_{ij} = \langle w_i, w_j \rangle$. Let the two bases be related by $w_j = \sum_i a_{ij} v_i$, i.e., let $[a_{ij}] = \begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}$. Then we have

$$C_{ij} = \langle w_i, w_j \rangle = \Big\langle \sum_k a_{ki} v_k, \sum_l a_{lj} v_l \Big\rangle = \sum_{k,l} a_{ki} a_{lj} \langle v_k, v_l \rangle = \sum_{k,l} a_{ki} B_{kl} a_{lj}.$$

Translating this formula into matrix form, we obtain the following proposition.

Proposition 6.1. Let $\langle \cdot , \cdot \rangle$ be a bilinear form on a finite-dimensional vector space V, let Γ and Δ be ordered bases of V, and let B and C be the respective matrices of $\langle \cdot , \cdot \rangle$ relative to Γ and Δ. Then

$$C = \begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}^t B \begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}.$$

The qualitative conclusion about the matrices may be a little unexpected. It is not that they are similar but that they are related by $C = S^t B S$ for some nonsingular square matrix S. In particular, B and C need not have the same determinant.

Guided by the circle of ideas around the Riesz Representation Theorem for inner products (Theorem 3.12), let us examine what happens when we fix one of the variables of a bilinear form and work with the resulting linear map. Thus again let $\langle \cdot , \cdot \rangle$ be a bilinear form on V. For fixed u in V, $v \mapsto \langle u, v \rangle$ is a linear functional on V, thus a member of the dual space V' of V. If we write $L(u)$ for this linear functional, then L is a function from V to V' satisfying $L(u)(v) = \langle u, v \rangle$. The formula for L shows that L is in fact a linear function. We define the **left radical**, lrad, of $\langle \cdot , \cdot \rangle$ to be the kernel of L; thus

$$\operatorname{lrad}\big(\langle \cdot , \cdot \rangle\big) = \{u \in V \mid \langle u, v \rangle = 0 \text{ for all } v \in V\}.$$

Similarly we let $R : V \to V'$ be the linear map $R(v)(u) = \langle u, v \rangle$, and we define the **right radical**, rrad, of $\langle \cdot , \cdot \rangle$ to be the kernel of R; thus

$$\operatorname{rrad}\big(\langle \cdot , \cdot \rangle\big) = \{v \in V \mid \langle u, v \rangle = 0 \text{ for all } u \in V\}.$$

EXAMPLE 1, CONTINUED. The vector space V is the space $\mathbb{K}^n$ of n-dimensional column vectors, the dual V' is the space of n-dimensional row vectors, A is an n-by-n matrix with entries in $\mathbb{K}$, and $\langle\,\cdot\,,\,\cdot\,\rangle$ is given by $\langle u, v\rangle = u^t A v = L(u)(v) = R(v)(u)$ for u and v in $\mathbb{K}^n$. Explicit formulas for L and R are given by

$$L(u) = u^t A = (A^t u)^t$$

and

$$R(v) = (Av)^t.$$

Thus

$$\operatorname{lrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big) = \ker L = \text{null space}(A^t),$$
$$\operatorname{rrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big) = \ker R = \text{null space}(A).$$

Since A is square and since the row rank and column rank of A are equal, the dimensions of the null spaces of A and A^t are equal. Hence

$$\dim \operatorname{lrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big) = \dim \operatorname{rrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big).$$

This equality of dimensions for the case of $\mathbb{K}^n$ extends to general V, as is noted in the next proposition.

Proposition 6.2. If $\langle\,\cdot\,,\,\cdot\,\rangle$ is any bilinear form on a finite-dimensional vector space V, then

$$\dim \operatorname{lrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big) = \dim \operatorname{rrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big).$$

PROOF. We saw above that computations with bilinear forms of V reduce, once we choose an ordered basis for V, to computations with matrices, row vectors, and column vectors. Thus the argument just given in the continuation of Example 1 is completely general, and the proposition is proved. □

A bilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ is said to be **nondegenerate** if its left radical is 0. In view of the Proposition 6.2, it is equivalent to require that the right radical be 0. When the radicals are 0, the associated linear maps L and R from V to V' are one-one. Since $\dim V = \dim V'$, it follows that L and R are onto V'. Thus a nondegenerate bilinear form on V sets up two canonical isomorphisms of V with its dual V'.

For definiteness let us work with the linear mapping $L : V \to V'$ given by $L(u)(v) = \langle u, v\rangle$. If $U \subseteq V$ is a vector subspace, define

$$U^\perp = \{u \in V \mid \langle u, v\rangle = 0 \text{ for all } v \in U\}.$$

It is apparent from the definitions that

$$\boxed{U \cap U^\perp = \operatorname{lrad}\big(\langle\,\cdot\,,\,\cdot\,\rangle\big)\big|_{U\times U}\,.}$$

In contrast to the special case that $\mathbb{K} = \mathbb{R}$ and the bilinear form is an inner product, $U \cap U^\perp$ may be nonzero even if $\langle \cdot\,, \cdot \rangle$ is nondegenerate. For example let $V = \mathbb{R}^2$, define

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 - x_2 y_2,$$

and suppose that U is the 1-dimensional vector subspace $U = \left\{ \begin{pmatrix} x_1 \\ x_1 \end{pmatrix} \right\}$. The matrix of the bilinear form in the standard ordered basis is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; since the matrix is nonsingular, the bilinear form is nondegenerate. Direct calculation shows that $U^\perp = \left\{ \begin{pmatrix} y_1 \\ y_1 \end{pmatrix} \right\} = U$, so that $U \cap U^\perp \neq 0$. Nevertheless, in the nondegenerate case the dimensions of U and $U^\perp$ behave as if $U^\perp$ were an orthogonal complement. The precise result is as follows.

Proposition 6.3. If $\langle \cdot\,, \cdot \rangle$ is a nondegenerate bilinear form on the finite-dimensional vector space V and if U is a vector subspace of V, then

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

PROOF. Define $\ell : V \to U'$ by $\ell(v)(u) = \langle v, u \rangle$ for $v \in V$ and $u \in U$. The definition of $U^\perp$ shows that $\ker \ell = U^\perp$. To see that image $\ell = U'$, choose a vector subspace U_1 of V with $V = U \oplus U_1$, let u' be in U', and define v' in V' by

$$v' = \begin{cases} u' & \text{on } U, \\ 0 & \text{on } U_1. \end{cases}$$

Since $\langle \cdot\,, \cdot \rangle$ is nondegenerate, the linear mapping $L : V \to V'$ is onto V'. Thus we can choose $v \in V$ with $L(v) = v'$. Then

$$\ell(v)(u) = \langle v, u \rangle = L(v)(u) = v'(u) = u'(u)$$

for all u in U, and hence $\ell(v) = u'$. Therefore image $\ell = U'$, and we conclude that

$$\dim V = \dim(\ker \ell) + \dim(\text{image } \ell) = \dim U^\perp + \dim U' = \dim U^\perp + \dim U.$$

□

Corollary 6.4. If $\langle \cdot\,, \cdot \rangle$ is a nondegenerate bilinear form on the finite-dimensional vector space V and if U is a vector subspace of V, then $V = U \oplus U^\perp$ if and only if $\langle \cdot\,, \cdot \rangle\big|_{U \times U}$ is nondegenerate.

PROOF. Corollary 2.29 and Proposition 6.3 together give

$$\dim(U + U^\perp) + \dim(U \cap U^\perp) = \dim U + \dim U^\perp = \dim V.$$

Thus $U + U^\perp = V$ if and only if $U \cap U^\perp = 0$, if and only if $\langle \cdot\,, \cdot \rangle\big|_{U \times U}$ is nondegenerate. The result therefore follows from Proposition 2.30. □

2. Symmetric Bilinear Forms

We continue with the setting in which $\mathbb{K}$ is a field and all vector spaces of interest are defined over $\mathbb{K}$ and are finite-dimensional.

A bilinear form $\langle \cdot\,, \cdot \rangle$ on V is said to be **symmetric** if $\langle u, v\rangle = \langle v, u\rangle$ for all u and v in V, **skew-symmetric** if $\langle u, v\rangle = -\langle v, u\rangle$ for all u and v in V, and **alternating** if $\langle u, u\rangle = 0$ for all u in V.

"Alternating" always implies "skew-symmetric." In fact, if $\langle \cdot\,, \cdot \rangle$ is alternating, then $0 = \langle u+v, u+v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle = \langle u, v\rangle + \langle v, u\rangle$; thus $\langle \cdot\,, \cdot \rangle$ is skew-symmetric. If $\mathbb{K}$ has characteristic different from 2, then the converse is valid: "skew-symmetric" implies "alternating." In fact, if $\langle \cdot\,, \cdot \rangle$ is skew-symmetric, then $\langle u, u\rangle = -\langle u, u\rangle$ and hence $2\langle u, u\rangle = 0$; thus $\langle u, u\rangle = 0$, and $\langle \cdot\,, \cdot \rangle$ is alternating.

Let us examine further the effect of the characteristic of $\mathbb{K}$. If, on the one hand, $\mathbb{K}$ has characteristic different from 2, the most general bilinear form $\langle \cdot\,, \cdot \rangle$ is the sum of the symmetric form $\langle \cdot\,, \cdot \rangle_s$ and the alternating form $\langle \cdot\,, \cdot \rangle_a$ given by

$$\langle u, v\rangle_s = \tfrac{1}{2}(\langle u, v\rangle + \langle v, u\rangle),$$
$$\langle u, v\rangle_a = \tfrac{1}{2}(\langle u, v\rangle - \langle v, u\rangle).$$

In this sense the symmetric and alternating bilinear forms are the extreme cases among all bilinear forms, and we shall study the two cases separately.

If, on the other hand, $\mathbb{K}$ has characteristic 2, then "alternating" implies "skew-symmetric" but not conversely. "Alternating" is a serious restriction, and we shall be able to deal with it. However, "symmetric" and "skew-symmetric" are equivalent since $1 = -1$, and thus neither condition is much of a restriction; we shall not attempt to say anything insightful in these cases.

In this section we study symmetric bilinear forms, obtaining results when $\mathbb{K}$ has characteristic different from 2. From the symmetry it is apparent that the left and right radicals of a symmetric bilinear form are the same, and we call this vector subspace the **radical** of the form. By way of an example, here is a continuation of Example 1 from the previous section.

EXAMPLE. Let $V = \mathbb{K}^n$, let A be a **symmetric** n-by-n matrix (i.e., one with $A^t = A$), and let $\langle u, v\rangle = u^t Av$. The computation $\langle v, u\rangle = v^t Au = (v^t Au)^t = u^t A^t v = u^t Av = \langle u, v\rangle$ shows that the bilinear form $\langle \cdot\,, \cdot \rangle$ is symmetric; the second equality $v^t Au = (v^{\text{th}} Au)^t$ holds since $v^t Au$ is a 1-by-1 matrix.

Again the example is completely general. In fact, if $\Gamma = (v_1, \dots, v_n)$ is an ordered basis of a vector space V and if $\langle \cdot\,, \cdot \rangle$ is a given symmetric bilinear form on V, then the matrix of the form has entries $A_{ij} = \langle v_i, v_j\rangle$, and these evidently satisfy $A_{ij} = A_{ji}$. So A is a symmetric matrix, and computations with the bilinear form are reduced to those used in the example.

Theorem 6.5 (Principal Axis Theorem). Suppose that $\mathbb{K}$ has characteristic different from 2.

(a) If $\langle\cdot,\ \cdot\rangle$ is a symmetric bilinear form on a finite-dimensional vector space V, then there exists an ordered basis of V in which the matrix of $\langle\cdot,\ \cdot\rangle$ is diagonal.

(b) If A is an n-by-n symmetric matrix, then there exists a nonsingular n-by-n matrix M such that M^tAM is diagonal.

REMARKS. Because computations with general symmetric bilinear forms reduce to computations in the special case of a symmetric matrix and because Proposition 6.1 tells the effect of a change of ordered basis, (a) and (b) amount to the same result; nevertheless, we give two proofs of Theorem 6.5—a proof via matrices and a proof via linear maps. A hint of the validity of the theorem comes from the case that $\mathbb{K}=\mathbb{R}$. For the field $\mathbb{R}$ when the bilinear form is an inner product, the Spectral Theorem (Theorem 3.21) says that there is an orthonormal basis of eigenvectors and hence that (a) holds. When $\mathbb{K}=\mathbb{R}$, the same theorem says that there exists an orthogonal matrix M with $M^{-1}AM$ diagonal; since any orthogonal matrix M satisfies $M^{-1}=M^t$, the Spectral Theorem is saying that (b) holds.

PROOF VIA MATRICES. If A is an n-by-n symmetric matrix, we seek a nonsingular M with M^tAM diagonal. We induct on the size of A, the base case of the induction being $n=1$, where there is nothing to prove. Assume the result to be known for size $n-1$, and write the given n-by-n matrix A in block form as $A=\begin{pmatrix} a & b \\ b^t & d\end{pmatrix}$ with d of size 1-by-1. If $d\neq 0$, let x be the column vector $-d^{-1}b$. Then

$$\begin{pmatrix} I & x \\ 0 & 1\end{pmatrix}\begin{pmatrix} a & b \\ b^t & d\end{pmatrix}\begin{pmatrix} I & 0 \\ x^t & 1\end{pmatrix}=\begin{pmatrix} * & 0 \\ 0 & d\end{pmatrix},$$

and the induction goes through. If $d=0$, we argue in a different way. We may assume that $b\neq 0$ since otherwise the result is immediate by induction. Say $b_i\neq 0$ with $1\leq i\leq n-1$. Let y be an $(n-1)$-dimensional row vector with i^{th} entry a member δ of $\mathbb{K}$ to be specified and with other entries 0. Then

$$\begin{pmatrix} I & 0 \\ y & 1\end{pmatrix}\begin{pmatrix} a & b \\ b^t & 0\end{pmatrix}\begin{pmatrix} I & y^t \\ 0 & 1\end{pmatrix}=\begin{pmatrix} * & * \\ * & yay^t+b^ty^t+yb\end{pmatrix}=\begin{pmatrix} * & * \\ * & \delta^2a_{ii}+2\delta b_i\end{pmatrix}.$$

Since $\mathbb{K}$ has characteristic different from 2, $2b_i$ is not 0; thus there is some value of δ for which $\delta^2a_{ii}+2\delta b_i\neq 0$. Then we are reduced to the case $d\neq 0$, which we have already handled, and the induction goes through. □

PROOF VIA LINEAR MAPS. We may assume that the given symmetric bilinear form is not identically 0, since otherwise any basis will do. Let the radical of the form be denoted by $\text{rad}=\text{rad}\big(\langle\cdot,\ \cdot\rangle\big)$. Choose a vector subspace S of V such that $V=\text{rad}\oplus S$, and put $[\cdot,\ \cdot]=\langle\cdot,\ \cdot\rangle\big|_{S\times S}$. Then $[\cdot,\ \cdot]$ is a symmetric bilinear form on S, and it is nondegenerate. In fact, $[u, \cdot\,] = 0$ means $\langle u, v\rangle = 0$ for all $v \in S$; since $\langle u, v\rangle = 0$ for v in rad anyway, $\langle u, v\rangle = 0$ for all $v \in V$, u is in rad as well as S, and $u = 0$.

Since $\langle\,\cdot\,,\,\cdot\,\rangle$ is not identically 0, the subspace S is not 0. Thus the nondegenerate symmetric bilinear form $[\,\cdot\,,\,\cdot\,]$ on S is not 0. Since

$$[u, v] = \tfrac{1}{2}\big([u + v, u + v] - [u, u] - [v, v]\big),$$

it follows that $[v, v] \neq 0$ for some v in S. Put $U_1 = \mathbb{K}v$. Then $[\,\cdot\,,\,\cdot\,]\big|_{U_1\times U_1}$ is nondegenerate, and Corollary 6.4 implies that $S = U_1 \oplus U_1^\perp$. Applying the converse direction of the same corollary to $U_1^\perp$, we see that $[\,\cdot\,,\,\cdot\,]\big|_{U_1^\perp\times U_1^\perp}$ is nondegenerate. Repeating this construction with $U^\perp$ and iterating, we obtain

$$V = \mathrm{rad} \oplus U_1 \oplus \cdots \oplus U_k$$

with $\langle U_i, U_j\rangle = 0$ for $i \neq j$ and with $\dim U_i = 1$ for all i. This completes the proof. □

Theorem 6.5 fails in characteristic 2. Problem 2 at the end of the chapter illustrates the failure.

Let us examine the matrix version of Theorem 6.5 more closely when $\mathbb{K}$ is $\mathbb{C}$ or $\mathbb{R}$. The theorem says that if A is n-by-n symmetric, then we can find a nonsingular M with $B = M^tAM$ diagonal. Taking D diagonal and forming $C = D^tBD$, we see that we can adjust the diagonal entries of B by arbitrary nonzero squares. Over $\mathbb{C}$, we can therefore arrange that C is of the form $\mathrm{diag}(1, \dots, 1, 0, \dots, 0)$. The number of 1's equals the rank, and this has to be the same as the rank of the given matrix A. The form is nondegenerate if and only if there are no 0's. Thus we understand everything about the diagonal form.

Over $\mathbb{R}$, matters are more subtle. We can arrange that C is of the form $\mathrm{diag}(\pm1, \dots, \pm1, 0, \dots, 0)$, the various signs ostensibly not being correlated. Replacing C by P^tCP with P a permutation matrix, we may assume that our diagonal matrix is of the form $\mathrm{diag}(+1, \dots, +1, -1, \dots, -1, 0, \dots, 0)$. The number of $+1$'s and -1's together is again the rank of A, and the form is nondegenerate if and only if there are no 0's. But what about the separate numbers of $+1$'s and -1's? The triple given by

$$(p, m, z) = \big(\#(+1)\text{'s}, \#(-1)\text{'s}, \#(0)\text{'s}\big)$$

is called the **signature** of A when $\mathbb{K} = \mathbb{R}$. A similar notion can be defined in the case of a symmetric bilinear form over $\mathbb{R}$.

Theorem 6.6 (Sylvester's Law). The signature of an n-by-n symmetric matrix over $\mathbb{R}$ is well defined.

PROOF. The integer $p + m$ is the rank, which does not change under a transformation $A \mapsto M^t AM$ if M is nonsingular. Thus we may take z as known. Let (p', m', z) and (p, m, z) be two signatures for a symmetric matrix A, with $p' \leq p$. Define the corresponding symmetric bilinear form on $\mathbb{R}^n$ by $\langle u, v\rangle = u^t Av$. Let $(v_1', \dots, v_n')$ and $(v_1, \dots, v_n)$ be ordered bases of $\mathbb{R}^n$ diagonalizing the bilinear form and exhibiting the resulting signature, i.e., having $\langle v_i', v_j'\rangle = \langle v_i, v_j\rangle = 0$ for $i \neq j$ and having

$$\langle v_j', v_j'\rangle = \begin{cases} +1 & \text{for } 1 \leq j \leq p', \\ -1 & \text{for } p'+1 \leq j \leq n-z, \\ 0 & \text{for } n-z+1 \leq j \leq n, \end{cases}$$

$$\langle v_j, v_j\rangle = \begin{cases} +1 & \text{for } 1 \leq j \leq p, \\ -1 & \text{for } p+1 \leq j \leq n-z, \\ 0 & \text{for } n-z+1 \leq j \leq n. \end{cases}$$

We shall prove that $\{v_1, \dots, v_p, v'_{p'+1}, \dots, v_n'\}$ is linearly independent, and then we must have $p' \geq p$. Reversing the roles of p and p', we see that $p' = p$ and $m' = m$, and the theorem is proved. Thus suppose we have a linear dependence:

$$a_1 v_1 + \cdots + a_p v_p = b_{p'+1} v'_{p'+1} + \cdots + b_n v_n'.$$

Let v be the common value of the two sides of this equation. Then

$$\langle v, v\rangle = \langle a_1 v_1 + \cdots + a_p v_p,\ a_1 v_1 + \cdots + a_p v_p\rangle = \sum_{j=1}^{p} a_j^2 \geq 0$$

and

$$\langle v, v\rangle = \langle b_{p'+1} v'_{p'+1} + \cdots + b_n v_n',\ b_{p'+1} v'_{p'+1} + \cdots + b_n v_n'\rangle = -\sum_{j=p'+1}^{n-z} b_j^2 \leq 0.$$

We conclude that $\langle v, v\rangle = 0$, $\sum_{j=1}^p a_j^2 = 0$, and $a_1 = \cdots = a_p = 0$. Thus $v = 0$ and $b_{p'+1} v'_{p'+1} + \cdots + b_n v_n' = 0$. Since $\{v'_{p'+1}, \dots, v_n'\}$ is linearly independent, we obtain also $b_{p'+1} = \cdots = b_n = 0$. Therefore $\{v_1, \dots, v_p, v'_{p'+1}, \dots, v_n'\}$ is a linearly independent set, and the proof is complete. □

3. Alternating Bilinear Forms

We continue with the setting in which $\mathbb{K}$ is a field and all vector spaces of interest are defined over $\mathbb{K}$ and are finite-dimensional.

In this section we study alternating bilinear forms, imposing no restriction on the characteristic of $\mathbb{K}$. From the skew symmetry of any alternating bilinear form it is apparent that the left and right radicals of such a form are the same, and we call this vector subspace the **radical** of the form. First let us consider examples given in terms of matrices. Temporarily let us separate matters according to the characteristic.

EXAMPLE 1 OF SECTION 1 WITH $\mathbb{K}$ OF CHARACTERISTIC $\neq 2$. Let $V = \mathbb{K}^n$, let A be a **skew-symmetric** n-by-n matrix (i.e., one with $A^t = -A$), and let $\langle u, v\rangle = u^t A v$. The computation $\langle v, u\rangle = v^t A u = (v^t A u)^t = u^t A^t v = -u^t A v = -\langle u, v\rangle$ shows that the bilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ is skew-symmetric, hence alternating.

EXAMPLE 1 OF SECTION 1 WITH $\mathbb{K}$ OF CHARACTERISTIC $= 2$. Let $V = \mathbb{K}^n$, let A be an n-by-n matrix, and define $\langle u, v\rangle = u^t A v$. We suppose that A is skew-symmetric; it is the same to assume that A is symmetric since the characteristic is 2. In order to have $\langle e_i, e_i\rangle = 0$ for each standard basis vector, we shall assume that $A_{ii} = 0$ for all i. If u is a column vector with entries $u_1, \dots, u_n$, then $\langle u, u\rangle = u^t A u = \sum_{i,j} u_i A_{ij} u_j = \sum_{i\neq j} u_i A_{ij} u_j = \sum_{i<j} (A_{ij} u_i u_j + A_{ji} u_i u_j) = \sum_{i<j} 2A_{ij} u_i u_j = 0$. Hence the bilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ is alternating.

Again the examples are completely general. In fact, if $\Gamma = (v_1, \dots, v_n)$ is an ordered basis of a vector space V and if $\langle\,\cdot\,,\,\cdot\,\rangle$ is a given alternating bilinear form, then the matrix of the form has entries $A_{ij} = \langle v_i, v_j\rangle$ that evidently satisfy $A_{ij} = -A_{ji}$ and $A_{ii} = 0$. So A is a skew-symmetric matrix with 0's on the diagonal, and computations with the bilinear form are reduced to those used in the examples. To keep the terminology parallel, let us say that a square matrix is **alternating** if it is skew-symmetric and has 0's on the diagonal.

Theorem 6.7.

(a) If $\langle\,\cdot\,,\,\cdot\,\rangle$ is an alternating bilinear form on a finite-dimensional vector space V, then there exists an ordered basis of V in which the matrix of $\langle\,\cdot\,,\,\cdot\,\rangle$ has the form

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

If $\langle\,\cdot\,,\,\cdot\,\rangle$ is nondegenerate, then $\dim V$ is even.

(b) If A is an n-by-n alternating matrix, then there exists a nonsingular n-by-n matrix M such that $M^t AM$ is as in (a).

PROOF. It is enough to prove (a). Let rad be the radical of the given form $\langle\,\cdot\,,\,\cdot\,\rangle$, and choose a vector subspace S of V with $V = \mathrm{rad} \oplus S$. The restriction of $\langle\,\cdot\,,\,\cdot\,\rangle$ to S is then alternating and nondegenerate. We may now proceed by induction on $\dim V$ under the assumption that $\langle\,\cdot\,,\,\cdot\,\rangle$ is nondegenerate. For $\dim V = 1$, the form is degenerate. For $\dim V = 2$, we can find u and v with $\langle u, v\rangle \neq 0$, and we can normalize one of the vectors to make $\langle u, v\rangle = 1$. Then (u, v) is the required ordered basis.

Assuming the result in the nondegenerate case for dimension $< n$, suppose that $\dim V = n$. Again choose u and v with $\langle u, v\rangle = 1$, and define $U = \mathbb{K}u \oplus \mathbb{K}v$. Then $\langle\,\cdot\,,\,\cdot\,\rangle\big|_{U\times U}$ has matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and is nondegenerate. By Corollary 6.4, $V = U \oplus U^\perp$, and an application of the converse of the corollary shows that $\langle\,\cdot\,,\,\cdot\,\rangle\big|_{U^\perp\times U^\perp}$ is nondegenerate. The induction hypothesis applies to $U^\perp$, and we obtain the desired matrix for the given form. □

4. Hermitian Forms

In this section the field will be $\mathbb{C}$, and V will be a finite-dimensional vector space over $\mathbb{C}$.

A **sesquilinear form** $\langle\,\cdot\,,\,\cdot\,\rangle$ on V is a function from $V \times V$ into $\mathbb{C}$ that is linear in the first variable and conjugate linear in the second.[1] Sesquilinear forms do not make sense for general fields because of the absence of a universal analog of complex conjugation, and we shall consequently work only with the field $\mathbb{C}$ in this section.[2]

A sesquilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ is **Hermitian** if $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all u and v in V. The form is **skew-Hermitian** if instead $\langle u, v\rangle = -\overline{\langle v, u\rangle}$ for all u and v in V. Hermitian and skew-Hermitian forms are the extreme types of sesquilinear forms since any sesquilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ is the sum of a Hermitian form $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathrm{h}}$ and a skew-Hermitian form $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathrm{sh}}$ given by

$$\langle u, v\rangle_{\mathrm{h}} = \tfrac{1}{2}(\langle u, v\rangle + \overline{\langle v, u\rangle}),$$
$$\langle u, v\rangle_{\mathrm{sh}} = \tfrac{1}{2}(\langle u, v\rangle - \overline{\langle v, u\rangle}).$$

[1] Some authors, particularly in mathematical physics, reverse the roles of the two variables and assume the conjugate linearity in the first variable instead of the second.

[2] Sesquilinear forms make sense in number fields like $\mathbb{Q}[\sqrt{2}\,]$ that have an automorphism of order 2 (see Section IV.1), but sesquilinear forms in this kind of setting will not concern us here.

In addition, any skew-Hermitian form becomes a Hermitian form simply by multiplying by i. Specifically if $\langle\,\cdot\,,\,\cdot\,\rangle_{\text{sh}}$ is skew-Hermitian, then $i\langle\,\cdot\,,\,\cdot\,\rangle_{\text{sh}}$ is sesquilinear and Hermitian, as is readily checked. Consequently the study of skew-Hermitian forms immediately reduces to the study of Hermitian forms.

EXAMPLE. Let $V = \mathbb{C}^n$, and let A be a **Hermitian** matrix, i.e., one with $A^* = A$, where A^* is the conjugate transpose of A. Then it is a simple matter to check that $\langle u, v\rangle = v^* A u$ defines a Hermitian form on $\mathbb{C}^n$.

Again the example with a matrix is completely general. In fact, let $\langle\,\cdot\,,\,\cdot\,\rangle$ be a Hermitian form on V, let $\Gamma = (v_1, \dots, v_n)$ be an ordered basis of V, and define $A_{ij} = \langle v_i, v_j\rangle$. Then A is a Hermitian matrix, and $\langle u, v\rangle = u^t A \bar{v}$, where $\bar{v}$ is the entry-by-entry complex conjugate of v.

If $\Delta = (w_1, \dots, w_n)$ is a second ordered basis, then the formula for changing basis may be derived as follows: Write $w_j = \sum_i c_{ij} v_i$, so that $[c_{ij}]$ is the matrix $\begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}$. If $B_{ij} = \langle w_i, w_j\rangle$, then $B_{ij} = \langle w_i, w_j\rangle = \sum_{kl} c_{ki}\langle v_k, v_l\rangle \bar{c}_{lj}$, and hence

$$B = \begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}^t A\, \overline{\begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}}.$$

Thus two Hermitian matrices A and B represent the same Hermitian form in different bases if and only if $B = M^* A M$ for some nonsingular matrix M.

Proposition 6.8.

(a) If $\langle\,\cdot\,,\,\cdot\,\rangle$ is a Hermitian form on a finite-dimensional vector space V over $\mathbb{C}$, then there exists an ordered basis of V in which the matrix of $\langle\,\cdot\,,\,\cdot\,\rangle$ is diagonal with real entries.

(b) If A is an n-by-n Hermitian matrix, then there exists a nonsingular n-by-n matrix M such that $M^* A M$ is diagonal.

PROOF. The above considerations show that (a) and (b) are reformulations of the same result. Hence it is enough to prove (b). By the Spectral Theorem (Theorem 3.21), there exists a unitary matrix U such that $U^{-1} A U$ is diagonal with real entries. Since U is unitary, $U^{-1} = U^*$. Thus we can take $M = U$ to prove (b). □

Just as with symmetric bilinear forms over $\mathbb{R}$, we can do a little better than Proposition 6.8 indicates. If B is Hermitian and diagonal with diagonal entries b_i, and if D is diagonal with positive entries d_i, then $C = D^* B D$ is diagonal with diagonal entries $d_i^2 b_i$. Choosing D suitably and then replacing C by $P^t C P$ for a suitable permutation matrix P, we may assume that $P^t C P$ is of the form $\text{diag}(+1, \ldots, +1, -1, \ldots, -1, 0, \ldots, 0)$. The number of $+1$'s and -1's together is the rank of A, and the form is nondegenerate if and only if there are no 0's. The triple given by

$$(p, m, z) = \big(\#(+1)\text{'s}, \#(-1)\text{'s}, \#(0)\text{'s}\big)$$

is again called the **signature** of A. A similar notion can be defined in the case of a Hermitian form, as opposed to a Hermitian matrix.

Theorem 6.9 (Sylvester's Law). The signature of an n-by-n Hermitian matrix is well defined.

The proof is the same as for Theorem 6.6 except for adjustments in notation.

5. Groups Leaving a Bilinear Form Invariant

Although it is not logically necessary to do so, we digress in this section to introduce some important groups that are defined by means of bilinear or Hermitian forms. These groups arise in many areas of mathematics, both pure and applied, and their detailed structure constitutes a topic in the fields of Lie groups, algebraic groups, and finite groups that is beyond the scope of this book. Thus the best place to define them seems to be now.

We limit our comments on applications to just these: When the underlying field in the definition of these groups is $\mathbb{R}$ or $\mathbb{C}$, the group is quite often a "simple Lie group," one of the basic building blocks of the theory of the continuous groups that so often arise in topology, geometry, differential equations, and mathematical physics. When the underlying field is a number field in the sense of Example 9 of Section IV.1, the group quite often plays a role in algebraic number theory. When the underlying field is a finite field, the group is often closely related to a finite simple group; an example of this relationship occurred in Problems 55–62 at the end of Chapter IV, where it was shown that the group $\text{PSL}(2, \mathbb{K})$, built in an easy way from the general linear group $\text{GL}(2, \mathbb{K})$, is simple if the field $\mathbb{K}$ has more than 5 elements. More general examples of finite simple groups produced by analogous constructions are said to be of "Lie type." A celebrated theorem of the late twentieth century classified the finite simple groups—establishing that the only such groups are the cyclic groups of prime order, the alternating groups on 5 or more letters, the simple groups of Lie type, and 26 so-called sporadic simple groups.

If $\langle\,\cdot\,,\,\cdot\,\rangle$ is a bilinear form on an n-dimensional vector space V over a field $\mathbb{K}$, a nonsingular linear map $g : V \to V$ is said to **leave the bilinear form invariant** if

$$\langle g(u), g(v)\rangle = \langle u, v\rangle$$

for all u and v in V. Fix an ordered basis Γ of V, let A be the matrix of the bilinear form in this basis, let $g' = \begin{pmatrix} g \\ \Gamma\Gamma \end{pmatrix}$ be the member of $\mathrm{GL}(n, \mathbb{K})$ corresponding to g, and abbreviate $\begin{pmatrix} w \\ \Gamma \end{pmatrix}$ as w' for any w in V. To translate the invariance condition into one concerning matrices, we use the formula $\langle u, v\rangle = u'^t Av'$, the corresponding formula for $\langle g(u), g(v)\rangle$, and the formula $g(w)' = g'(w')$ from Theorem 2.14. Then we obtain $u'^t g'^t Ag'v' = u'^t Av'$. Taking u to be the i^{th} member of the ordered basis Γ and v to be the j^{th} member, we obtain equality of the $(i, j)^{\text{th}}$ entry of the two matrices $g'^t Ag'$ and A. Thus the matrix form of the invariance condition is that a nonsingular matrix g' satisfy

$$g'^t Ag' = A.$$

We know that changing the ordered basis Γ amounts to replacing A by $M^t AM$ for some nonsingular matrix M. If g' satisfies the invariance condition $g'^t Ag' = A$ relative to A, then $M^{-1}g'M$ satisfies

$$(M^{-1}g'M)^t(M^t AM)(M^{-1}g'M) = M^t AM.$$

Thus we are led to a conjugate subgroup within $\mathrm{GL}(n, \mathbb{K})$. A conjugate subgroup is not something substantially new, and thus we might as well make a convenient choice of basis so that A looks particularly special.

The interesting cases are that the given bilinear form is symmetric or alternating, hence that the matrix A is symmetric or alternating. Let us restrict our attention to them. The left and right radicals coincide in these cases, and the first thing to do is to take the two-sided radical into account. Returning to the original bilinear form, we write $V = \mathrm{rad} \oplus S$, where rad is the radical and S is some vector subspace of S, and we choose an ordered basis $(v_1, \dots, v_p, v_{p+1}, \dots, v_n)$ such that $v_1, \dots, v_p$ are in S and $v_{p+1}, \dots, v_n$ are in rad. Then $\langle v_i, v_j\rangle = 0$ if $i > p$ or $j > p$, and consequently A has its only nonzero entries in the upper left p-by-p block. The same argument as in the proofs of Theorems 6.5 and 6.7 shows that the restriction of the bilinear form to S is nondegenerate, and consequently the upper left p-by-p block of A is nonsingular. Changing notation slightly, suppose that g is an n-by-n matrix written in block form as $g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$ with g_{11} of size p-by-p, suppose that $\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$ is another matrix written in the same block form, suppose that the p-by-p matrix A is nonsingular, and suppose that $g^t \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} g = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$. Making a brief computation, we find that necessary and sufficient conditions on g are that g_{11} be nonsingular and have $g_{11}^t Ag_{11} = A$, that $g_{12} = 0$, that g_{22} be arbitrary nonsingular, and that g_{21} be arbitrary. In other words, the only interesting condition $g_{11}^t A g_{11} = A$ is a reflection of what happens in the nonsingular case.

Consequently the interesting cases are that the given bilinear form is nondegenerate, as well as either symmetric or alternating. If A is symmetric and nonsingular, then the group of all nonsingular matrices g such that $g^t A g = A$ is called the **orthogonal group** relative to A. If A is alternating and nonsingular, then the group of all nonsingular matrices g such that $g^t A g = A$ is called the **symplectic group** relative to A.

For the symplectic case it is customary to invoke Theorem 6.7 and take A to be

$$J = \begin{pmatrix} \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & & \\ & \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix}} \end{pmatrix},$$

except possibly for a permutation of the rows and columns and possibly for a multiplication by -1. Two conflicting notations are in common use for the symplectic group, namely $Sp(n, \mathbb{K})$ and $Sp(\frac{1}{2}n, \mathbb{K})$, and one always has to check a particular author's definitions.

For the orthogonal case the notation is less standardized. Theorem 6.5 says that we may take A to be diagonal except when $\mathbb{K}$ has characteristic 2. But the theorem does not tell us exactly which A's are representative of the same bilinear form. When $\mathbb{K} = \mathbb{C}$, we know that we can take A to be the identity matrix I. The group is known as the complex orthogonal group and is denoted by $\mathrm{O}(n, \mathbb{C})$. When $\mathbb{K} = \mathbb{R}$, we can take A to be diagonal with diagonal entries ± 1. Sylvester's Law (Theorem 6.6) says that the form determines the number of $+1$'s and the number of -1's. The groups are called indefinite orthogonal groups and are denoted by $\mathrm{O}(p, q)$, where p is the number of $+1$'s and q is the number of -1's. When $q = 0$, we obtain the ordinary orthogonal group of matrices relative to an inner product.

A similar analysis applies to Hermitian forms. The field is now $\mathbb{C}$, the invariance condition with the form is still $\langle g(u), g(v) \rangle = \langle u, v \rangle$, and the corresponding condition with matrices is $g^t A \bar{g} = A$. The interesting case is that the Hermitian form is nondegenerate. Proposition 6.8 and Sylvester's Law (Theorem 6.9) together show that we may take A to be diagonal with diagonal entries ± 1 and that the Hermitian form determines the number of $+1$'s and the number of -1's. The groups are the indefinite unitary groups and are denoted by $\mathrm{U}(p, q)$, where p is the number of $+1$'s and q is the number of -1's. When $q = 0$, we obtain the ordinary unitary group of matrices relative to an inner product.

6. Tensor Product of Two Vector Spaces

If E is a vector space over $\mathbb{K}$, then the set of all bilinear forms on E is a vector space under addition and scalar multiplication of the values, i.e., it is a vector subspace of the set of all functions from $E \times E$ into $\mathbb{K}$. In this section we introduce a vector space called the "tensor product" of E with itself, whose dual, even if E is infinite-dimensional, is canonically isomorphic to this vector space of bilinear forms.

Matters will be clearer if we work initially with something slightly more general than bilinear forms on a single vector space E. Thus fix a field $\mathbb{K}$, and let E and F be vector spaces over $\mathbb{K}$. A function from $E \times F$ into a vector space U over $\mathbb{K}$ is said to be **bilinear** if it is linear in each of the two variables when the other one is held fixed. Such a space of bilinear functions is a vector space over $\mathbb{K}$ under addition and scalar multiplication of the values. The bilinear functions are called **bilinear forms** when the range space U is $\mathbb{K}$ itself. More generally, if $E_1, \dots, E_k$ are vector spaces over $\mathbb{K}$, a function from $E_1 \times \cdots \times E_k$ into a vector space over $\mathbb{K}$ is said to be k-**linear** or k-**multilinear** if it is linear in each of its k variables when the other $k-1$ variables are held fixed. Again the word "form" is used in the scalar-valued case, and all of these spaces of multilinear functions are vector spaces over $\mathbb{K}$.

In this section we shall introduce the tensor product of two vector spaces E and F over $\mathbb{K}$, ultimately denoting it by $E \otimes_{\mathbb{K}} F$. The dual of this tensor product will be canonically isomorphic to the vector space of bilinear forms on $E \times F$. More generally the space of linear functions from the tensor product into a vector space U will be canonically isomorphic to the vector space of bilinear functions on $E \times F$ with values in U.

Following the habit encouraged by Chapter IV, we want to arrange that tensor product is a functor. If $\mathcal{V}$ denotes the category of vector spaces over $\mathbb{K}$ and if $\mathcal{V} \times \mathcal{V}$ denotes the category described in Section IV.11 as $\mathcal{V}^S$ for a two-element set S, then tensor product is to be a functor from $\mathcal{V} \times \mathcal{V}$ into $\mathcal{V}$. Hence we will want to examine the effect of tensor products on morphisms, i.e., on linear maps.

As in similar constructions in Chapter IV, the effect of tensor product on linear maps is captured by defining the tensor product by means of a universal mapping property. The appropriate universal mapping property rephrases the statement above that the space of linear functions from the tensor product into any vector space U is canonically isomorphic to the vector space of bilinear functions on $E \times F$ with values in U.

If E and F are vector spaces over $\mathbb{K}$, a **tensor product** of E and F is a pair (V, ι) consisting of a vector space V over $\mathbb{K}$ together with a bilinear function $\iota : E \times F \to V$, with the following **universal mapping property**: whenever b is a bilinear mapping of $E \times F$ into a vector space U over $\mathbb{K}$, then there exists a unique linear mapping B of V into U such that the diagram in Figure 6.1 commutes, i.e., such that $B\iota = b$ holds in the diagram. When ι is understood, one frequently refers to V itself as the tensor product. The linear mapping $B : V \to U$ is called the **linear extension** of b to the tensor product.

$$\begin{array}{ccc} E \times F & \xrightarrow{\;b\;} & U \\ \iota \big\downarrow & \nearrow B & \\ V & & \end{array}$$

FIGURE 6.1. Universal mapping property of a tensor product.

Theorem 6.10. If E and F are vector spaces over $\mathbb{K}$, then a tensor product of E and F exists and is unique up to canonical isomorphism in this sense: if (V_1, ι_1) and (V_2, ι_2) are tensor products, then there exists a unique linear mapping $B : V_2 \to V_1$ with $B\iota_2 = \iota_1$, and B is an isomorphism. Any tensor product is spanned linearly by the image of $E \times F$ in it.

REMARKS. As usual, uniqueness will follow readily from the universal mapping property. What is really needed is a proof of existence. This will be carried out by an explicit construction. Later, in Chapter X, we shall reintroduce tensor products, taking the basic construction to be that of the tensor product of two abelian groups, and then the tensor product of two vector spaces will in effect be obtained in a slightly different way. However, the exact construction does not matter, only the existence; the uniqueness allows us to match the results of any two constructions.

$$\begin{array}{ccc} E \times F & \xrightarrow{\;\iota_2\;} & V_2 \\ \iota_1 \big\downarrow & \nearrow B_2 & \\ V_1 & & \end{array} \qquad \text{and} \qquad \begin{array}{ccc} E \times F & \xrightarrow{\;\iota_1\;} & V_1 \\ \iota_2 \big\downarrow & \nearrow B_1 & \\ V_2 & & \end{array}$$

FIGURE 6.2. Diagrams for uniqueness of a tensor product.

PROOF OF UNIQUENESS. Let (V_1, ι_1) and (V_2, ι_2) be tensor products. Set up the diagrams in Figure 6.2, and use the universal mapping property to obtain linear maps $B_2 : V_1 \to V_2$ and $B_1 : V_2 \to V_1$ extending ι_2 and ι_1. Then $B_1B_2 : V_1 \to V_1$ has $B_1B_2\iota_1 = B_1\iota_2 = \iota_1$, and $1_{V_1} : V_1 \to V_1$ has $(1_{V_1})\iota_1 = \iota_1$. By the assumed uniqueness within the universal mapping property, $B_1B_2 = 1_{V_1}$ on V_1. Similarly $B_2B_1 = 1_{V_2}$ on V_2. Then $B_1 : V_2 \to V_1$ gives the canonical isomorphism. Because of the isomorphism the image of $E \times F$ will span an arbitrary tensor product if it spans some particular tensor product. $\square$

PROOF OF EXISTENCE. Let $V_1 = \bigoplus_{(e,f)} \mathbb{K}(e, f)$, the direct sum being taken over all ordered pairs (e, f) with $e \in E$ and $f \in F$. Then V_1 is a vector space over $\mathbb{K}$ with a basis consisting of all ordered pairs (e, f). We think of all identities that the elements of V_1 must satisfy to be a tensor product, writing each as some expression set equal to 0, and then we assemble those expressions into a vector subspace to factor out from V_1. Namely, let V_0 be the vector subspace of V_1 generated by all elements of any of the kinds

$$\begin{gathered}(e_1 + e_2, f) - (e_1, f) - (e_2, f),\\ (ce, f) - c(e, f),\\ (e, f_1 + f_2) - (e, f_1) - (e, f_2),\\ (e, cf) - c(e, f),\end{gathered}$$

the understanding being that c is in $\mathbb{K}$, the elements e, e_1, e_2 are in E, and the elements f, f_1, f_2 are in F. Define $V = V_1/V_0$, and define $\iota : E \times F \to V_1/V_0$ by $\iota(e, f) = (e, f) + V_0$. We shall prove that (V, ι) is a tensor product of E and F. The definitions show that the image of ι spans V linearly.

Let $b : E \times F \to U$ be given as in Figure 6.1. To see that a linear extension B exists and is unique, define B_1 on V_1 by

$$B_1\Big(\sum_{\text{(finite)}} c_i(e_i, f_i)\Big) = \sum_{\text{(finite)}} c_i b(e_i, f_i).$$

The bilinearity of b shows that B_1 maps V_0 to 0. By Proposition 2.25, B_1 descends to a linear map $B : V_1/V_0 \to U$, and we have $B\iota = b$. Hence B exists as required.

To check uniqueness of B, we observe again that the cosets $(e, f) + V_0$ within V_1/V_0 span V; since commutativity of the diagram in Figure 6.1 forces

$$B((e, f) + V_0) = B(\iota(e, f)) = b(e, f),$$

B is unique. This completes the proof. □

A tensor product of E and F is denoted by $(E \otimes_{\mathbb{K}} F, \iota)$, with the bilinear map ι given by $\iota(e, f) = e \otimes f$; the map ι is frequently dropped from the notation when there is no chance of ambiguity. The tensor product that was constructed in the proof of existence in Theorem 6.10 is not given any special notation to distinguish it from any other tensor product. The elements $e \otimes f$ span $E \otimes_{\mathbb{K}} F$, as was noted in the statement of the theorem. Elements of the form $e \otimes f$ are sometimes called **pure tensors**.

Not every element need be a pure tensor, but every element in $E \otimes_{\mathbb{K}} F$ is a finite sum of pure tensors. We shall see in Proposition 6.14 that if $\{u_i\}$ is a basis of E and $\{v_j\}$ is a basis of F, then the pure tensors $u_i \otimes v_j$ form a basis of $E \otimes_{\mathbb{K}} F$. In particular the dimension of the tensor product is the product of the dimensions of the factors. We could have defined the tensor product in this way—by taking bases and declaring that $u_i \otimes v_j$ is to be a basis of the desired space. The difficulty is that we would be forever wedded to our choice of those particular bases, or we would constantly have to prove that our definitions are independent of bases. The definition by means of Theorem 6.10 avoids this difficulty.

To make tensor product $(E, F) \mapsto E \otimes_{\mathbb{K}} F$ into a functor, we have to describe the effect on linear mappings. To aid in that discussion, let us reintroduce some notation first used in Chapter II: if U and V are vector spaces over $\mathbb{K}$, then $\mathrm{Hom}_{\mathbb{K}}(U, V)$ is defined to be the vector space of $\mathbb{K}$ linear maps from U to V.

Corollary 6.11. If E, F, and V are vector spaces over $\mathbb{K}$, then the vector space $\mathrm{Hom}_{\mathbb{K}}(E \otimes_{\mathbb{K}} F, V)$ is canonically isomorphic (via restriction to pure tensors) to the vector space of all V-valued bilinear functions on $E \times F$.

PROOF. Restriction is a linear mapping from $\mathrm{Hom}_{\mathbb{K}}(E \otimes_{\mathbb{K}} F, V)$ to the vector space of all V-valued bilinear functions on $E \times F$, and it is one-one since the image of $E \times F$ in $E \otimes_{\mathbb{K}} F$ spans $E \otimes_{\mathbb{K}} F$. It is onto since any bilinear function from $E \times F$ to V has a linear extension to $E \otimes_{\mathbb{K}} F$, by Theorem 6.10. □

Corollary 6.12. If E and F are vector spaces over $\mathbb{K}$, then the vector space of all bilinear forms on $E \times F$ is canonically isomorphic to $(E \otimes_{\mathbb{K}} F)'$, the dual of the vector space $E \otimes_{\mathbb{K}} F$.

PROOF. This is the special case of Corollary 6.11 in which $V = \mathbb{K}$. □

Corollary 6.13. If E, F, and V are vector spaces over $\mathbb{K}$, then there is a canonical $\mathbb{K}$ linear isomorphism Φ of left side to right side in

$$\mathrm{Hom}_{\mathbb{K}}(E \otimes_{\mathbb{K}} F,\ V) \cong \mathrm{Hom}_{\mathbb{K}}(E,\ \mathrm{Hom}_{\mathbb{K}}(F, V))$$

such that

$$\Phi(\varphi)(e)(f) = \varphi(e \otimes f)$$

for all $\varphi \in \mathrm{Hom}_{\mathbb{K}}(E \otimes_{\mathbb{K}} F,\ V)$, $e \in E$, and $f \in F$.

REMARK. This result is just a restatement of Corollary 6.11, but let us prove it anyway, writing the proof in the language of the statement.

PROOF. The map Φ is well defined and $\mathbb{K}$ linear, and it carries the left side to the right side. For ψ in the right side, define $\Psi(\psi)(e, f) = \psi(e)(f)$. Then $\Psi(\psi)$ is a bilinear map from $E \times F$ into V, and we let $\widetilde{\Psi}(\psi)$ be the linear extension from $E \otimes_{\mathbb{K}} F$ into V given in Theorem 6.10. Then $\widetilde{\Psi}$ is a two-sided inverse to Φ, and the corollary follows. □

Let us now make $(E, F) \mapsto E \otimes_{\mathbb{K}} F$ into a covariant functor. If (E_1, F_1) and (E_2, F_2) are objects in $\mathcal{V} \times \mathcal{V}$, i.e., if they are two ordered pairs of vector spaces, then a morphism from the first to the second is a pair (L, M) of linear maps of the form $L : E_1 \to E_2$ and $M : F_1 \to F_2$. To (L, M), we are to associate a linear map from $E_1 \otimes_{\mathbb{K}} F_1$ into $E_2 \otimes_{\mathbb{K}} F_2$; this linear map will be denoted by $L \otimes M$. We use Corollary 6.11 to define $L \otimes M$ as the member of $\operatorname{Hom}_{\mathbb{K}}(E_1 \otimes_{\mathbb{K}} F_1, E_2 \otimes_{\mathbb{K}} F_2)$ that corresponds under restriction to the bilinear map $(e_1, f_1) \mapsto L(e_1) \otimes M(f_1)$ of $E_1 \times F_1$ into $E_2 \otimes_{\mathbb{K}} F_2$. In terms of pure tensors, the map $L \otimes M$ satisfies

$$(L \otimes M)(e_1 \otimes f_1) = L(e_1) \otimes M(f_1),$$

and this formula completely determines $L \otimes M$ because of the uniqueness of linear extensions of bilinear maps.

To check that this definition of the effect of tensor product on pairs of linear maps makes $(E, F) \mapsto E \otimes_{\mathbb{K}} F$ into a covariant functor, we have to check the effect on the identity map and the effect on composition. For the effect on the identity map $(1_{E_1}, 1_{F_1})$ when $E_1 = E_2$ and $F_1 = F_2$, we see from the above displayed formula that $(1_{E_1} \otimes 1_{F_1})(e_1 \otimes f_1) = 1_{E_1}(e_1) \otimes 1_{F_1}(f_1) = e_1 \otimes f_1 = 1_{E_1 \otimes_{\mathbb{K}} F_1}(e_1 \otimes f_1)$. Since elements of the form $e_1 \otimes f_1$ span $E_1 \otimes_{\mathbb{K}} F_1$, we conclude that $1_{E_1} \otimes 1_{F_1} = 1_{E_1 \otimes_{\mathbb{K}} F_1}$.

For the effect on composition, let $(L_1, M_1) : (E_1, F_1) \to (E_2, F_2)$ and $(L_2, M_2) : (E_2, F_2) \to (E_3, F_3)$ be given. Then we have

$$\begin{aligned}(L_2 \otimes M_2)(L_1 \otimes M_1)(e_1 \otimes f_1) &= (L_2 \otimes M_2)(L_1(e_1) \otimes M_1(f_1)) \\ &= (L_2L_1)(e_1) \otimes (M_2M_1)(f_1) = (L_2L_1 \otimes M_2M_1)(e_1 \otimes f_1).\end{aligned}$$

Since elements of the form $e_1 \otimes f_1$ span $E_1 \otimes_{\mathbb{K}} F_1$, we conclude that

$$(L_2 \otimes M_2)(L_1 \otimes M_1) = L_2L_1 \otimes M_2M_1.$$

Therefore $(E, F) \mapsto E \otimes_{\mathbb{K}} F$ is a covariant functor.

In particular, $E \mapsto E \otimes_{\mathbb{K}} F$ and $F \mapsto E \otimes_{\mathbb{K}} F$ are covariant functors from $\mathcal{V}$ into itself. For these two functors from $\mathcal{V}$ into itself, the effect on linear mappings is especially nice, namely that

$$L_1 \mapsto L_1 \otimes M_1 \qquad \begin{cases} \text{is } \mathbb{K} \text{ linear from } \operatorname{Hom}_{\mathbb{K}}(E_1, E_2) \\ \text{into } \operatorname{Hom}_{\mathbb{K}}(E_1 \otimes_{\mathbb{K}} F_1,\ E_2 \otimes_{\mathbb{K}} F_2), \end{cases}$$

$$M_1 \mapsto L_1 \otimes M_1 \qquad \begin{cases} \text{is } \mathbb{K} \text{ linear from } \operatorname{Hom}_{\mathbb{K}}(F_1, F_2) \\ \text{into } \operatorname{Hom}_{\mathbb{K}}(E_1 \otimes_{\mathbb{K}} F_1,\ E_2 \otimes_{\mathbb{K}} F_2). \end{cases}$$

To prove the first of these assertions, for example, we observe that the sum of the linear extensions of

$$(e_1, f_1) \mapsto L_1(e_1) \otimes M_1(f_1) \qquad \text{and} \qquad (e_1, f_1) \mapsto L_1'(e_1) \otimes M_1(f_1)$$

is a linear extension of $(e_1, f_1) \mapsto (L_1+L_1')(e_1)\otimes M_1(f_1)$, and the uniqueness in the universal mapping property implies that $(L_1+L_1')\otimes M_1 = L_1\otimes M_1+L_1'\otimes M_1$. Similar remarks apply to multiplication by scalars.

Let us mention some identities satisfied by $\otimes_{\mathbb{K}}$. There is a canonical isomorphism

$$E \otimes_{\mathbb{K}} F \cong F \otimes_{\mathbb{K}} E$$

given by taking the linear extension of $(e, f) \mapsto f \otimes e$ as the map from left to right. The linear extension of $(f, e) \mapsto e\otimes f$ gives a two-sided inverse. Category theory has a way of capturing the idea that this isomorphism is systematic, rather than randomly dependent on E and F. The two sides of the above isomorphism may be regarded as the values of the covariant functors $(E, F) \mapsto E \otimes_{\mathbb{K}} F$ and $(E, F) \mapsto F \otimes_{\mathbb{K}} E$. The notion in category theory capturing "systematic" is called "naturality." It makes precise the fact that the system of isomorphisms respects linear maps, as well as the vector spaces. Here is the general definition. Its usefulness will be examined later in this section.

Let $\mathcal{C}$ and $\mathcal{D}$ be two categories, and let $\Phi : \mathcal{C} \to \mathcal{D}$ and $\Psi : \mathcal{C} \to \mathcal{D}$ be covariant functors. Suppose that for each X in $\mathrm{Obj}(\mathcal{C})$, a morphism T_X in $\mathrm{Morph}_{\mathcal{D}}(\Phi(X), \Psi(X))$ is given. Then the system $\{T_X\}$ is called a **natural transformation** of Φ into Ψ if for each pair of objects X_1 and X_2 in $\mathcal{C}$ and each h in $\mathrm{Morph}_{\mathcal{C}}(X_1, X_2)$, the diagram in Figure 6.3 commutes. If furthermore each T_X is an isomorphism, then it is immediate that the system $\{T_X^{-1}\}$ is a natural transformation of Ψ into Φ, and we say that $\{T_X\}$ is a **natural isomorphism**.

$$\begin{array}{ccc} \Phi(X_1) & \xrightarrow{\Phi(h)} & \Phi(X_2) \\ {\scriptstyle T_{X_1}}\downarrow & & \downarrow{\scriptstyle T_{X_2}} \\ \Psi(X_1) & \xrightarrow{\Psi(h)} & \Psi(X_2) \end{array}$$

FIGURE 6.3. Commutative diagram of a natural transformation $\{T_X\}$.

If Φ and Ψ are contravariant functors, then the system $\{T_X\}$ is called a **natural transformation** of Φ into Ψ if the diagram obtained from Figure 6.3 by reversing the horizontal arrows commutes. The system is a **natural isomorphism** if furthermore each T_x is an isomorphism.

In the case we are studying, we have $\mathcal{C} = \mathcal{V}\times\mathcal{V}$ and $\mathcal{D} = \mathcal{V}$. Objects X in $\mathcal{C}$ are pairs (E, F) of vector spaces, and Φ and Ψ are the covariant functors with $\Phi(E, F) = E\otimes_{\mathbb{K}} F$ and $\Psi(E, F) = F\otimes_{\mathbb{K}} E$. The mapping $T_{(E,F)} : E\otimes_{\mathbb{K}} F \to F \otimes_{\mathbb{K}} E$ is uniquely determined by the condition that $T_{(E,F)}(e \otimes f) = f \otimes e$ for all $e \in E$ and $f \in F$. A morphism of pairs from (E_1, F_1) to (E_2, F_2) is of the form $h = (L, M)$ with $L \in \mathrm{Hom}_{\mathbb{K}}(E_1, E_2)$ and $M \in \mathrm{Hom}_{\mathbb{K}}(F_1, F_2)$. Our constructions above show that

$$\Phi(L, M) = L \otimes M \in \mathrm{Hom}_{\mathbb{K}}(E_1 \otimes_{\mathbb{K}} F_1, E_2 \otimes_{\mathbb{K}} F_2)$$

and
$$\Psi(L, M) = M \otimes L \in \mathrm{Hom}_{\mathbb{K}}(F_1 \otimes_{\mathbb{K}} E_1, F_2 \otimes_{\mathbb{K}} E_2).$$

In Figure 6.3 the two routes from top left to bottom right in the diagram have

$$\begin{aligned} T_{(E_2,F_2)}\Phi(L, M)(e_1 \otimes f_1) &= T_{(E_2,F_2)}(L \otimes M)(e_1 \otimes f_1) \\ &= T_{(E_2,F_2)}(L(e_1) \otimes M(f_1)) = M(f_1) \otimes L(e_1) \end{aligned}$$

and

$$\begin{aligned} \Psi(L, M)T_{(E_1,F_1)}(e_1 \otimes f_1) &= \Psi(L, M)(f_1 \otimes e_1) \\ &= (M \otimes L)(f_1 \otimes e_1) = M(f_1) \otimes L(e_1). \end{aligned}$$

The results are equal, and therefore the diagram commutes. Consequently the isomorphism

$$E \otimes_{\mathbb{K}} F \cong F \otimes_{\mathbb{K}} E$$

is natural in the pair (E, F).

Another canonical isomorphism of interest is

$$E \otimes_{\mathbb{K}} \mathbb{K} \cong E.$$

Here the map from left to right is the linear extension of $(e, c) \mapsto ce$, while the map from right to left is $e \mapsto e \otimes 1$. In view of the previous canonical isomorphism, we have $\mathbb{K} \otimes_{\mathbb{K}} E \cong E$ also. Each of these isomorphisms is natural in E.

Next let us consider how $\otimes_{\mathbb{K}}$ interacts with direct sums. The result is that tensor product distributes over direct sums, even infinite direct sums:

$$E \otimes_{\mathbb{K}} \Big(\bigoplus_{s\in S} F_s\Big) \cong \bigoplus_{s\in S} (E \otimes_{\mathbb{K}} F_s).$$

The map from left to right is the linear extension of the bilinear map $(e, \{f_s\}_{s\in S}) \mapsto \{e \otimes f_s\}_{s\in S}$. For the definition of the inverse, the constructions of Section II.6 show that we have only to define the map on each $E \otimes_{\mathbb{K}} F_s$, where it is the linear extension of $(e, f_s) \mapsto e \otimes \{i_s(f_s))\}_{s\in S}$; here $i_{s_0} : F_{s_0} \to \bigoplus_s F_s$ is the one-one linear map carrying the s_0^{th} vector space into the direct sum. Once again it is possible to prove that the isomorphism is natural; we omit the details.

It follows from the displayed isomorphism and the isomorphism $E \otimes_{\mathbb{K}} \mathbb{K} \cong E$ that if $\{x_i\}$ is a basis of E and $\{y_j\}$ is a basis of F, then $\{x_i \otimes y_j\}$ is a basis of $E \otimes_{\mathbb{K}} F$. This proves the following result.

Proposition 6.14. If E and F are vector spaces over $\mathbb{K}$, then

$$\dim(E \otimes_{\mathbb{K}} F) = (\dim E)(\dim F).$$

If $\{y_j\}$ is a basis of F, then the most general member of $E \otimes_{\mathbb{K}} F$ is of the form $\sum_j e_j \otimes y_j$ with all e_j in E.

We turn to a consideration of $\mathrm{Hom}_{\mathbb{K}}$ from the point of view of functors. In the examples in Section IV.11, we saw that $V \mapsto \mathrm{Hom}_{\mathbb{K}}(U, V)$ is a covariant functor from $\mathcal{V}$ to itself and that $U \mapsto \mathrm{Hom}_{\mathbb{K}}(U, V)$ is a contravariant functor from $\mathcal{V}$ to itself. If we are not squeamish about mixing the two types—covariant and contravariant—then we can consider $(U, V) \mapsto \mathrm{Hom}_{\mathbb{K}}(U, V)$ as a functor[3] from $\mathcal{V} \times \mathcal{V}$ to $\mathcal{V}$. At any rate if L is in $\mathrm{Hom}_{\mathbb{K}}(U_1, U_2)$ and M is in $\mathrm{Hom}_{\mathbb{K}}(V_1, V_2)$, then $\mathrm{Hom}(L, M)$ carries $\mathrm{Hom}_{\mathbb{K}}(U_2, V_1)$ into $\mathrm{Hom}_{\mathbb{K}}(U_1, V_2)$ and is given by

$$\mathrm{Hom}(L, M)(h) = MhL \qquad \text{for } h \in \mathrm{Hom}_{\mathbb{K}}(U_2, V_1).$$

It is evident that the result is $\mathbb{K}$ linear as a function of h, and hence

$$\mathrm{Hom}(L, M) \qquad \text{is in } \mathrm{Hom}_{\mathbb{K}}\big(\mathrm{Hom}_{\mathbb{K}}(U_2, V_1), \mathrm{Hom}_{\mathbb{K}}(U_1, V_2)\big).$$

When we look for analogs for the functor $\mathrm{Hom}_{\mathbb{K}}$ of the identity $E \otimes_{\mathbb{K}} \mathbb{K} \cong E$ for the functor $\otimes_{\mathbb{K}}$, we are led to two identities. One is just the definition of the dual of a vector space:

$$\mathrm{Hom}_{\mathbb{K}}(U, \mathbb{K}) = U'.$$

The other is the natural isomorphism

$$\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}, V) \cong V.$$

In the proof of the latter identity, the mapping from left to right is given by sending a linear $h : \mathbb{K} \to V$ to $h(1)$, and the mapping from right to left is given by sending v in V to h with $h(c) = cv$.

Next let us consider how $\mathrm{Hom}_{\mathbb{K}}$ interacts with direct sums and direct products. The construction $\mathrm{Hom}_{\mathbb{K}}(U, V)$ distributes over finite direct sums in each variable, but the situation with infinite direct sums or direct products is more subtle. Valid identities are

$$\mathrm{Hom}_{\mathbb{K}}\Big(\bigoplus_{s \in S} U_s,\ V\Big) \cong \prod_{s \in S} \mathrm{Hom}_{\mathbb{K}}(U_s, V)$$

and

$$\mathrm{Hom}_{\mathbb{K}}\Big(U, \prod_{s \in S} V_s\Big) \cong \prod_{s \in S} \mathrm{Hom}_{\mathbb{K}}(U, V_s),$$

[3] Readers who care about this point can regard U as in the category $\mathcal{V}^{\mathrm{opp}}$ defined in Problems 78–80 at the end of Chapter IV. Then $(U, V) \mapsto \mathrm{Hom}_{\mathbb{K}}(U, V)$ is a covariant functor from $\mathcal{V}^{\mathrm{opp}} \times \mathcal{V}$ into $\mathcal{V}$.

and these are natural isomorphisms. Proofs of these identities for all S and counterexamples related to them when S is infinite appear in Problems 7–8 at the end of the chapter.

We have already checked that the isomorphism $E \otimes_{\mathbb{K}} F \cong F \otimes_{\mathbb{K}} E$ is natural in (E, F), and we have asserted naturality in some other situations in which it is easy to check. The next proposition asserts naturality for the identity of Corollary 6.13, which combines $\otimes_{\mathbb{K}}$ and $\mathrm{Hom}_{\mathbb{K}}$ in a nontrivial way. After the proof of the result, we shall digress for a moment to indicate the usefulness of natural isomorphisms.

Proposition 6.15. Let E, F, V, E_1, F_1, and V_1 be vector spaces over $\mathbb{K}$, and let $L_{E_1} : E_1 \to E$, $L_{F_1} : F_1 \to F$, and $L_V : V \to V_1$ be $\mathbb{K}$ linear maps. Then the isomorphism Φ of Corollary 6.13 is natural in the sense that the diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbb{K}}(E \otimes_{\mathbb{K}} F,\ V) & \xrightarrow{\ \Phi\ } & \mathrm{Hom}_{\mathbb{K}}(E,\ \mathrm{Hom}_{\mathbb{K}}(F, V)) \\ {\scriptstyle \mathrm{Hom}(L_{E_1} \otimes L_{F_1},\, L_V)}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{Hom}(L_{E_1}, \mathrm{Hom}(L_{F_1}, L_V))} \\ \mathrm{Hom}_{\mathbb{K}}(E_1 \otimes_{\mathbb{K}} F_1,\ V_1) & \xrightarrow{\ \Phi\ } & \mathrm{Hom}_{\mathbb{K}}(E_1,\ \mathrm{Hom}_{\mathbb{K}}(F_1, V_1)) \end{array}$$

commutes.

REMARKS. Observe that the first two linear maps L_{E_1} and L_{F_1} go in the opposite direction to the two vertical maps, while L_V goes in the same direction as the vertical maps. This is a reflection of the fact that both sides of the identity in Corollary 6.13 are contravariant in the first two variables and covariant in the third variable.

PROOF. For φ in $\mathrm{Hom}_{\mathbb{K}}(E \otimes_{\mathbb{K}} F,\ V)$, e_1 in E_1, and f_1 in F_1, we have

$$\begin{aligned} (\mathrm{Hom}(L_{E_1}, \mathrm{Hom}(L_{F_1}, L_V)) \circ \Phi)(\varphi)(e_1)(f_1) \\ &= (\mathrm{Hom}(L_{F_1}, L_V) \circ \Phi(\varphi) \circ L_{E_1})(e_1)(f_1) \\ &= (\mathrm{Hom}(L_{F_1}, L_V) \circ (\Phi(\varphi) \circ L_{E_1}))(e_1)(f_1) \\ &= L_V(\Phi(\varphi)(L_{E_1}(e_1))(L_{F_1}(f_1))) \\ &= L_V(\varphi(L_{E_1}(e_1) \otimes L_{F_1}(f_1))) \\ &= (L_V \circ \varphi \circ (L_{E_1} \otimes L_{F_1}))(e_1 \otimes f_1) \\ &= (\mathrm{Hom}(L_{E_1} \otimes L_{F_1},\ L_V)(\varphi))(e_1 \otimes f_1) \\ &= \Phi(\mathrm{Hom}(L_{E_1} \otimes L_{F_1},\ L_V) \circ \varphi)(e_1)(f_1). \end{aligned}$$

This proves the proposition. $\square$

Let us now discuss naturality in a wider context. In a general category $\mathcal{D}$, if we have two objects U and U' such that $\mathrm{Morph}(U, V)$ and $\mathrm{Morph}(U', V)$ have the same cardinality for each object V, then we cannot really say anything about the relationship between U and U'. But under a hypothesis that the isomorphism of sets has a certain naturality to it, then, according to Proposition 6.16 below, U and U' are isomorphic objects. Thus naturality of a system of weak-looking set-theoretic isomorphisms can lead to a much stronger-looking isomorphism. Corollary 6.17 goes on to make a corresponding assertion about functors. The assertion about functors in the corollary is a helpful tool for establishing natural isomorphisms of functors, and an example appears below in Proposition 6.20′.

Proposition 6.16. Let $\mathcal{D}$ be a category, and suppose that U and U' are objects in $\mathcal{D}$ with the following property: to each object V in $\mathcal{D}$ corresponds a one-one onto function

$$T_V : \mathrm{Morph}(U, V) \to \mathrm{Morph}(U', V)$$

with the system $\{T_V\}$ natural in V in the sense that whenever σ is in $\mathrm{Morph}(V, V')$, then the diagram

$$\begin{array}{ccc} \mathrm{Morph}(U, V) & \xrightarrow{T_V} & \mathrm{Morph}(U', V) \\ {\scriptstyle \text{left-by-}\sigma}\downarrow & & \downarrow{\scriptstyle \text{left-by-}\sigma} \\ \mathrm{Morph}(U, V') & \xrightarrow{T_{V'}} & \mathrm{Morph}(U', V') \end{array}$$

commutes. Then U is isomorphic to U' as an object in $\mathcal{D}$, an isomorphism from U to U' being the member $T_{U'}^{-1}(1_{U'})$ of $\mathrm{Morph}(U, U')$.

REMARKS.

(1) Another way of formulating this result is as follows: Let $\mathcal{D}$ be any category, let $\mathcal{S}$ be the category of sets, and let U and U' be objects in $\mathcal{D}$. Define a covariant functor $H_U : \mathcal{D} \to \mathcal{S}$ by $H_U(V) = \mathrm{Morph}_{\mathcal{D}}(U, V)$ and $H_U(\sigma) = \text{left-by-}\sigma$ for $\sigma \in \mathrm{Morph}_{\mathcal{D}}(V, V')$, and define $H_{U'}$ similarly. If H_U and $H_{U'}$ are naturally isomorphic functors, then U and U' are isomorphic objects in $\mathcal{D}$.

(2) A similar result is valid when H_U and $H_{U'}$ are contravariant functors, H_U being defined by $H_U(V) = \mathrm{Hom}_{\mathcal{D}}(V, U)$ and $H_U(\sigma) = \text{right-by-}\sigma$ for $\sigma \in \mathrm{Morph}_{\mathcal{D}}(V, V')$. The result in this case follows immediately by applying Proposition 6.16 to the opposite category $\mathcal{D}^{\mathrm{opp}}$ of $\mathcal{D}$ as defined in Problems 78–80 at the end of Chapter IV.

PROOF. Let φ be the element $T_{U'}^{-1}(1_{U'})$ of $\mathrm{Morph}(U, U')$, and let ψ be the element $T_U(1_U)$ of $\mathrm{Morph}(U', U)$. To prove the proposition, it is enough to show that $\varphi\psi = 1_{U'}$ and $\psi\varphi = 1_U$.

For σ in $\text{Morph}(V, V')$, form the commutative diagram in the statement of the proposition. The commutativity says that

$$\sigma T_V(h) = T_{V'}(\sigma h) \qquad \text{for } h \in \text{Morph}(U, V). \tag{$*$}$$

Taking $V = U$, $V' = U'$, $\sigma = \varphi$, and $h = 1_U$ in $(*)$ proves the second equality of the chain

$$\varphi\psi = \varphi T_U(1_U) = T_{U'}(\varphi 1_U) = T_{U'}(\varphi) = 1_{U'}.$$

Taking $V = U'$, $V' = U$, $\sigma = \psi$, and $h = \varphi$ in $(*)$ proves the first equality of the chain

$$T_U(\psi\varphi) = \psi T_{U'}(\varphi) = \psi 1_{U'} = \psi = T_U(1_U);$$

Applying T_U^{-1}, we obtain $\psi\varphi = 1_U$, as required. □

Corollary 6.17. Let $\mathcal{C}$ and $\mathcal{D}$ be categories, and let $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{C} \to \mathcal{D}$ be covariant functors. Suppose that to each pair of objects (A, V) in $\mathcal{C} \times \mathcal{D}$ corresponds a one-one onto function

$$T_{A,V} : \text{Morph}(F(A), V) \to \text{Morph}(G(A), V)$$

with the system $\{T_{A,V}\}$ natural in (A, V). Then the functors F and G are naturally isomorphic.

REMARKS. A similar result is valid if $T_{A,V}$ carries $\text{Morph}(V, F(A))$ to $\text{Morph}(V, G(A))$ and/or if F and G are contravariant. To handle these situations, we apply the corollary to the opposite categories $\mathcal{D}^{\text{opp}}$ and/or $\mathcal{C}^{\text{opp}}$, as defined in Problems 78–80 at the end of Chapter IV, instead of to the categories $\mathcal{D}$ and/or $\mathcal{C}$.

PROOF. By Proposition 6.16 and the hypotheses, the member $T_{A,G(A)}^{-1}(1_{G(A)})$ of $\text{Morph}_{\mathcal{D}}(F(A), G(A))$ is an isomorphism. We are to prove that the system $\{T_{A,G(A)}\}$ is natural in A. If σ in $\text{Morph}_{\mathcal{C}}(A, A')$ is given, then the naturality of $T_{A,V}$ in the V variable implies that the diagram

$$\begin{array}{ccc}
\text{Morph}_{\mathcal{D}}(F(A), G(A)) & \xrightarrow{T_{A,G(A)}} & \text{Morph}_{\mathcal{D}}(G(A), G(A)) \\
{\scriptstyle \text{left-by-}G(\sigma)}\downarrow & & \downarrow{\scriptstyle \text{left-by-}G(\sigma)} \\
\text{Morph}_{\mathcal{D}}(F(A), G(A')) & \xrightarrow{T_{A,G(A')}} & \text{Morph}_{\mathcal{D}}(G(A), G(A'))
\end{array}$$

commutes. Evaluating at $T_{A,G(A)}^{-1}(1_{G(A)}) \in \text{Morph}_{\mathcal{D}}(F(A), G(A))$ the two equal compositions in the diagram, we obtain

$$G(\sigma) = G(\sigma)1_{G(A)} = T_{A,G(A')}\big(G(\sigma)T_{A,G(A)}^{-1}(1_{G(A)})\big). \tag{$*$}$$

With σ as above, the naturality of $T_{A,V}$ in the A variable implies that the diagram

$$\begin{array}{ccc} \mathrm{Morph}_{\mathcal{D}}(F(A'), G(A')) & \xrightarrow{T_{A',G(A')}} & \mathrm{Morph}_{\mathcal{D}}(G(A'), G(A')) \\ {\scriptstyle \text{right-by-}F(\sigma)}\big\downarrow & & \big\downarrow{\scriptstyle \text{right-by-}G(\sigma)} \\ \mathrm{Morph}_{\mathcal{D}}(F(A), G(A')) & \xrightarrow{T_{A,G(A')}} & \mathrm{Morph}_{\mathcal{D}}(G(A), G(A')) \end{array}$$

commutes. Evaluating at $T^{-1}_{A',G(A')}(1_{G(A')}) \in \mathrm{Morph}_{\mathcal{D}}(F(A'), G(A'))$ the two equal compositions in the diagram, we obtain

$$G(\sigma) = 1_{G(A')}G(\sigma) = T_{A,G(A')}\big(T^{-1}_{A',G(A')}(1_{G(A')})F(\sigma)\big). \qquad (**)$$

Equations $(*)$ and $(**)$, together with the fact that $T_{A,G(A')}$ is invertible, say that

$$G(\sigma)T^{-1}_{A,G(A)}(1_{G(A)}) = T^{-1}_{A',G(A')}(1_{G(A')})F(\sigma).$$

In other words, the isomorphism $\widetilde{T}_A \in \mathrm{Morph}_{\mathcal{D}}(F(A), G(A))$ given by $\widetilde{T}_A = T^{-1}_{A,G(A)}(1_{G(A)})$ makes the diagram

$$\begin{array}{ccc} F(A) & \xrightarrow{\widetilde{T}_A} & G(A) \\ {\scriptstyle F(\sigma)}\big\downarrow & & \big\downarrow{\scriptstyle G(\sigma)} \\ F(A') & \xrightarrow{\widetilde{T}_{A'}} & G(A') \end{array}$$

commute. Thus F is naturally isomorphic to G. □

Tensor product provides a device for converting a real vector space canonically into a complex vector space, so that a basis over $\mathbb{R}$ in the original space becomes a basis over $\mathbb{C}$ in the new space. If E is the given real vector space, then the complex vector space, called the **complexification** of E, is the space $E^{\mathbb{C}} = E \otimes_{\mathbb{R}} \mathbb{C}$ with multiplication by a complex number c in $E^{\mathbb{C}}$ defined to be $1 \otimes (z \mapsto cz)$.

This construction works more generally when we have any inclusion of fields $\mathbb{K} \subseteq \mathbb{L}$. In this situation, $\mathbb{L}$ becomes a vector space over $\mathbb{K}$ if scalar multiplication $\mathbb{K} \times \mathbb{L} \to \mathbb{L}$ is defined as the restriction of the multiplication $\mathbb{L} \times \mathbb{L} \to \mathbb{L}$ within $\mathbb{L}$. For any vector space E over $\mathbb{K}$, we define $E^{\mathbb{L}} = E \otimes_{\mathbb{K}} \mathbb{L}$, initially as a vector space over $\mathbb{K}$. For $c \in \mathbb{L}$, we then define

$$(\text{multiplication by } c \text{ in } E \otimes_{\mathbb{K}} \mathbb{L}) = 1 \otimes (\text{multiplication by } c \text{ in } \mathbb{L}).$$

The above identities concerning tensor products of linear maps allow one easily to prove the following identities:

$$\begin{aligned} c_1(c_2v) &= (c_1c_2)v, \\ c(u+v) &= cu+cv, \\ (c_1+c_2)v &= c_1v+c_2v, \\ 1v &= v. \end{aligned}$$

Together these identities say that $E^{\mathbb{L}} = E \otimes_{\mathbb{K}} \mathbb{L}$, with its vector-space addition and the above definition of multiplication by scalars in $\mathbb{L}$, is a vector space over $\mathbb{L}$. The further identity

$$c(e \otimes 1) = ce \otimes 1 \qquad \text{if } c \text{ is in } \mathbb{K} \text{ and } e \text{ is in } E$$

shows that its scalar multiplication is consistent with scalar multiplication in E when the scalars are in $\mathbb{K}$ and E is identified with the subset $E \otimes 1$ of $E^{\mathbb{L}}$.

Let us say that the pair $(E^{\mathbb{L}}, \iota)$, where $\iota : E \to E^{\mathbb{L}}$ is the mapping $e \mapsto e \otimes 1$, is obtained by **extension of scalars**. This construction is characterized by a universal mapping property as follows.

Proposition 6.18. Let $\mathbb{K} \subseteq \mathbb{L}$ be an inclusion of fields, and let E be a vector space over $\mathbb{K}$.

(a) If $(E^{\mathbb{L}}, \iota)$ is formed by extension of scalars, then $(E^{\mathbb{L}}, \iota)$ has the following universal mapping property: whenever U is a vector space over $\mathbb{L}$ and $\varphi : E \to U$ is a $\mathbb{K}$ linear map, there exists a unique $\mathbb{L}$ linear map $\Phi : E^{\mathbb{L}} \to U$ such that $\Phi\,\iota = \varphi$.

(b) Suppose that (V, j) is any pair in which V is a vector space over $\mathbb{L}$ and $j : E \to V$ is a $\mathbb{K}$ linear function such that the following universal mapping property holds: whenever U is a vector space over $\mathbb{L}$ and $\varphi : E \to U$ is a $\mathbb{K}$ linear map, there exists a unique $\mathbb{L}$ linear map $\Phi : V \to U$ such that $\Phi j = \varphi$. Then there exists a unique isomorphism $\Psi : E^{\mathbb{L}} \to V$ of $\mathbb{L}$ vector spaces such that $\Psi\,\iota = j$.

PROOF. In (a), for the uniqueness of Φ, we must have $\Phi(e \otimes c) = c\Phi(e \otimes 1) = c(\Phi\,\iota)(e) = c\varphi(e)$. Hence Φ is determined by φ on pure tensors in $E \otimes_{\mathbb{K}} \mathbb{L}$ and therefore everywhere.

For existence let $\Phi : E \otimes_{\mathbb{K}} \mathbb{L} \to U$ be the $\mathbb{K}$ linear extension of the $\mathbb{K}$ bilinear function of $E \times \mathbb{L}$ into U given by

$$(e, c) \mapsto c\varphi(e) \qquad \text{for } e \in E \text{ and } c \in \mathbb{L}.$$

In the $\mathbb{L}$ vector space $E \otimes_{\mathbb{K}} \mathbb{L}$, multiplication by a member c_0 of $\mathbb{L}$ is defined to be $1 \otimes$ (multiplication by c_0). On a pure tensor $e \otimes c$, we therefore have

$$\Phi(c_0(e \otimes c)) = \Phi(e \otimes c_0c) = (c_0c)\varphi(e) = c_0(c\varphi(e)) = c_0(\Phi(e \otimes c)).$$

Since $E \otimes_{\mathbb{K}} \mathbb{L}$ is generated by pure tensors, Φ is $\mathbb{L}$ linear. By the construction of Φ, $\varphi(e) = \Phi(e \otimes 1) = (\Phi\,\iota)(e)$. Thus Φ has the required properties.

In (b), let (V, j) have the same universal mapping property as $(E^{\mathbb{L}}, \iota)$. We apply the universal mapping property of $(E^{\mathbb{L}}, \iota)$ to the $\mathbb{K}$ linear map $j : E \to V$ to obtain an $\mathbb{L}$ linear $\Phi : E^{\mathbb{L}} \to V$ with $\Phi\,\iota = j$, and we apply the universal mapping property of (V, j) to the $\mathbb{K}$ linear map $\iota : E \to E^{\mathbb{L}}$ to obtain an $\mathbb{L}$ linear $\Phi' : V \to E^{\mathbb{L}}$ with $\Phi' j = \iota$. From $(\Phi'\Phi)\,\iota = \Phi' j = \iota$ and $1_{E^{\mathbb{L}}}\,\iota = \iota$, the uniqueness in the universal mapping property for $(E^{\mathbb{L}}, \iota)$ implies $\Phi'\Phi = 1_{E^{\mathbb{L}}}$. Arguing similarly, we obtain $\Phi\Phi' = 1_V$. Thus Φ is an isomorphism with the required properties.

If $\Psi : E^{\mathbb{L}} \to V$ is another isomorphism with $\Psi\,\iota = j$, then the argument just given shows that $\Phi'\Psi = 1_{E^{\mathbb{L}}}$ and $\Psi\Phi' = 1_V$. Hence $\Psi = (\Phi')^{-1} = \Phi$, and Ψ is unique. □

To make $E \mapsto E^{\mathbb{L}}$ into a covariant functor from vector spaces over $\mathbb{K}$ to vector spaces over $\mathbb{L}$, we must examine the effect on linear maps. The tool is Proposition 6.18a. Thus let E and F be two vector spaces over $\mathbb{K}$, and let $M : E \to F$ be a $\mathbb{K}$ linear map between them. We extend scalars for E and F. The proposition applies to the composition $E \to F \to F^{\mathbb{L}}$ and shows that the composition extends uniquely to an $\mathbb{L}$ linear map from $E^{\mathbb{L}}$ to $F^{\mathbb{L}}$. A quick look at the proof shows that this $\mathbb{L}$ linear map is $M \otimes 1$. Actually, we can see directly that $M \otimes 1$ is indeed linear over $\mathbb{L}$ and not just over $\mathbb{K}$: we just use our identity for compositions of tensor products to write

$$\begin{aligned}(M \otimes 1)(I \otimes \text{(multiplication by } c)) &= M \otimes \text{(multiplication by } c) \\ &= (I \otimes \text{(multiplication by } c))(M \otimes 1).\end{aligned}$$

In any event, the explicit form of the extended linear map as $M \otimes 1$ shows immediately that the identity linear map goes to the identity and that compositions go to compositions. Thus $E \mapsto E^{\mathbb{L}}$ is a covariant functor.

In the special case that the vector spaces are $\mathbb{K}^n$ and $\mathbb{K}^m$, extension of scalars has a particularly simple interpretation. The new spaces may be viewed as $\mathbb{L}^n$ and $\mathbb{L}^m$. Thus column vectors with entries in $\mathbb{K}$ get replaced by column vectors with entries in $\mathbb{L}$. What happens with linear mappings is even more transparent. A linear map $M : E \to F$ is given by an m-by-n matrix A with entries in $\mathbb{K}$, and the linear map $M \otimes 1 : E^{\mathbb{L}} \to F^{\mathbb{L}}$ is the one given by the same matrix A. Now the entries of A are to be regarded as members of the larger field $\mathbb{L}$. Viewed this way, extension of scalars might look as if it is dependent on choices of bases, but the tensor-product formalism shows that it is not.

A related notion to extension of scalars is that of **restriction of scalars**. Again with an inclusion $\mathbb{K} \subseteq \mathbb{L}$ of fields, a vector space E over the larger field $\mathbb{L}$ becomes a vector space $E_{\mathbb{K}}$ over the smaller field $\mathbb{K}$ by ignoring unnecessary scalar multiplications. Although this notion is related to extension of scalars, it is not inverse to it. For example, if the two fields are $\mathbb{R}$ and $\mathbb{C}$ and if we start with an n-dimensional vector space E over $\mathbb{R}$, then $E^{\mathbb{C}}$ is a complex vector space of dimension n and $(E^{\mathbb{C}})_{\mathbb{R}}$ is a real vector space of dimension $2n$. We thus do not get back to the original space E.

7. Tensor Algebra

Just as polynomial rings are often used in the construction of more general commutative rings, so "tensor algebras" are often used in the construction of more general rings that may not be commutative. In this section we construct the "tensor algebra" of a vector space as a direct sum of iterated tensor products of the vector space with itself, and we establish its properties. We shall proceed with care, in order to provide a complete proof of the associativity of the multiplication.

Let A, B, and C be vector spaces over a field $\mathbb{K}$. A **triple tensor product** $V = A \otimes_{\mathbb{K}} B \otimes_{\mathbb{K}} C$ is a vector space over $\mathbb{K}$ with a 3-linear map $\iota : A \times B \times C \to V$ having the following universal mapping property: whenever t is a 3-linear mapping of $A \times B \times C$ into a vector space U over $\mathbb{K}$, then there exists a linear mapping T of V into U such that the diagram in Figure 6.4 commutes.

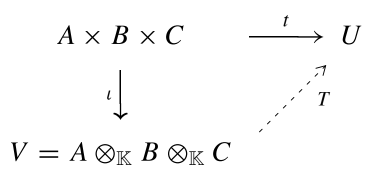


FIGURE 6.4. Commutative diagram of a triple tensor product.

The usual argument with universal mapping properties shows that there is at most one triple tensor product up to a well-determined isomorphism, and one can give an explicit construction of it that is similar to the one for ordinary tensor products $E \otimes_{\mathbb{K}} F$. We shall not need that particular proof of existence since Proposition 6.19a below will give us an alternative argument. Once we have that statement, we shall use the uniqueness of triple tensor products to establish in Proposition 6.19b an associativity formula for ordinary iterated tensor products.

A shorter proof of Proposition 6.19b, which avoids Proposition 6.19a and uses naturality, will be given after the proof of Proposition 6.20.

Proposition 6.19. If $\mathbb{K}$ is a field and A, B, C are vector spaces over $\mathbb{K}$, then

(a) $(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$ and $A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C)$ are triple tensor products.

(b) there exists a unique $\mathbb{K}$ isomorphism Φ from left to right in

$$(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C \cong A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C)$$

such that $\Phi((a \otimes b) \otimes c) = a \otimes (b \otimes c)$ for all $a \in A$, $b \in B$, and $c \in C$.

PROOF. In (a), consider $(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$. Let $t : A \times B \times C \to U$ be 3-linear. For $c \in C$, define $t_c : A \times B \to U$ by $t_c(a, b) = t(a, b, c)$. Then t_c is bilinear and hence extends to a linear $T_c : A \otimes_{\mathbb{K}} B \to U$. Since t is 3-linear, $t_{c_1+c_2} = t_{c_1} + t_{c_2}$ and $t_{xc} = xt_c$ for scalar x; thus uniqueness of the linear extension forces $T_{c_1+c_2} = T_{c_1} + T_{c_2}$ and $T_{xc} = xT_c$. Consequently

$$t' : (A \otimes_{\mathbb{K}} B) \times C \to U$$

given by $t'(d, c) = T_c(d)$ is bilinear and therefore extends to a linear $T : (A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C \to U$. This T proves existence of the linear extension of the given t. Uniqueness is trivial, since the elements $(a \otimes b) \otimes c$ span $(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$. So $(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$ is a triple tensor product. In a similar fashion, $A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C)$ is a triple tensor product.

For (b), set up the diagram of the universal mapping property for a triple tensor product, using $V = (A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$, $U = A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C)$, and $t(a, b, c) = a \otimes (b \otimes c)$. We have just seen in (a) that V is a triple tensor product with $\iota(a, b, c) = (a \otimes b) \otimes c$. Thus there exists a linear $T : V \to U$ with $T\iota(a, b, c) = t(a, b, c)$. This equation means that $T((a \otimes b) \otimes c) = a \otimes (b \otimes c)$. Interchanging the roles of $(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$ and $A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C)$, we obtain a two-sided inverse for T. Thus T will serve as Φ in (b), and existence is proved. Uniqueness is trivial, since the elements $(a \otimes b) \otimes c$ span $(A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C$. □

When there is no danger of confusion, Proposition 6.19 allows us to write a triple tensor product without parentheses as $A \otimes_{\mathbb{K}} B \otimes_{\mathbb{K}} C$. The same argument as in Corollaries 6.11 and 6.12 shows that the vector space of 3-linear forms on $A \times B \times C$ is canonically isomorphic to the dual of the vector space $A \otimes_{\mathbb{K}} B \otimes_{\mathbb{K}} C$.

Just as with Corollary 6.13 and Proposition 6.15, the result of Proposition 6.19 can be improved by saying that the isomorphism is natural in the variables A, B, and C, as follows.

Proposition 6.20. Let A, B, C, A_1, B_1, and C_1 be vector spaces over a field $\mathbb{K}$, and let $L_A : A \to A_1$, $L_B : B \to B_1$, and $L_C : C \to C_1$ be linear maps. Then the isomorphism Φ of Proposition 6.19b is natural in the triple (A, B, C) in the sense that the diagram

$$\begin{array}{ccc} (A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C & \xrightarrow{\Phi} & A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C) \\ {\scriptstyle (L_A \otimes L_B) \otimes L_C} \downarrow & & \downarrow {\scriptstyle L_A \otimes (L_B \otimes L_C)} \\ (A_1 \otimes_{\mathbb{K}} B_1) \otimes_{\mathbb{K}} C_1 & \xrightarrow{\Phi} & A_1 \otimes_{\mathbb{K}} (B_1 \otimes_{\mathbb{K}} C_1) \end{array}$$

commutes.

PROOF. We have

$$\begin{aligned} ((L_A \otimes (L_B \otimes L_C)) \circ \Phi)((a \otimes b) \otimes c) \\ &= (L_A \otimes (L_B \otimes L_C))(a \otimes (b \otimes c)) \\ &= L_A a \otimes (L_B \otimes L_C)(b \otimes c) \\ &= L_A a \otimes (L_B b \otimes L_C c) \\ &= \Phi((L_A a \otimes L_B b) \otimes L_C c) \\ &= \Phi((L_A \otimes L_B)(a \otimes b) \otimes L_C c) \\ &= (\Phi \circ ((L_A \otimes L_B) \otimes L_C))((a \otimes b) \otimes c), \end{aligned}$$

and the proposition follows. □

The treatment of Propositions 6.19 and 6.20 can be shortened if we are willing to bypass the notion of a triple tensor product and use what was proved about naturality in the previous section. The result and the proof are as follows.

Proposition 6.20′. Let A, B, and C be vector spaces over a field $\mathbb{K}$. Then there is an isomorphism $\Phi : (A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C \to A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C)$ that is natural in the triple (A, B, C) and satisfies $\Phi(a \otimes (b \otimes c)) = a \otimes (b \otimes c)$.

PROOF. Writing $\cong$ for "naturally isomorphic in all variables" and applying Proposition 6.15 and other natural isomorphisms of the previous section repeatedly, we have

$$\begin{aligned} \mathrm{Hom}_{\mathbb{K}}\big((A \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} C, V\big) &\cong \mathrm{Hom}_{\mathbb{K}}\big(A \otimes_{\mathbb{K}} B,\ \mathrm{Hom}_{\mathbb{K}}(C, V)\big) \\ &\cong \mathrm{Hom}_{\mathbb{K}}\big(B,\ \mathrm{Hom}_{\mathbb{K}}(A, \mathrm{Hom}_{\mathbb{K}}(C, V))\big) \\ &\cong \mathrm{Hom}_{\mathbb{K}}\big(B,\ \mathrm{Hom}_{\mathbb{K}}(A \otimes_{\mathbb{K}} C, V)\big) \\ &\cong \mathrm{Hom}_{\mathbb{K}}\big(B,\ \mathrm{Hom}_{\mathbb{K}}(C \otimes_{\mathbb{K}} A, V)\big) \\ &\cong \mathrm{Hom}_{\mathbb{K}}\big((C \otimes_{\mathbb{K}} B) \otimes_{\mathbb{K}} A, V\big) \quad \text{by symmetry} \\ &\cong \mathrm{Hom}_{\mathbb{K}}\big(A \otimes_{\mathbb{K}} (C \otimes_{\mathbb{K}} B), V\big) \\ &\cong \mathrm{Hom}_{\mathbb{K}}\big(A \otimes_{\mathbb{K}} (B \otimes_{\mathbb{K}} C), V\big). \end{aligned}$$

Then the existence of the natural isomorphism follows from Corollary 6.17. Using the explicit formula for the isomorphism in Proposition 6.16 and tracking matters down, we see that $\Phi(a \otimes (b \otimes c)) = a \otimes (b \otimes c)$. □

There is no difficulty in generalizing matters to n-fold tensor products by induction. An n-**fold tensor product** is to be universal for n-multilinear maps. Again it is unique up to canonical isomorphism, as one proves by an argument that runs along familiar lines. A direct construction of an n-fold tensor product is possible in the style of the proof for ordinary tensor products, but such a construction will not be needed. Instead, we can form an n-fold tensor product as the $(n-1)$-fold tensor product of the first $n-1$ spaces, tensored with the n^{th} space. Proposition 6.19b allows us to regroup parentheses (inductively) in any fashion we choose, and the same argument as in Corollaries 6.11 and 6.12 yields the following proposition.

Proposition 6.21. If $E_1, \dots, E_n$, and V are vector spaces over $\mathbb{K}$, then the vector space $\operatorname{Hom}_{\mathbb{K}}(E_1 \otimes_{\mathbb{K}} \cdots \otimes_{\mathbb{K}} E_n,\ V)$ is canonically isomorphic (via restriction to pure tensors) to the vector space of all V-valued n-multilinear functions on $E_1 \times \cdots \times E_n$. In particular the vector space of all n-multilinear forms on $E_1 \times \cdots \times E_n$ is canonically isomorphic to $(E_1 \otimes_{\mathbb{K}} \cdots \otimes_{\mathbb{K}} E_n)'$.

Iterated application of Proposition 6.20 shows that we get also a well-defined notion of a linear map $L_1 \otimes \cdots \otimes L_n$, the tensor product of n linear maps. Thus $(E_1, \dots, E_n) \mapsto E_1 \otimes_{\mathbb{K}} \cdots \otimes_{\mathbb{K}} E_n$ is a functor. There is no need to write out the details.

We turn to the question of defining a multiplication operation on tensors. If $\mathbb{K}$ is a field, an **algebra**[4] over $\mathbb{K}$ is a vector space V over $\mathbb{K}$ with a **multiplication** or **product** operation $V \times V \to V$ that is $\mathbb{K}$ bilinear. The additive part of the $\mathbb{K}$ bilinearity means that the product operation satisfies the distributive laws

$$a(b+c) = ab + ac \qquad \text{and} \qquad (b+c)a = ba + ca \qquad \text{for all } a, b, c \text{ in } V,$$

and the scalar-multiplication part of the $\mathbb{K}$ bilinearity means that

$$(ka)b = k(ab) = a(kb) \qquad \text{for all } k \text{ in } \mathbb{K} \text{ and } a, b \text{ in } V.$$

Within the text of the book, we shall work mostly just with **associative algebras**, i.e., those algebras satisfying the usual associative law

$$a(bc) = (ab)c \qquad \text{for all } a, b, c \text{ in } V.$$

[4]Some authors use the term "algebra" to mean what we shall call an "associative algebra."

An associative algebra is therefore a ring and a vector space, the scalar multiplication and the ring multiplication being linked by the requirement that $(ka)b = k(ab) = a(kb)$ for all scalars k. Some commutative examples of associative algebras over $\mathbb{K}$ are any field $\mathbb{L}$ containing $\mathbb{K}$, the polynomial algebra $\mathbb{K}[X_1, \dots, X_n]$, and the algebra of all $\mathbb{K}$-valued functions on a nonempty set S. Two noncommutative examples of associative algebras over $\mathbb{K}$ are the matrix algebra $M_n(\mathbb{K})$, with matrix multiplication as its product, and $\operatorname{Hom}_{\mathbb{K}}(V, V)$ for any vector space V, with composition as its product. The division ring $\mathbb{H}$ of quaternions (Example 10 in Section IV.1) is another example of a noncommutative associative algebra over $\mathbb{R}$.

Despite our emphasis on algebras that are associative, certain kinds of nonassociative algebras are of great importance in applications, and consequently several problems at the end of the chapter make use of nonassociative algebras. A nonassociative algebra is determined by its vector-space structure and the multiplication table for the members of a $\mathbb{K}$ basis. There is no restriction on the multiplication table; all multiplication tables define algebras. Perhaps the best-known nonassociative algebra is the 3-dimensional algebra over $\mathbb{R}$ determined by **vector product** in $\mathbb{R}^3$. A basis is $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, the multiplication operation is denoted by $\times$, and the multiplication table is

$$\begin{array}{lll} \mathbf{i}\times\mathbf{i}=0, & \mathbf{i}\times\mathbf{j}=\mathbf{k}, & \mathbf{i}\times\mathbf{k}=-\mathbf{j}, \\ \mathbf{j}\times\mathbf{i}=-\mathbf{k}, & \mathbf{j}\times\mathbf{j}=0, & \mathbf{j}\times\mathbf{k}=\mathbf{i}, \\ \mathbf{k}\times\mathbf{i}=\mathbf{j}, & \mathbf{k}\times\mathbf{j}=-\mathbf{i}, & \mathbf{k}\times\mathbf{k}=0. \end{array}$$

Since $\mathbf{i}\times(\mathbf{i}\times\mathbf{k}) = \mathbf{i}\times(-\mathbf{j}) = -\mathbf{k}$ and $(\mathbf{i}\times\mathbf{i})\times\mathbf{k} = 0$, vector product is not associative. The vector-product algebra is a special case of a Lie algebra; Lie algebras are defined in Problems 31–35 at the end of the chapter.

Tensor algebras, which we shall now construct, will be associative algebras. Fix a vector space E over $\mathbb{K}$, and for integers $n \geq 1$, let $T^n(E)$ be the n-fold tensor product of E with itself. In the case $n = 0$, we let $T^0(E)$ be the field $\mathbb{K}$. Define, initially as a vector space, $T(E)$ to be the direct sum

$$T(E) = \bigoplus_{n=0}^{\infty} T^n(E).$$

The elements that lie in one or another $T^n(E)$ are called **homogeneous**. We define a bilinear multiplication on homogeneous elements

$$T^m(E) \times T^n(E) \to T^{m+n}(E)$$

to be the restriction of the canonical isomorphism

$$T^m(E) \otimes_{\mathbb{K}} T^n(E) \to T^{m+n}(E)$$

resulting from iterating Proposition 6.19b. This multiplication, denoted by $\otimes$, is associative, as far as it goes, because the restriction of the $\mathbb{K}$ isomorphism

$$T^l(E) \otimes_{\mathbb{K}} (T^m(E) \otimes_{\mathbb{K}} T^n(E)) \to (T^l(E) \otimes_{\mathbb{K}} T^m(E)) \otimes_{\mathbb{K}} T^n(E)$$

to $T^l(E) \times (T^m(E) \times T^n(E))$ factors through the map

$$T^l(E) \times (T^m(E) \times T^n(E)) \to (T^l(E) \times T^m(E)) \times T^n(E)$$

given by $(r, (s,t)) \mapsto ((r,s),t)$.

This much tells how to multiply homogeneous elements in $T(E)$. Since each element t in $T(E)$ has a unique expansion as a finite sum $t = \sum_{k=0}^{n} t_k$ with $t_k \in T^k(E)$, we can define the product of this t and the element $t' = \sum_{k=0}^{n'} t'_k$ to be the element $t \otimes t' = \sum_{l=0}^{n+n'} \sum_{k+k'=l} (t_k \otimes t'_k)$; the expression $\sum_{k+k'=l} (t_k \otimes t'_k)$ is the component of the product in $T^l(E)$.

Multiplication is thereby well defined in $T(E)$, and it satisfies the distributive laws and is associative. Thus $T(E)$ becomes an associative algebra with a (two-sided) identity, namely the element 1 in $T^0(E)$. In the presence of the identification $\iota : E \to T^1(E)$, $T(E)$ is known as the **tensor algebra** of E. The pair $(T(E), \iota)$ has the **universal mapping property** given in Proposition 6.22 and pictured in Figure 6.5.

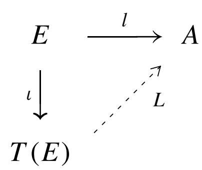


FIGURE 6.5. University mapping property of a tensor algebra.

Proposition 6.22. The pair $(T(E), \iota)$ has the following universal mapping property: whenever $l : E \to A$ is a linear map from E into an associative algebra with identity, then there exists a unique associative algebra homomorphism $L : T(E) \to A$ with $L(1) = 1$ such that the diagram in Figure 6.5 commutes.

PROOF. Uniqueness is clear, since E and 1 generate $T(E)$ as an algebra. For existence we define $L^{(n)}$ on $T^n(E)$ to be the linear extension of the n-multilinear map

$$(v_1, v_2, \ldots, v_n) \mapsto l(v_1)l(v_2)\cdots l(v_n),$$

and we let $L = \bigoplus L^{(n)}$ in obvious notation. Let $u_1 \otimes \cdots \otimes u_m$ be in $T^m(E)$ and $v_1 \otimes \cdots \otimes v_n$ be in $T^n(E)$. Then we have

$$L^{(m)}(u_1 \otimes \cdots \otimes u_m) = l(u_1)\cdots l(u_m),$$

$$L^{(n)}(v_1 \otimes \cdots \otimes v_n) = l(v_1)\cdots l(v_n),$$

$$L^{(m+n)}(u_1 \otimes \cdots \otimes u_m \otimes v_1 \otimes \cdots \otimes v_n) = l(u_1)\cdots l(u_m)l(v_1)\cdots l(v_n).$$

Hence

$$L^{(m)}(u_1 \otimes \cdots \otimes u_m) L^{(n)}(v_1 \otimes \cdots \otimes v_n) = L^{(m+n)}(u_1 \otimes \cdots \otimes u_m \otimes v_1 \otimes \cdots \otimes v_n).$$

Taking linear combinations, we see that L is a homomorphism. □

Proposition 6.22 allows us to make $E \mapsto T(E)$ into a functor from the category of vector spaces over $\mathbb{K}$ to the category of associative algebras with identity over $\mathbb{K}$. To carry out the construction, we suppose that $\varphi : E \to F$ is a linear map between two vector spaces over $\mathbb{K}$. If $i : E \to T(E)$ and $j : F \to T(F)$ are the inclusion maps, then $j\varphi$ is a linear map from E into $T(F)$, and Proposition 6.22 produces a unique algebra homomorphism $\Phi : T(E) \to T(F)$ carrying 1 to 1 and satisfying $\Phi i = j\varphi$. Then the tensor-product functor is defined to carry the linear map φ to the homomorphism Φ of associative algebras with identity.

For the situation in which R is a commutative ring with identity, Section IV.5 introduced the ring $R[X_1, \ldots, X_n]$ of polynomials in n commuting indeterminates with coefficients in R. This ring was characterized by a universal mapping property saying that if a ring homomorphism of R into a commutative ring with identity were given and if n elements $t_1, \ldots, t_n$ were given, then the ring homomorphism of R could be extended uniquely to a ring homomorphism of $R[X_1, \ldots, X_n]$ carrying X_j into t_j for each j.

Proposition 6.22 yields a noncommutative version of this result, except that the ring of coefficients is assumed this time to be a field $\mathbb{K}$. To arrange for $X_1, \ldots, X_n$ to be *noncommuting* indeterminates, we form a vector space with $\{X_1, \ldots, X_n\}$ as a basis. Thus we let $E = \bigoplus_{j=1}^n \mathbb{K}X_j$. If $t_1, \ldots, t_n$ are arbitrary elements of an associative algebra A with identity, then the formulas $l(X_j) = t_j$ for $1 \le j \le n$ define a linear map $l : E \to A$. The associative-algebra homomorphism $L : T(E) \to A$ produced by the proposition extends the inclusion of $\mathbb{K}$ into the subfield $\mathbb{K}1$ of A and carries each X_j to t_j.

8. Symmetric Algebra

We continue to allow $\mathbb{K}$ to be an arbitrary field. Let E be a vector space over $\mathbb{K}$, and let $T(E)$ be the tensor algebra. We begin by defining the symmetric algebra $S(E)$. This is to be a version of $T(E)$ in which the elements, which are called symmetric tensors, commute with one another. It will not be canonically an algebra of polynomials, as we shall see presently, and thus we make no use of polynomial rings in the construction.

Just as the vector space of n-multilinear forms $E \times \cdots \times E \to \mathbb{K}$ is canonically the dual of $T^n(E)$, so the vector space of **symmetric** n-multilinear forms will be canonically the dual of $S^n(E)$. Here "symmetric" means that $f(x_1, \dots, x_n) = f(x_{\tau(1)}, \dots, x_{\tau(n)})$ for every permutation τ in the symmetric group $\mathfrak{S}_n$.

Since tensor algebras are supposed to be universal devices for constructing associative algebras over $\mathbb{K}$, whether commutative or not, we seek to form $S(E)$ as a quotient of $T(E)$. If q is the quotient homomorphism, we want to have $q(u \otimes v) = q(v \otimes u)$ in $S(E)$ whenever u and v are in $\iota(E) = T^1(E)$. Hence every element $u \otimes v - v \otimes u$ is to be in the kernel of the homomorphism. On the other hand, we do not want to impose any unnecessary conditions on our quotient, and so we factor out only what the elements $u \otimes v - v \otimes u$ force us to factor out. Thus we define the **symmetric algebra** by

$$S(E) = T(E)/I,$$

where
$$I = \begin{pmatrix} \text{two-sided ideal generated by all} \\ u \otimes v - v \otimes u \text{ with } u \text{ and } v \\ \text{in } T^1(E) \end{pmatrix}.$$

Then $S(E)$ is an associative algebra with identity.

Let us see that the fact that the generators of the ideal I are homogeneous elements (all being in $T^2(E)$) implies that

$$I = \bigoplus_{n=0}^{\infty} (I \cap T^n(E)).$$

In fact, each $I \cap T^n(E)$ is contained in I, and hence I contains the right side. On the other hand, if x is any element of I, then x is a sum of terms of the form $a \otimes (u \otimes v - v \otimes u) \otimes b$, and we may assume that each a and b is homogeneous. Any individual term $a \otimes (u \otimes v - v \otimes u) \otimes b$ is in some $I \cap T^n(E)$, and x is exhibited as a sum of members of the various intersections $I \cap T^n(E)$.

An ideal with the property $I = \bigoplus_{n=0}^{\infty} (I \cap T^n(E))$ is said to be **homogeneous**. Since I is homogeneous,

$$S(E) = \bigoplus_{n=0}^{\infty} T^n(E)/(I \cap T^n(E)).$$

We write $S^n(E)$ for the n^{th} summand on the right side, so that

$$S(E) = \bigoplus_{n=0}^{\infty} S^n(E).$$

Since $I \cap T^1(E) = 0$, the map of $E \to T^1(E) \to S^1(E)$ into first-order elements is one-one onto. The product operation in $S(E)$ is written without a product sign, the image in $S^n(E)$ of $v_1 \otimes \cdots \otimes v_n$ in $T^n(E)$ being written as $v_1 \cdots v_n$. If a is in $S^m(E)$ and b is in $S^n(E)$, then ab is in $S^{m+n}(E)$. Moreover, $S^n(E)$ is generated by elements $v_1 \cdots v_n$ with all v_j in $S^1(E) \cong E$, since $T^n(E)$ is generated by corresponding elements $v_1 \otimes \cdots \otimes v_n$. The defining relations for $S(E)$ make $v_i v_j = v_j v_i$ for v_i and v_j in $S^1(E)$, and it follows that the associative algebra $S(E)$ is commutative. □

Proposition 6.23. Let E be a vector space over the field $\mathbb{K}$.

(a) Let ι be the n-multilinear function $\iota(v_1, \ldots, v_n) = v_1 \cdots v_n$ of $E \times \cdots \times E$ into $S^n(E)$. Then $(S^n(E), \iota)$ has the following **universal mapping property**: whenever l is any symmetric n-multilinear map of $E \times \cdots \times E$ into a vector space U, then there exists a unique linear map $L : S^n(E) \to U$ such that the diagram

$$\begin{array}{ccc} E \times \cdots \times E & \xrightarrow{\ l\ } & U \\ {\scriptstyle \iota}\downarrow & \nearrow{\scriptstyle L} & \\ S^n(E) & & \end{array}$$

commutes.

(b) Let ι be the one-one linear function that embeds E as $S^1(E) \subseteq S(E)$. Then $(S(E), \iota)$ has the following **universal mapping property**: whenever l is any linear map of E into a commutative associative algebra A with identity, then there exists a unique algebra homomorphism $L : S(E) \to A$ with $L(1) = 1$ such that the diagram

$$\begin{array}{ccc} E & \xrightarrow{\ l\ } & A \\ {\scriptstyle \iota}\downarrow & \nearrow{\scriptstyle L} & \\ S(E) & & \end{array}$$

commutes.

PROOF. In both cases uniqueness is trivial. For existence we use the universal mapping properties of $T^n(E)$ and $T(E)$ to produce $\widetilde{L}$ on $T^n(E)$ or $T(E)$. If we can show that $\widetilde{L}$ annihilates the appropriate subspace so as to descend to $S^n(E)$ or $S(E)$, then the resulting map can be taken as L, and we are done. For (a), we have $\widetilde{L} : T^n(E) \to U$, and we are to show that $\widetilde{L}(T^n(E) \cap I) = 0$, where I is generated by all $u \otimes v - v \otimes u$ with u and v in $T^1(E)$. A member of $T^n(E) \cap I$ is thus of the form $\sum a_i \otimes (u_i \otimes v_i - v_i \otimes u_i) \otimes b_i$ with each term in $T^n(E)$. Each term here is a sum of pure tensors

$$x_1 \otimes \cdots \otimes x_r \otimes u_i \otimes v_i \otimes y_1 \otimes \cdots \otimes y_s - x_1 \otimes \cdots \otimes x_r \otimes v_i \otimes u_i \otimes y_1 \otimes \cdots \otimes y_s \quad (*)$$

with $r+2+s=n$. Since l by assumption takes equal values on

$$x_1 \times \cdots \times x_r \times u_i \times v_i \times y_1 \times \cdots \times y_s$$

and

$$x_1 \times \cdots \times x_r \times v_i \times u_i \times y_1 \times \cdots \times y_s,$$

$\widetilde{L}$ vanishes on $(*)$, and it follows that $\widetilde{L}(T^n(E) \cap I) = 0$.

For (b) we are to show that $\widetilde{L} : T(E) \to A$ vanishes on I. Since $\ker \widetilde{L}$ is an ideal, it is enough to check that $\widetilde{L}$ vanishes on the generators of I. But $\widetilde{L}(u \otimes v - v \otimes u) = l(u)l(v) - l(v)l(u) = 0$ by the commutativity of A, and thus $L(I) = 0$. □

Corollary 6.24. If E and F are vector spaces over the field $\mathbb{K}$, then the vector space $\operatorname{Hom}_{\mathbb{K}}(S^n(E), F)$ is canonically isomorphic (via restriction to pure tensors) to the vector space of all F-valued symmetric n-multilinear functions on $E \times \cdots \times E$.

PROOF. Restriction is linear and one-one. It is onto by Proposition 6.23a. □

Corollary 6.25. If E is a vector space over the field $\mathbb{K}$, then the dual $(S^n(E))'$ of $S^n(E)$ is canonically isomorphic (via restriction to pure tensors) to the vector space of symmetric n-multilinear forms on $E \times \cdots \times E$.

PROOF. This is a special case of Corollary 6.24. □

If $\varphi : E \to F$ is a linear map between vector spaces, then we can use Proposition 6.23b to define a corresponding homomorphism $\Phi : S(E) \to S(F)$ of associative algebras with identity. In this way, we can make $E \mapsto S(E)$ into a functor from the category of vector spaces over $\mathbb{K}$ to the category of commutative associative algebras with identity over $\mathbb{K}$. The details appear in Problem 14 at the end of the chapter.

Next we shall identify a basis for $S^n(E)$ as a vector space. The union of such bases as n varies will then be a basis of $S(E)$. Let $\{u_i\}_{i\in A}$ be a basis of E, possibly infinite. As noted in Section A5 of the appendix, a **simple ordering** on the index set A is a partial ordering in which every pair of elements is comparable and in which $a \le b$ and $b \le a$ together imply $a = b$.

Proposition 6.26. Let E be a vector space over the field $\mathbb{K}$, let $\{u_i\}_{i\in A}$ be a basis of E, and suppose that a simple ordering has been imposed on the index set A. Then the set of all monomials $u_{i_1}^{j_1} \cdots u_{i_k}^{j_k}$ with $i_1 < \cdots < i_k$ and $\sum_m j_m = n$ is a basis of $S^n(E)$.

REMARK. In particular if E is finite-dimensional with $(u_1, \dots, u_N)$ as an ordered basis, then the monomials $u_1^{j_1} \cdots u_N^{j_N}$ of total degree n form a basis of $S^n(E)$.

PROOF. Since $S(E)$ is commutative and since n-fold products of elements $\iota(u_i)$ in $T^1(E)$ span $T^n(E)$, the indicated set of monomials spans $S^n(E)$. Let us see that the set is linearly independent. Take any finite subset $F \subseteq A$ of indices. The map $\sum_{i \in A} c_i u_i \mapsto \sum_{i \in F} c_i X_i$ of E into the polynomial algebra $\mathbb{K}[\{X_i\}_{i \in F}]$ is linear into a commutative algebra with identity. Its extension via Proposition 6.23b maps all monomials in the u_i for $i \in F$ into distinct monomials in $\mathbb{K}[\{X_i\}_{i \in F}]$, which are necessarily linearly independent. Hence any finite subset of the monomials in the statement of the proposition is linearly independent, and the whole set must be linearly independent. Therefore our spanning set is a basis. $\square$

The proof of Proposition 6.26 shows that $S(E)$ may be identified with polynomials in indeterminates identified with members of E once a basis has been chosen, but this identification depends on the choice of basis. Indeed, if we think of E as specified in advance, then the isomorphism was set up by mapping the set $\{X_i\}_{i \in A}$ to the specified basis of E, and the result certainly depended on what basis was used. Nevertheless, if E is finite-dimensional, there is still an isomorphism that is independent of basis; it is between $S(E')$, where E' is the *dual* of E, and a natural basis-free notion of "polynomials" *on* E. We return to this point after one application of Proposition 6.26.

Corollary 6.27. Let E be a finite-dimensional vector space over $\mathbb{K}$ of dimension N. Then

(a) $\dim S^n(E) = \binom{n+N-1}{N-1}$ for $0 \le n < \infty$,

(b) $S^n(E')$ is canonically isomorphic to $S^n(E)'$ in such a way that

$$(f_1 \cdots f_n)(w_1 \cdots w_n) = \sum_{\tau \in \mathfrak{S}_n} \prod_{j=1}^{n} f_j(w_{\tau(j)}),$$

for any $f_1, \dots, f_n$ in E' and any $w_1, \dots, w_n$ in E, provided $\mathbb{K}$ has characteristic 0; here $\mathfrak{S}_n$ is the symmetric group on n letters.

PROOF. For (a), a basis has been described in Proposition 6.26. To see its cardinality, we recognize that picking out $N-1$ objects from $n+N-1$ to label as dividers is a way of assigning exponents to the u_j's in an ordered basis; thus the cardinality of the indicated basis is $\binom{n+N-1}{N-1}$.

For (b), let $f_1, \dots, f_n$ be in E' and $w_1, \dots, w_n$ be in E, and define

$$l_{f_1,\dots,f_n}(w_1, \dots, w_n) = \sum_{\tau \in \mathfrak{S}_n} \prod_{j=1}^{n} f_j(w_{\tau(j)}).$$

Then $l_{f_1,\dots,f_n}$ is symmetric n-multilinear from $E \times \cdots \times E$ into $\mathbb{K}$ and extends by Proposition 6.23a to a linear $L_{f_1,\dots,f_n} : S^n(E) \to \mathbb{K}$. Thus $l(f_1, \dots, f_n) = L_{f_1,\dots,f_n}$ defines a symmetric n-multilinear map of $E' \times \cdots \times E'$ into $S^n(E)'$. Its linear extension L maps $S^n(E')$ into $S^n(E)'$.

To complete the proof, we shall show that L carries basis to basis. Let $u_1, \dots, u_N$ be an ordered basis of E, and let $u_1', \dots, u_N'$ be the dual basis. Part (a) shows that the elements $(u_1')^{j_1} \cdots (u_N')^{j_N}$ with $\sum_m j_m = n$ form a basis of $S^n(E')$ and that the elements $(u_1)^{k_1} \cdots (u_N)^{k_N}$ with $\sum_m k_m = n$ form a basis of $S^n(E)$. We show that L of the basis of $S^n(E')$ is the dual basis of the basis of $S^n(E)$, except for positive-integer factors. Thus let all of $f_1, \dots, f_{j_1}$ be u_1', let all of $f_{j_1+1}, \dots, f_{j_1+j_2}$ be u_2', and so on. Similarly let all of $w_1, \dots, w_{k_1}$ be u_1, let all of $w_{k_1+1}, \dots, w_{k_1+k_2}$ be u_2, and so on. Then

$$\begin{aligned} L((u_1')^{j_1} \cdots (u_N')^{j_N})((u_1)^{k_1} \cdots (u_N)^{k_N}) &= L(f_1 \cdots f_n)(w_1 \cdots w_n) \\ &= l(f_1, \dots, f_n)(w_1 \cdots w_n) \\ &= \sum_{\tau \in \mathfrak{S}_n} \prod_{i=1}^{n} f_i(w_{\tau(i)}). \end{aligned}$$

For given τ, the product on the right side is 0 unless, for each index i, an inequality $j_{m-1} + 1 \le i \le j_m$ implies that $k_{m-1} + 1 \le \tau(i) \le k_m$. In this case the product is 1; so the right side counts the number of such τ's. For given τ, obtaining a nonzero product forces $k_m = j_m$ for all m. And when $k_m = j_m$ for all m, the choice $\tau = 1$ does lead to product 1. Hence the members of L of the basis are positive-integer multiples of the members of the dual basis, as asserted. □

Let us return to the question of introducing a basis-free notion of polynomials on the vector space E under the assumption that E is finite-dimensional. We take a cue from Corollary 4.32, which tells us that the evaluation homomorphism carrying $\mathbb{K}[X_1, \dots, X_n]$ to the algebra of $\mathbb{K}$-valued polynomial functions of $(t_1, \dots, t_n)$ is one-one if $\mathbb{K}$ is an infinite field. We regard the latter as the algebra of polynomial functions on $\mathbb{K}^n$, and we check what happens when we identify the vector space E with $\mathbb{K}^n$ by fixing a basis. Let $\Gamma = \{x_1, \dots, x_n\}$ be a basis of E, and let $\Gamma' = \{x_1', \dots, x_n'\}$ be the dual basis of E'. If $e = t_1x_1 + \cdots + t_nx_n$ is the expansion of a member of E in terms of Γ, then we have $x_j'(e) = t_j$. Thus the polynomial functions t_j are given by the members of the dual basis. The vector space of all homogeneous first-degree polynomial functions is the set of linear combinations of the t_j's, and these are given by arbitrary linear functionals on E. Thus the vector space of homogeneous first-degree polynomial functions on E is just the dual space E', and this conclusion does not depend on the choice of basis. The algebra of all polynomial functions on E is then the algebra of all $\mathbb{K}$-valued functions on E generated by E' and the constant functions.

This discussion tells us unambiguously what polynomial *functions* on E are to be, and we want to backtrack to handle abstract polynomials on E. Although the evaluation homomorphism from $\mathbb{K}[X_1, \ldots, X_n]$ to the algebra of polynomial functions on $\mathbb{K}^n$ may fail to be one-one if $\mathbb{K}$ is a finite field, its restriction to homogeneous first-degree polynomials *is* one-one. Thus, whatever we might mean by the vector space of homogeneous first-degree polynomials on E, the evaluation mapping should exhibit this space as isomorphic to E'.

Armed with these clues, we define the **polynomial algebra** $P(E)$ on E to be the symmetric algebra $S(E')$ if E is finite-dimensional. We need an evaluation mapping for each point e of E, and we obtain this from the universal mapping property of symmetric algebras (Proposition 6.23b): With e fixed, we have a linear map l from the vector space E' to the commutative associative algebra $\mathbb{K}$ given with $l(e') = e'(e)$. The universal mapping property gives us a unique algebra homomorphism $L : S(E') \to \mathbb{K}$ that extends l and carries 1 to 1. The algebra homomorphism L is then a multiplicative linear functional on $P(E) = S(E')$ that carries 1 to 1 and agrees with evaluation at e on homogeneous first-degree polynomials. We write this homomorphism as $p \mapsto p(e)$, and we define $P^n(E) = S^n(E')$; this is the vector space of homogeneous n^{th}-degree polynomials on E. A confirmation that $P(E)$ is indeed to be regarded as the algebra of abstract polynomials on E comes from the following.

Proposition 6.28. If E is a finite-dimensional vector space over the field $\mathbb{K}$, then the system of evaluation homomorphisms $P(E) \to \mathbb{K}$ on polynomials given by $p \mapsto \{p(e)\}_{e\in E}$ is an algebra homomorphism of $P(E)$ onto the algebra of $\mathbb{K}$-valued polynomial functions on E that carries the identity to the constant function 1, and it is one-one if $\mathbb{K}$ is an infinite field.

PROOF. Certainly $p \mapsto \{p(e)\}_{e\in E}$ is an algebra homomorphism of $P(E)$ into the algebra of $\mathbb{K}$-valued polynomial functions on E, and it carries the identity to the constant function 1. We have seen that the image of $P^1(E)$ is exactly E', and hence the image of $P(E)$ is the algebra of $\mathbb{K}$-valued functions on E generated by E' and the constants. This is exactly the algebra of all $\mathbb{K}$-valued polynomial functions, and hence the mapping is onto.

Suppose that $\mathbb{K}$ is infinite. The restriction of $p \mapsto \{p(e)\}_{e\in E}$ to the finite-dimensional subspace $P^n(E)$ of $P(E)$ maps into the finite-dimensional subspace of all polynomial functions on E homogeneous of degree n, and this restriction must therefore be onto. We can read off the dimension of the space of all polynomial functions on E homogeneous of degree n from Corollary 4.32 and Corollary 6.27a. This dimension matches the dimension of $P^n(E)$, according to Corollary 6.27a. Since the mapping is onto and the finite dimensions match, the restricted mapping is one-one. Hence $p \mapsto \{p(e)\}_{e\in E}$ is one-one. $\square$

We have defined the symmetric algebra $S(E)$ as a quotient of the tensor algebra $T(E)$. Now let us suppose that $\mathbb{K}$ has characteristic 0. With this hypothesis we shall be able to identify an explicit vector subspace of $T(E)$ that maps one-one onto $S(E)$ during the passage to the quotient. This subspace of $T(E)$ can therefore be viewed as a version of $S(E)$ for some purposes.

We define an n-multilinear function from $E \times \cdots \times E$ into $T^n(E)$ by

$$(v_1, \dots, v_n) \mapsto \frac{1}{n!} \sum_{\tau \in \mathfrak{S}_n} v_{\tau(1)} \otimes \cdots \otimes v_{\tau(n)},$$

and let $\sigma : T^n(E) \to T^n(E)$ be its linear extension. We call σ the **symmetrizer** operator. The image of σ in $T(E)$ is denoted by $\widetilde{S}^n(E)$, and the members of this subspace are called **symmetrized** tensors.

Proposition 6.29. Let the field $\mathbb{K}$ have characteristic 0, and let E be a vector space over $\mathbb{K}$. Then the symmetrizer operator σ satisfies $\sigma^2 = \sigma$. The kernel of σ on $T^n(E)$ is exactly $T^n(E) \cap I$, and therefore

$$T^n(E) = \widetilde{S}^n(E) \oplus (T^n(E) \cap I).$$

REMARK. In view of this corollary, the quotient map $T^n(E) \to S^n(E)$ carries $\widetilde{S}^n(E)$ one-one onto $S^n(E)$. Thus $\widetilde{S}^n(E)$ can be viewed as a copy of $S^n(E)$ embedded as a direct summand of $T^n(E)$.

PROOF. We have

$$\begin{aligned}
\sigma^2(v_1 \otimes \cdots \otimes v_n) &= \frac{1}{(n!)^2} \sum_{\rho,\tau \in \mathfrak{S}_n} v_{\rho\tau(1)} \otimes \cdots \otimes v_{\rho\tau(n)} \\
&= \frac{1}{(n!)^2} \sum_{\rho \in \mathfrak{S}_n} \sum_{\substack{\omega \in \mathfrak{S}_n, \\ (\omega = \rho\tau)}} v_{\omega(1)} \otimes \cdots \otimes v_{\omega(n)} \\
&= \frac{1}{n!} \sum_{\rho \in \mathfrak{S}_n} \sigma(v_1 \otimes \cdots \otimes v_n) \\
&= \sigma(v_1 \otimes \cdots \otimes v_n).
\end{aligned}$$

Hence $\sigma^2 = \sigma$. Thus σ fixes any member of image σ, and it follows that image $\sigma \cap \ker\sigma = 0$. Consequently $T^n(E)$ is the direct sum of image σ and $\ker\sigma$. We are left with identifying $\ker\sigma$ as $T^n(E) \cap I$.

The subspace $T^n(E) \cap I$ is spanned by elements

$$x_1 \otimes \cdots \otimes x_r \otimes u \otimes v \otimes y_1 \otimes \cdots \otimes y_s - x_1 \otimes \cdots \otimes x_r \otimes v \otimes u \otimes y_1 \otimes \cdots \otimes y_s$$

with $r + 2 + s = n$, and the symmetrizer σ certainly vanishes on such elements. Hence $T^n(E) \cap I \subseteq \ker\sigma$. Suppose that the inclusion is strict, say with t in $\ker\sigma$ but t not in $T^n(E) \cap I$. Let q be the quotient map $T^n(E) \to S^n(E)$. The kernel of q is $T^n(E) \cap I$, and thus $q(t) \neq 0$. From Proposition 6.26 the $T(E)$ monomials in basis elements from E with increasing indices map onto a basis of $S(E)$. Since $\mathbb{K}$ has characteristic 0, the symmetrized versions of these monomials map to nonzero multiples of the images of the initial monomials. Consequently q carries $\widetilde{S}^n(E) = \text{image}\,\sigma$ onto $S^n(E)$. Thus choose $t' \in \widetilde{S}^n(E)$ with $q(t') = q(t)$. Then $t' - t$ is in $\ker q = T^n(E) \cap I \subseteq \ker\sigma$. Since $\sigma(t) = 0$, we see that $\sigma(t') = 0$. Consequently t' is in $\ker\sigma \cap \text{image}\,\sigma = 0$, and we obtain $t' = 0$ and $q(t) = q(t') = 0$, contradiction. □

9. Exterior Algebra

We turn to a discussion of the exterior algebra. Let $\mathbb{K}$ be an arbitrary field, and let E be a vector space over $\mathbb{K}$. The construction, results, and proofs for the exterior algebra $\bigwedge(E)$ are similar to those for the symmetric algebra $S(E)$. The elements of $\bigwedge(E)$ are to be all the alternating tensors (= skew-symmetric if $\mathbb{K}$ has characteristic $\neq 2$), and so we want to force $v \otimes v = 0$. Thus we define the **exterior algebra** by

$$\textstyle\bigwedge(E) = T(E)/I',$$

where $$I' = \begin{pmatrix} \text{two-sided ideal generated by all} \\ v \otimes v \text{ with } v \text{ in } T^1(E) \end{pmatrix}.$$

Then $\bigwedge(E)$ is an associative algebra with identity.

It is clear that I' is homogeneous: $I' = \bigoplus_{n=0}^{\infty} (I' \cap T^n(E))$. Thus we can write

$$\textstyle\bigwedge(E) = \bigoplus_{n=0}^{\infty} T^n(E)/(I' \cap T^n(E)).$$

We write $\bigwedge^n(E)$ for the n^{th} summand on the right side, so that

$$\textstyle\bigwedge(E) = \bigoplus_{n=0}^{\infty} \bigwedge^n(E).$$

Since $I' \cap T^1(E) = 0$, the map of E into first-order elements $\bigwedge^1(E)$ is one-one onto. The product operation in $\bigwedge(E)$ is denoted by $\wedge$ rather than $\otimes$, the image in $\bigwedge^n(E)$ of $v_1 \otimes \cdots v_n$ in $T^n(E)$ being denoted by $v_1 \wedge \cdots \wedge v_n$. If a is in $\bigwedge^m(E)$ and b is in $\bigwedge^n(E)$, then $a \wedge b$ is in $\bigwedge^{m+n}(E)$. Moreover, $\bigwedge^n(E)$ is generated by elements $v_1 \wedge \cdots \wedge v_n$ with all v_j in $\bigwedge^1(E) \cong E$, since $T^n(E)$ is generated by corresponding elements $v_1 \otimes \cdots \otimes v_n$. The defining relations for $\bigwedge(E)$ make $v_i \wedge v_j = -v_j \wedge v_i$ for v_i and v_j in $\bigwedge^1(E)$, and it follows that

$$a \wedge b = (-1)^{mn} b \wedge a \qquad \text{for } a \in \textstyle\bigwedge^m(E) \text{ and } b \in \textstyle\bigwedge^n(E).$$

Proposition 6.30. Let E be a vector space over the field $\mathbb{K}$.

(a) Let ι be the n-multilinear function $\iota(v_1, \ldots, v_n) = v_1 \wedge \cdots \wedge v_n$ of $E \times \cdots \times E$ into $\bigwedge^n(E)$. Then $(\bigwedge^n(E), \iota)$ has the following **universal mapping property**: whenever l is any alternating n-multilinear map of $E \times \cdots \times E$ into a vector space U, then there exists a unique linear map $L : \bigwedge^n(E) \to U$ such that the diagram

$$\begin{array}{ccc} E \times \cdots \times E & \xrightarrow{\;l\;} & U \\ {\scriptstyle\iota}\big\downarrow & \nearrow {\scriptstyle L} & \\ \bigwedge^n(E) & & \end{array}$$

commutes.

(b) Let ι be the function that embeds E as $\bigwedge^1(E) \subseteq \bigwedge(E)$. Then $(\bigwedge(E), \iota)$ has the following **universal mapping property**: whenever l is any linear map of E into an associative algebra A with identity such that $l(v)^2 = 0$ for all $v \in E$, then there exists a unique algebra homomorphism $L : \bigwedge(E) \to A$ with $L(1) = 1$ such that the diagram

$$\begin{array}{ccc} E & \xrightarrow{\;l\;} & A \\ {\scriptstyle\iota}\big\downarrow & \nearrow {\scriptstyle L} & \\ \bigwedge(E) & & \end{array}$$

commutes.

PROOF. The proof is completely analogous to the proof of Proposition 6.23. $\square$

Corollary 6.31. If E and F are vector spaces over the field $\mathbb{K}$, then the vector space $\mathrm{Hom}_{\mathbb{K}}(\bigwedge^n(E), F)$ is canonically isomorphic (via restriction to pure tensors) to the vector space of all F-valued alternating n-multilinear functions on $E \times \cdots \times E$.

PROOF. Restriction is linear and one-one. It is onto by Proposition 6.30a. $\square$

Corollary 6.32. If E is a vector space over the field $\mathbb{K}$, then the dual $(\bigwedge^n(E))'$ of $\bigwedge^n(E)$ is canonically isomorphic (via restriction to pure tensors) to the vector space of alternating n-multilinear forms on $E \times \cdots \times E$.

PROOF. This is a special case of Corollary 6.31. $\square$

If $\varphi : E \to F$ is a linear map between vector spaces, then we can use Proposition 6.30b to define a corresponding homomorphism $\Phi : \bigwedge(E) \to \bigwedge(F)$ of associative algebras with identity. In this way, we can make $E \mapsto \bigwedge(E)$ into a functor from the category of vector spaces over $\mathbb{K}$ to the category of commutative associative algebras with identity over $\mathbb{K}$. We omit the details, which are similar to those for symmetric tensors.

Next we shall identify a basis for $\bigwedge^n(E)$ as a vector space. The union of such bases as n varies will then be a basis of $\bigwedge(E)$.

Proposition 6.33. Let E be a vector space over the field $\mathbb{K}$, let $\{u_i\}_{i \in A}$ be a basis of E, and suppose that a simple ordering has been imposed on the index set A. Then the set of all monomials $u_{i_1} \wedge \cdots \wedge u_{i_n}$ with $i_1 < \cdots < i_n$ is a basis of $\bigwedge^n(E)$.

PROOF. Since multiplication in $\bigwedge(E)$ satisfies $a \wedge b = (-1)^{mn} b \wedge a$ for $a \in \bigwedge^m(E)$ and $b \in \bigwedge^n(E)$ and since monomials span $T^n(E)$, the indicated set spans $\bigwedge^n(E)$. Let us see that the set is linearly independent. For $i \in A$, let u_i' be the member of E' with $u_i'(u_j)$ equal to 1 for $j = i$ and equal to 0 for $j \neq i$. Fix $r_1 < \cdots < r_n$, and define

$$l(w_1, \ldots, w_n) = \det\{u_{r_i}'(w_j)\} \qquad \text{for } w_1, \ldots, w_n \text{ in } E.$$

Then l is alternating n-multilinear from $E \times \cdots \times E$ into $\mathbb{K}$ and extends by Proposition 6.30a to $L : \bigwedge^n(E) \to \mathbb{K}$. If $k_1 < \cdots < k_n$, then

$$L(u_{k_1} \wedge \cdots \wedge u_{k_n}) = l(u_{k_1}, \ldots, u_{k_n}) = \det\{u_{r_i}'(u_{k_j})\},$$

and the right side is 0 unless $r_1 = k_1, \ldots, r_n = k_n$, in which case it is 1. This proves that the $u_{r_1} \wedge \cdots \wedge u_{r_n}$ are linearly independent in $\bigwedge^n(E)$. $\square$

Corollary 6.34. Let E be a finite-dimensional vector space over $\mathbb{K}$ of dimension N. Then

(a) $\dim \bigwedge^n(E) = \binom{N}{n}$ for $0 \le n \le N$ and $= 0$ for $n > N$,

(b) $\bigwedge^n(E')$ is canonically isomorphic to $\bigwedge^n(E)'$ by

$$(f_1 \wedge \cdots \wedge f_n)(w_1, \ldots, w_n) = \det\{f_i(w_j)\}.$$

PROOF. Part (a) is an immediate consequence of Proposition 6.33, and (b) is proved in the same way as Corollary 6.27b, using Proposition 6.30a as a tool. The "positive-integer multiples" that arise in the proof of Corollary 6.27b are all 1 in the current proof, and hence no restriction on the characteristic of $\mathbb{K}$ is needed. □

Now let us suppose that $\mathbb{K}$ has characteristic 0. We define an n-multilinear function from $E \times \cdots \times E$ into $T^n(E)$ by

$$(v_1, \dots, v_n) \mapsto \frac{1}{n!} \sum_{\tau \in \mathfrak{S}_n} (\operatorname{sgn} \tau) v_{\tau(1)} \otimes \cdots \otimes v_{\tau(n)},$$

and let $\sigma' : T^n(E) \to T^n(E)$ be its linear extension. We call σ' the **antisymmetrizer** operator. The image of σ' in $T(E)$ is denoted by $\widetilde{\bigwedge}^n(E)$, and the members of this subspace are called **antisymmetrized** tensors.

Proposition 6.35. Let the field $\mathbb{K}$ have characteristic 0, and let E be a vector space over $\mathbb{K}$. Then the antisymmetrizer operator σ' satisfies $\sigma'^2 = \sigma'$. The kernel of σ' on $T^n(E)$ is exactly $T^n(E) \cap I'$, and therefore

$$T^n(E) = \widetilde{\bigwedge}^n(E) \oplus (T^n(E) \cap I').$$

REMARK. In view of this corollary, the quotient map $T^n(E) \to \bigwedge^n(E)$ carries $\widetilde{\bigwedge}^n(E)$ one-one onto $\bigwedge^n(E)$. Thus $\widetilde{\bigwedge}^n(E)$ can be viewed as a copy of $\bigwedge^n(E)$ embedded as a direct summand of $T^n(E)$.

PROOF. We have

$$\begin{aligned} \sigma'^2(v_1 \otimes \cdots \otimes v_n) &= \frac{1}{(n!)^2} \sum_{\rho, \tau \in \mathfrak{S}_n} (\operatorname{sgn} \rho\tau) v_{\rho\tau(1)} \otimes \cdots \otimes v_{\rho\tau(n)} \\ &= \frac{1}{(n!)^2} \sum_{\rho \in \mathfrak{S}_n} \sum_{\substack{\omega \in \mathfrak{S}_n, \\ (\omega = \rho\tau)}} (\operatorname{sgn} \omega) v_{\omega(1)} \otimes \cdots \otimes v_{\omega(n)} \\ &= \frac{1}{n!} \sum_{\rho \in \mathfrak{S}_n} \sigma'(v_1 \otimes \cdots \otimes v_n) \\ &= \sigma'(v_1 \otimes \cdots \otimes v_n). \end{aligned}$$

Hence $\sigma'^2 = \sigma'$. Consequently $T^n(E)$ is the direct sum of image σ' and $\ker \sigma'$, and we are left with identifying $\ker \sigma'$ as $T^n(E) \cap I'$.

The subspace $T^n(E) \cap I'$ is spanned by elements

$$x_1 \otimes \cdots \otimes x_r \otimes v \otimes v \otimes y_1 \otimes \cdots \otimes y_s$$

with $r+2+s=n$, and the antisymmetrizer σ' certainly vanishes on such elements. Hence $T^n(E)\cap I' \subseteq \ker\sigma'$. Suppose that the inclusion is strict, say with t in $\ker\sigma'$ but t not in $T^n(E)\cap I'$. Let q be the quotient map $T^n(E)\to \bigwedge^n(E)$. The kernel of q is $T^n(E)\cap I'$, and thus $q(t)\neq 0$. From Proposition 6.33 the $T(E)$ monomials with strictly increasing indices map onto a basis of $\bigwedge(E)$. Since $\mathbb{K}$ has characteristic 0, the antisymmetrized versions of these monomials map to nonzero multiples of the images of the initial monomials. Consequently q carries $\widetilde{\bigwedge}^n(E) = \operatorname{image}\sigma'$ onto $\bigwedge^n(E)$. Thus choose $t'\in\widetilde{\bigwedge}^n(E)$ with $q(t')=q(t)$. Then $t'-t$ is in $\ker q = T^n(E)\cap I' \subseteq \ker\sigma'$. Since $\sigma'(t)=0$, we see that $\sigma'(t')=0$. Consequently t' is in $\ker\sigma'\cap\operatorname{image}\sigma' = 0$, and we obtain $t'=0$ and $q(t)=q(t')=0$, contradiction. □

10. Problems

1. Let V be a vector space over a field $\mathbb{K}$, and let $\langle\cdot\,,\,\cdot\rangle$ be a nondegenerate bilinear form on V.
   (a) Prove that every member v' of V is of the form $v'(w)=\langle v, w\rangle$ for one and only one member v of V.
   (b) Suppose that $(\cdot\,,\,\cdot)$ is another bilinear form on V. Prove that there is some linear function $L: V\to V$ such that $(v,w)=\langle L(v), w\rangle$ for all v and w in V.
2. The matrix $A=\begin{pmatrix}0&1\\1&0\end{pmatrix}$ with entries in $\mathbb{F}_2$ is symmetric. Prove that there is no nonsingular M with M^tAM diagonal.
3. This problem shows that one possible generalization of Sylvester's Law to other fields is not valid. Over the field $\mathbb{F}_3$, show that there is a nonsingular matrix M such that $\begin{pmatrix}-1&0\\0&-1\end{pmatrix} = M^t\begin{pmatrix}1&0\\0&1\end{pmatrix}M$. Conclude that the number of squares in $\mathbb{K}^\times$ among the diagonal entries of the diagonal form in Theorem 6.5 is not an invariant of the symmetric matrix.
4. Let V be a complex n-dimensional vector space, let $(\cdot\,,\,\cdot)$ be a Hermitian form on V, let $V_\mathbb{R}$ be the $2n$-dimensional real vector space obtained from V by restricting scalar multiplication to real scalars, and define $\langle\cdot\,,\,\cdot\rangle = \operatorname{Im}(\cdot\,,\,\cdot)$. Prove that
   (a) $\langle\cdot\,,\,\cdot\rangle$ is an alternating bilinear form on $V_\mathbb{R}$,
   (b) $\langle J(v_1), J(v_2)\rangle = \langle v_1, v_2\rangle$ for all v_1 and v_2 if $J: V_\mathbb{R}\to V_\mathbb{R}$ is what multiplication by i becomes when viewed as a linear map from $V_\mathbb{R}$ to itself,
   (c) $\langle\cdot\,,\,\cdot\rangle$ is nondegenerate on $V_\mathbb{R}$ if and only if $(\cdot\,,\,\cdot)$ is nondegenerate on V.
5. Let W be a $2n$-dimensional real vector space, and let $\langle\cdot\,,\,\cdot\rangle$ be a nondegenerate alternating bilinear form on W. Suppose that $J: W\to W$ is a linear map such

that $J^2 = -I$ and $\langle J(w_1), J(w_2)\rangle = \langle w_1, w_2\rangle$ for all w_1 and w_2 in W. Prove that W equals $V_{\mathbb{R}}$ for some n-dimensional complex vector space V possessing a Hermitian form whose imaginary part is $\langle \cdot\,, \cdot \rangle$.

6. This problem sharpens the result of Theorem 6.7 in the nondegenerate case. Let $\langle \cdot\,, \cdot \rangle$ be a nondegenerate alternating bilinear form on a $2n$-dimensional vector space V over $\mathbb{K}$. A vector subspace S of V is called an **isotropic** subspace if $\langle u, v\rangle = 0$ for all u and v in S. Prove that
   (a) any isotropic subspace of V that is maximal under inclusion has dimension n,
   (b) for any maximal isotropic subspace S_1, there exists a second maximal isotropic subspace S_2 such that $S_1 \cap S_2 = 0$.
   (c) if S_1 and S_2 are maximal isotropic subspaces of V such that $S_1 \cap S_2 = 0$, then the linear map $S_2 \to S_1'$ given by $s_2 \mapsto \langle \cdot\,, s_2\rangle\big|_{S_1}$ is an isomorphism of S_2 onto the dual space S_1'.
   (d) if S_1 and S_2 are maximal isotropic subspaces of V such that $S_1 \cap S_2 = 0$, then there exist bases $\{p_1, \dots, p_n\}$ of S_1 and $\{q_1, \dots, q_n\}$ of S_2 such that $\langle p_i, p_j\rangle = \langle q_i, q_j\rangle = 0$ and $\langle p_i, q_j\rangle = \delta_{ij}$ for all i and j. (The resulting basis $\{p_1, \dots, p_n, q_1, \dots, q_n\}$ of V is called a **Weyl basis** of V.)

7. Let S be a nonempty set, and let $\mathbb{K}$ be a field. For s in S, let U_s and V_s be vector spaces over $\mathbb{K}$, and let U and V be two further vector spaces over $\mathbb{K}$.
   (a) Prove that $\operatorname{Hom}_{\mathbb{K}}\big(\bigoplus_{s\in S} U_s, V\big) \cong \prod_{s\in S} \operatorname{Hom}_{\mathbb{K}}(U_s, V)$.
   (b) Prove that $\operatorname{Hom}_{\mathbb{K}}\big(U, \prod_{s\in S} V_s\big) \cong \prod_{s\in S} \operatorname{Hom}_{\mathbb{K}}(U, V_s)$.
   (c) Give examples to show that neither isomorphism in (a) and (b) need remain valid if all three direct products are changed to direct sums.

8. This problem continues Problem 1 at the end of Chapter V, which established a canonical-form theorem for an action of $GL(m, \mathbb{K}) \times GL(n, \mathbb{K})$ on m-by-n matrices. For the present problem, the group $GL(n, \mathbb{K})$ acts on $M_n(\mathbb{K})$ by $(g, x) \mapsto gxg^t$.
   (a) Verify that this is indeed a group action and that the vector subspaces $A_{nn}(\mathbb{K})$ of alternating matrices and $S_{nn}(\mathbb{K})$ of symmetric matrices are mapped into themselves under the group action.
   (b) Prove that two members of $A_{nn}(\mathbb{K})$ lie in the same orbit if and only if they have the same rank, and that the rank must be even. For each even rank $\leq n$, find an example of a member of $A_{nn}(\mathbb{K})$ with that rank.
   (c) Prove that two members of $S_{nn}(\mathbb{C})$ lie in the same orbit if and only if they have the same rank, and for each rank $\leq n$, find an example of a member of $S_{nn}(\mathbb{C})$ with that rank.

9. Let U and V be vector spaces over $\mathbb{K}$, and let U' be the dual of U. The bilinear map $(u', v) \mapsto u'(\cdot)v$ of $U' \times V$ into $\operatorname{Hom}_{\mathbb{K}}(U, V)$ extends to a linear map $T_{UV} : U' \otimes_{\mathbb{K}} V \to \operatorname{Hom}_{\mathbb{K}}(U, V)$. Do the following:

(a) Prove that T_{UV} is one-one.
(b) Prove that T_{UV} is onto $\mathrm{Hom}_{\mathbb{K}}(U, V)$ if U is finite-dimensional.
(c) Give an example for which T_{UV} is not onto $\mathrm{Hom}_{\mathbb{K}}(U, V)$.
(d) Let $\mathcal{C}$ be the category of all vector spaces over $\mathbb{K}$, and let Φ and Ψ be the functors from $\mathcal{C}\times\mathcal{C}$ into $\mathcal{C}$ whose effects on objects are $\Phi(U, V) = U' \otimes_{\mathbb{K}} V$ and $\Psi(U, V) = \mathrm{Hom}_{\mathbb{K}}(U, V)$. Prove that the system $\{T_{UV}\}$ is a natural transformation of Φ into Ψ.
(e) In view of (c), can the system $\{T_{UV}\}$ be a natural isomorphism?

10. Let $\mathbb{K} \subseteq \mathbb{L}$ be an inclusion of fields, and let $\mathcal{V}_{\mathbb{K}}$ and $\mathcal{V}_{\mathbb{L}}$ be the categories of vector spaces over $\mathbb{K}$ and $\mathbb{L}$. Section 6 of the text defined extension of scalars as a covariant functor $\Phi(E) = E \otimes_{\mathbb{K}} \mathbb{L}$. Another definition of extension of scalars is $\Psi(E) = \mathrm{Hom}_{\mathbb{K}}(\mathbb{L}, E)$ with $(l\varphi)(l') = \varphi(ll')$. Verify that $\Psi(E)$ is a vector space over $\mathbb{L}$ and that Ψ is a functor.

11. A linear map $L : E \to F$ between finite-dimensional complex vector spaces becomes a linear map $L_{\mathbb{R}} : E_{\mathbb{R}} \to F_{\mathbb{R}}$ when we restrict attention to real scalars. Explain how to express a matrix for $L_{\mathbb{R}}$ in terms of a matrix for L.

12. **(Kronecker product of matrices)** Let $L : E_1 \to E_2$ and $M : F_1 \to F_2$ be linear maps between finite-dimensional vector spaces over $\mathbb{K}$, let Γ_1 and Γ_2 be ordered bases of E_1 and E_2, and let Δ_1 and Δ_2 be ordered bases of F_1 and F_2. Define matrices A and B by $A = \begin{pmatrix} L \\ \Gamma_2\Gamma_1 \end{pmatrix}$ and $B = \begin{pmatrix} M \\ \Delta_2\Delta_1 \end{pmatrix}$. Use Γ_1, Γ_2, Δ_1, and Δ_2 to define ordered bases Ω_1 and Ω_2 of $E_1 \otimes_{\mathbb{K}} F_1$ and $E_2 \otimes_{\mathbb{K}} F_2$, and describe how the matrix $C = \begin{pmatrix} L\otimes M \\ \Omega_2\Omega_1 \end{pmatrix}$ is related to A and B.

13. Let $\mathbb{K}$ be a field, and let E be the vector space $\mathbb{K}X \oplus \mathbb{K}Y$. Prove that the subalgebra of $T(E)$ generated by 1, Y, and $X^2 + XY + Y^2$ is isomorphic as an algebra with identity to $T(F)$ for some vector space F.

Problems 14–17 concern the functors $E \mapsto T(E)$, $E \mapsto S(E)$, and $E \mapsto \bigwedge E$ defined for vector spaces over a field $\mathbb{K}$.

14. If $\varphi : E \to F$ is a linear map between vector spaces over $\mathbb{K}$, Section 8 of the text indicated how to define a corresponding homomorphism $\Phi : S(E) \to S(F)$ of associative algebras with identity over $\mathbb{K}$, using Proposition 6.23b.
(a) Fill in the details of this application of Proposition 6.23b.
(b) Establish the appropriate conditions on mappings that complete the proof that $E \mapsto S(E)$ is a functor.
(c) Verify that Φ carries $S^n(E)$ linearly into $S^n(F)$ for all integers $n \geq 0$.

15. Suppose that a linear map $\varphi : E \to E$ is given. Let $\Phi : S(E) \to S(E)$ and $\widetilde{\Phi} : T(E) \to T(E)$ be the associated algebra homomorphisms of $S(E)$ into itself and of $T(E)$ into itself, and let $q : T(E) \to S(E)$ be the quotient homomorphism appearing in the definition of $S(E)$. These mappings are related by the equation $\Phi q(x) = q\widetilde{\Phi}(x)$ for x in $T(E)$. Proposition 6.29 shows for each $n \geq 0$ that

$T^n(E) = \widetilde{S}^n(E) \oplus (T^n(E) \cap I)$, where $\widetilde{S}^n(E)$ is the image of $T^n(E)$ under the symmetrizer mapping. The remark with the proposition observes that q carries $\widetilde{S}^n(E)$ one-one onto $S^n(E)$. Prove that $\widetilde{\Phi}$ carries $\widetilde{S}^n(E)$ into itself and that $\widetilde{\Phi}\big|_{\widetilde{S}^n(E)}$ matches $\Phi\big|_{S^n(E)}$ in the sense that $q\widetilde{\Phi}(x) = \Phi q(x)$ for all x in $\widetilde{S}^n(E)$.

16. With E finite-dimensional let $\varphi : E \to E$ be a linear mapping, and define $\Phi : \bigwedge E \to \bigwedge E$ to be the corresponding algebra homomorphism of $\bigwedge E$ sending 1 into 1. This carries each $\bigwedge^n E$ into itself. Prove that Φ acts as multiplication by the scalar $\det \varphi$ on the 1-dimensional space $\bigwedge^{\dim E}(E)$.

17. Suppose that G is a group, that the vector space E over $\mathbb{K}$ is finite-dimensional, and that $\varphi : G \to \mathrm{GL}(E)$ is a representation of G on E. The functors $E \mapsto T(E)$, $E \mapsto S(E)$, and $E \mapsto \bigwedge E$ yield, for each $\varphi(g)$, algebra homomorphisms of $T(E)$ into itself, $S(E)$ into itself, and $\bigwedge E$ into itself.
    (a) Show that as g varies, the result in each case is a representation of G.
    (b) Suppose that $E = \mathbb{K}^n$. Give a formula for the representation of G on a member of $P(\mathbb{K}^n) = S((\mathbb{K}^n)')$.

Problems 18–22 concern universal mapping properties. Let $\mathcal{A}$ and $\mathcal{V}$ be two categories, and let $\mathcal{F} : \mathcal{A} \to \mathcal{V}$ be a covariant functor. (In practice, $\mathcal{F}$ tends to be a relatively simple functor, such as one that simply ignores some of the structure of $\mathcal{A}$.) Let E be in $\mathrm{Obj}(\mathcal{V})$. A pair (S, ι) with S in $\mathrm{Obj}(\mathcal{A})$ and ι in $\mathrm{Morph}_{\mathcal{V}}(E, \mathcal{F}(S))$ is said to have the **universal mapping property** relative to E and $\mathcal{F}$ if the following condition is satisfied: whenever A is in $\mathrm{Obj}(\mathcal{A})$ and a member l of $\mathrm{Morph}_{\mathcal{V}}(E, \mathcal{F}(A))$ is given, there exists a unique member L of $\mathrm{Morph}_{\mathcal{A}}(S, A)$ such that $\mathcal{F}(L)\,\iota = l$.

18. (a) By suitably specializing $\mathcal{A}$, $\mathcal{V}$, $\mathcal{F}$, etc., show that the universal mapping property of the symmetric algebra of a vector space over $\mathbb{K}$ is an instance of what has been described.
    (b) How should the answer to (a) be adjusted so as to account for the universal mapping property of the exterior algebra of a vector space over $\mathbb{K}$?
    (c) How should the answer to (a) be adjusted so as to account for the universal mapping property of the coproduct of $\{X_j\}_{j\in J}$ in a category $\mathcal{C}$, the universal mapping property being as in Figure 4.12? (Educational note: For the *product* of $\{X_j\}_{j\in J}$ in $\mathcal{C}$, the above description does not apply directly because the morphisms go the wrong way. Instead, one applies the above description to the opposite categories $\mathcal{A}^{\mathrm{opp}}$ and $\mathcal{V}^{\mathrm{opp}}$, defined as in Problems 78–80 at the end of Chapter IV.)

19. If (S, ι) and (S', ι') are two pairs that each have the universal mapping property relative to E and $\mathcal{F}$, prove that S and S' are canonically isomorphic as objects in $\mathcal{A}$. More specifically prove that there exists a unique L in $\mathrm{Morph}_{\mathcal{A}}(S, S')$ such that $\mathcal{F}(L)\iota = \iota'$ and that L is an isomorphism whose inverse L' in $\mathrm{Morph}_{\mathcal{A}}(S', S)$ has $\mathcal{F}(L')\iota' = \iota$.

20. Suppose that the pair (S, ι) has the universal mapping property relative to E and $\mathcal{F}$. Let $\mathcal{S}$ be the category of sets, and define functors $F : \mathcal{A} \to \mathcal{S}$ and $G : \mathcal{A} \to \mathcal{S}$ by $F(A) = \operatorname{Morph}_{\mathcal{A}}(S, A)$, $F(\varphi)$ equals composition on the left by φ for $\varphi \in \operatorname{Morph}_{\mathcal{A}}(A, A')$, $G(A) = \operatorname{Morph}_{\mathcal{V}}(E, \mathcal{F}(A))$, and $G(\varphi)$ equals composition on the left by $\mathcal{F}(\varphi)$. Let $T_A : \operatorname{Morph}_{\mathcal{A}}(S, A) \to \operatorname{Morph}_{\mathcal{V}}(E, \mathcal{F}(A))$ be the one-one onto map given by the universal mapping property. Show that the system $\{T_A\}$ is a natural isomorphism of F into G.

21. Suppose that (S', ι) is a second pair having the universal mapping property relative to E and $\mathcal{F}$. Define $F' : \mathcal{A} \to \mathcal{S}$ by $F'(A) = \operatorname{Morph}_{\mathcal{A}}(S', A)$. Combining the previous problem and Proposition 6.16, obtain a second proof (besides the one in Problem 19) that S and S' are canonically isomorphic.

22. Suppose that for each E in $\operatorname{Obj}(\mathcal{V})$, there is some pair (S, ι) with the universal mapping property relative to E and $\mathcal{F}$. Fix such a pair (S, ι) for each E, calling it $(S(E), \iota_E)$. Making an appropriate construction for morphisms and carrying out the appropriate verifications, prove that $E \mapsto S(E)$ is a functor.

Problems 23–28 introduce the **Pfaffian** of a $(2n)$-by-$(2n)$ alternating matrix $X = [x_{ij}]$ with entries in a field $\mathbb{K}$. This is the polynomial in the entries of X with integer coefficients given by

$$\operatorname{Pfaff}(X) = \sum_{\substack{\text{some } \tau\text{'s} \\ \text{in } \mathfrak{S}_{2n}}} (\operatorname{sgn} \tau) \prod_{k=1}^{n} x_{\tau(2k-1), \tau(2k)},$$

where the sum is taken over those permutations τ such that $\tau(2k-1) < \tau(2k)$ for $1 \le k \le n$ and such that $\tau(1) < \tau(3) < \cdots < \tau(2n-1)$. It will be seen that $\det X$ is the square of this polynomial. Examples of Pfaffians are

$$\operatorname{Pfaff}\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} = x \qquad \text{and} \qquad \operatorname{Pfaff}\begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix} = af - be + cd.$$

The problems in this set will be continued at the end of Chapter VIII.

23. For the matrix J in Section 5, show that $\operatorname{Pfaff}(J) = 1$.

24. In the expansion $\det X = \sum_{\sigma \in \mathfrak{S}_{2n}} (\operatorname{sgn} \sigma) \prod_{l=1}^{2n} x_{l, \sigma(l)}$, prove that the value of the right side with X as above is not changed if the sum is extended only over those σ's whose expansion in terms of disjoint cycles involves only cycles of even length (and in particular no cycles of length 1).

25. Define $\sigma \in \mathfrak{S}_{2n}$ to be "good" if its expansion in terms of disjoint cycles involves only cycles of even length. If σ is good, show that there uniquely exist two disjoint subsets A and B of n elements each in $\{1, \ldots, 2n\}$ such that A contains the smallest-numbered index in each cycle and such that σ maps each set onto the other.

26. In the notation of the previous problem with σ good, let $y(\sigma)$ be the product of the monomials x_{ab} such that a is in A and $b = \sigma(a)$. For each factor x_{ij} of $y(\sigma)$ with $i > j$, replace the factor by $-x_{ji}$. In the resulting product, arrange the factors in order so that their first subscripts are increasing, and denote this expression by $s x_{i_1 i_2} x_{i_3 i_4} \cdots x_{i_{2n-1} i_{2n}}$, where s is a sign. Let τ be the permutation that carries each r to i_r, and define $s(\tau)$ to be the sign s. Similarly let $z(\sigma)$ be the product of the monomials x_{ba} such that b is in B and $a = \sigma(b)$. For each factor x_{ij} of $z(\sigma)$ with $i > j$, replace the factor by $-x_{ji}$. In the resulting product, arrange the factors in order so that their first subscripts are increasing, and denote this expression by $s' x_{j_1 j_2} x_{j_3 j_4} \cdots x_{j_{2n-1} j_{2n}}$, where s' is a sign. Let τ' be the permutation that carries each r to j_r, and define $s'(\tau')$ to be the sign s'. Prove, apart from signs, that the σ^{th} term in the expansion of $\det X$ matches the product of the τ^{th} term of $\text{Pfaff}(X)$ and the τ'^{th} term of $\text{Pfaff}(X)$.
27. In the previous problem, take the signs $s(\tau)$ and $s'(\tau')$ into account and show that the signs of σ, τ, and τ' work out so that the σ^{th} term in the expansion of $\det X$ is the product of the τ^{th} and τ'^{th} terms of $\text{Pfaff}(X)$.
28. Show that every term of the product of $\text{Pfaff}(X)$ with itself is accounted for once and only once by the construction in the previous three problems, and conclude that the alternating matrix X has $\det X = (\text{Pfaff}(X))^2$.

Problems 29–30 concern filtrations and gradings. A vector space V over $\mathbb{K}$ is said to be **filtered** when an increasing sequence of subspaces $V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots$ is specified with union V. In this case we put $V_{-1} = 0$ by convention. The space V is **graded** if a sequence of subspaces $V^0, V^1, V^2, \ldots$ is specified such that

$$V = \bigoplus_{n=0}^{\infty} V^n.$$

When V is graded, there is a natural filtration of V given by $V_n = \bigoplus_{k=0}^{n} V^k$. Examples of graded vector spaces are any tensor algebra $V = T(E)$, symmetric algebra $S(E)$, exterior algebra $\bigwedge(E)$, and polynomial algebra $P(E)$, the n^{th} subspace of the grading consisting of those elements that are homogeneous of degree n. Any polynomial algebra $\mathbb{K}[X_1, \ldots, X_n]$ is another example of a graded vector space, the grading being by total degree.

29. When V is a filtered vector space as in (A.34), the **associated graded vector space** is $\text{gr}\ V = \bigoplus_{n=0}^{\infty} V_n/V_{n-1}$. Let V and $V^\#$ be two filtered vector spaces, and let φ be a linear map between them such that $\varphi(V_n) \subseteq V_n^\#$ for all n. Since the restriction of φ to V_n carries V_{n-1} into $V_{n-1}^\#$, this restriction induces a linear map $\text{gr}^n\ \varphi : (V_n/V_{n-1}) \to (V_n^\#/V_{n-1}^\#)$. The direct sum of these linear maps is then a linear map $\text{gr}\ \varphi : \text{gr}\ V \to \text{gr}\ V^\#$ called the **associated graded map** for φ. Prove that if $\text{gr}\ \varphi$ is a vector-space isomorphism, then φ is a vector-space isomorphism.

30. Let A be an associative algebra over $\mathbb{K}$ with identity. If A has a filtration $A_0, A_1, \ldots$ of vector subspaces with $1 \in A_0$ such that $A_m A_n \subseteq A_{m+n}$ for all m and n, then one says that A is a **filtered associative algebra**; similarly if A is graded as $A = \bigoplus_{n=0}^{\infty} A^n$ in such a way that $A^m A^n \subseteq A^{m+n}$ for all m and n, then one says that A is a **graded associative algebra**. If A is a filtered associative algebra with identity, prove that the graded vector space $\operatorname{gr} A$ acquires a multiplication in a natural way, making it into a graded associative algebra with identity.

Problems 31–35 concern Lie algebras and their universal enveloping algebras. If $\mathbb{K}$ is a field, a **Lie algebra** $\mathfrak{g}$ over $\mathbb{K}$ is a nonassociative algebra whose product, called the **Lie bracket** and written $[x, y]$, is alternating as a function of the pair (x, y) and satisfies the **Jacobi identity** $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all x, y, z in $\mathfrak{g}$. The **universal enveloping algebra** $U(\mathfrak{g})$ of $\mathfrak{g}$ is the quotient $T(\mathfrak{g})/I''$, where I'' is the two-sided ideal generated by all elements $x \otimes y - y \otimes x - [x, y]$ with x and y in $T^1(\mathfrak{g})$. The grading for $T(\mathfrak{g})$ makes $U(\mathfrak{g})$ into a filtered associate algebra with identity. The product of x and y in $U(\mathfrak{g})$ is written xy.

31. If A is an associative algebra over $\mathbb{K}$, prove that A becomes a Lie algebra if the Lie bracket is defined by $[x, y] = xy - yx$. In particular, observe that $M_n(\mathbb{K})$ becomes a Lie algebra in this way.

32. Fix a matrix $A \in M_n(\mathbb{K})$, and let $\mathfrak{g}$ be the vector subspace of all members x of $M_n(\mathbb{K})$ with $x^t A + Ax = 0$.
    (a) Prove that $\mathfrak{g}$ is closed under the bracket operation of the previous problem and is therefore a Lie subalgebra of $M_n(\mathbb{K})$.
    (b) Deduce as a special case of (a) that the vector space of all skew-symmetric matrices in $M_n(\mathbb{K})$ is a Lie subalgebra of $M_n(\mathbb{K})$.

33. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$, and let ι be the linear map obtained as the composition of $\mathfrak{g} \to T^1(\mathfrak{g})$ and the passage to the quotient $U(\mathfrak{g})$. Prove that $(U(\mathfrak{g}), \iota)$ has the following universal mapping property: whenever l is any linear map of $\mathfrak{g}$ into an associative algebra A with identity satisfying the condition of being a Lie algebra homomorphism, namely $l[x, y] = l(x)l(y) - l(y)l(x)$ for all x and y in $\mathfrak{g}$, then there exists a unique associative algebra homomorphism $L : U(\mathfrak{g}) \to A$ with $L(1) = 1$ such that $L \circ \iota = l$.

34. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$, let $\{u_i\}_{i \in A}$ be a vector-space basis of $\mathfrak{g}$, and suppose that a simple ordering has been imposed on the index set A. Prove that the set of all monomials $u_{i_1}^{j_1} \cdots u_{i_k}^{j_k}$ with $i_1 < \cdots < i_k$ and $\sum_m j_m$ arbitrary is a spanning set for $U(\mathfrak{g})$.

35. For a Lie algebra $\mathfrak{g}$ over $\mathbb{K}$, the **Poincaré–Birkhoff–Witt Theorem** says that the spanning set for $U(\mathfrak{g})$ in the previous problem is actually a basis. Assuming this theorem, prove that $\operatorname{gr} U(\mathfrak{g})$ is isomorphic as a graded algebra to $S(\mathfrak{g})$.

Problems 36–40 introduce Clifford algebras. Let $\mathbb{K}$ be a field of characteristic $\neq 2$, let E be a finite-dimensional vector space over $\mathbb{K}$, and let $\langle \cdot\, , \cdot \rangle$ be a symmetric bilinear form on E. The **Clifford algebra** $\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$ is the quotient $T(E)/I''$, where I'' is the two-sided ideal generated by all elements[5] $v \otimes v + \langle v, v \rangle$ with v in E. The grading for $T(E)$ makes $\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$ into a filtered associative algebra with identity. Products in $\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$ are written as ab with no special symbol.

36. Let ι be the composition of the inclusion $E \subseteq T^1(E)$ and the passage to the quotient modulo I''. Prove that $(\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle), \iota)$ has the following universal mapping property: whenever l is any linear map of E into an associative algebra A with identity such that $l(v)^2 = -\langle v, v \rangle 1$ for all $v \in E$, then there exists a unique algebra homomorphism $L : \operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle) \to A$ with $L(1) = 1$ and such that $L \circ \iota = l$.
37. Let $\{u_1, \dots, u_n\}$ be a basis of E. Prove that the 2^n elements of $\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$ given by $u_{i_1} u_{i_2} \cdots u_{i_k}$ with $i_1 < \cdots < i_k$ form a spanning set of $\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$.
38. Using the Principal Axis Theorem, fix a basis $\{e_1, \dots, e_n\}$ of E such that $\langle e_i, e_j \rangle = d_i \delta_{ij}$ for all j. Introduce an algebra C over $\mathbb{K}$ of dimension 2^n with generators $e_1, \dots, e_n$ and with a basis parametrized by subsets of $\{1, \dots, n\}$ and given by all elements
$$e_{i_1} e_{i_2} \cdots e_{i_k} \qquad \text{with} \qquad i_1 < i_2 < \cdots < i_k,$$
with the multiplication that is implicit in the rules
$$e_i^2 = -d_i \qquad \text{and} \qquad e_i e_j = -e_j e_i \quad \text{if } i \neq j,$$
namely, to multiply two monomials $e_{i_1} e_{i_2} \cdots e_{i_k}$ and $e_{j_1} e_{j_2} \cdots e_{j_l}$, put them end to end, replace any occurrence of two e_k's by the scalar $-d_k$, and then permute the remaining e_k's until their indices are in increasing order, introducing a minus sign each time two distinct e_k's are interchanged. Prove that the algebra C is associative.
39. Prove that the associative algebra C of the previous problem is isomorphic as an algebra to $\operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$.
40. Prove that $\operatorname{gr} \operatorname{Cliff}(E, \langle \cdot\, , \cdot \rangle)$ is isomorphic as a graded algebra to $\bigwedge(E)$.

Problems 41–48 introduce finite-dimensional Heisenberg Lie algebras and the corresponding Weyl algebras. They make use of Problems 31–35 concerning Lie algebras and universal enveloping algebras. Let V be a finite-dimensional vector space over the field $\mathbb{K}$, and let $\langle \cdot\, , \cdot \rangle$ be a nondegenerate alternating bilinear form on $V \times V$. Write $2n$ for the dimension of V. Introduce an indeterminate X_0. The **Heisenberg Lie algebra** $H(V)$ on V is a Lie algebra whose underlying vector space is $\mathbb{K}X_0 \oplus V$ and whose Lie bracket is given by $[(cX_0, u), (dX_0, v)] = \langle u, v \rangle X_0$. Let $U(H(V))$ be its universal enveloping algebra. The **Weyl algebra** $W(V)$ on V is the quotient of the tensor algebra $T(V)$ by the two-sided ideal generated by all $u \otimes v - v \otimes u - \langle u, v \rangle 1$ with u and v in V; as such, it is a filtered associative algebra.

[5]Some authors factor out the elements $v \otimes v - \langle v, v \rangle$ instead. There is no generally accepted convention.

41. Verify when the field is $\mathbb{K} = \mathbb{R}$ that an example of a $2n$-dimensional V with its nondegenerate alternating bilinear form $\langle \cdot\,, \cdot \rangle$ is $V = \mathbb{C}^n$ with $\langle u, v\rangle = \operatorname{Im}(u, v)$, where $(\cdot\,, \cdot)$ is the usual inner product on $\mathbb{C}^n$. For this V, exhibit a Lie-algebra isomorphism of $H(V)$ with the Lie algebra of all complex $(n+1)$-by-$(n+1)$ matrices of the form $\begin{pmatrix} 0 & \bar{z}^t & ir \\ 0 & 0 & z \\ 0 & 0 & 0 \end{pmatrix}$ with $z \in \mathbb{C}^n$ and $r \in \mathbb{R}$.

42. In the general situation show that the linear map $\iota(cX_0, v) = c1 + v$ is a Lie algebra homomorphism of $H(V)$ into $W(V)$ and that its extension to an associative algebra homomorphism $\widetilde{\iota} : U(H(V)) \to W(V)$ is onto and has kernel equal to the two-sided ideal in $U(H(V))$ generated by $X_0 - 1$.

43. Prove that $W(V)$ has the following universal mapping property: whenever $\varphi : H(V) \to A$ is a Lie algebra homomorphism of $H(V)$ into an associative algebra A with identity such that $\varphi(X_0) = 1$, then there exists a unique associative algebra homomorphism $\widetilde{\varphi}$ of $W(V)$ into A such that $\varphi = \widetilde{\varphi} \circ \iota$.

44. Let $v_1, \dots, v_{2n}$ be any vector space basis of V. Prove that the elements $v_1^{k_1} \cdots v_{2n}^{k_{2n}}$ with integer exponents ≥ 0 span $W(V)$.

45. For $\mathbb{K} = \mathbb{R}$, let $\mathcal{S}$ be the vector space of all real-valued functions $P(x)e^{-\pi|x|^2}$, where $P(x)$ is a polynomial in n real variables. Show that $\mathcal{S}$ is mapped into itself by the linear operators $\partial/\partial x_i$ and $m_j =$ (multiplication by x_j).

46. With $\mathbb{K} = \mathbb{R}$, let $\{p_1, \dots, p_n, q_1, \dots, q_n\}$ be a Weyl basis of V in the terminology of Problem 6. In the notation of Problem 45, let $\varphi : V \to \operatorname{Hom}_{\mathbb{R}}(\mathcal{S}, \mathcal{S})$ be the linear map given by $\varphi(p_i) = \partial/\partial x_i$ and $\varphi(q_j) = m_j$. Use Problem 43 to extend φ to an algebra homomorphism $\widetilde{\varphi} : W(V) \to \operatorname{Hom}_{\mathbb{R}}(\mathcal{S}, \mathcal{S})$ with $\widetilde{\varphi}(1) = 1$, and use Problem 42 to obtain a representation of $H(V)$ on $\mathcal{S}$. Prove that this representation of $H(V)$ is irreducible in the sense that there is no proper nonzero vector subspace carried to itself by all members of $\widetilde{\varphi}(H(V))$.

47. In Problem 46 with $\mathbb{K} = \mathbb{R}$, prove that the associative algebra homomorphism $\widetilde{\varphi} : W(V) \to \operatorname{Hom}_{\mathbb{R}}(\mathcal{S}, \mathcal{S})$ is one-one. Conclude for $\mathbb{K} = \mathbb{R}$ that the elements $v_1^{k_1} \cdots v_{2n}^{k_{2n}}$ of Problem 44 form a vector-space basis of $W(V)$.

48. For $\mathbb{K} = \mathbb{R}$, prove that $\operatorname{gr} W(V)$ is isomorphic as a graded algebra to $S(V)$.

Problems 49–51 deal with Jordan algebras. Let $\mathbb{K}$ be a field of characteristic $\neq 2$. An algebra J over $\mathbb{K}$ with multiplication $a \cdot b$ is called a **Jordan algebra** if the identities $a \cdot b = b \cdot a$ and $a^2 \cdot (b \cdot a) = (a^2 \cdot b) \cdot a$ are always satisfied; here a^2 is an abbreviation for $a \cdot a$.

49. Let A be an associative algebra, and define $a \cdot b = \frac{1}{2}(ab + ba)$. Prove that A becomes a Jordan algebra under this new multiplication.

50. In the situation of the previous problem, suppose that $a \mapsto a^t$ is a one-one linear mapping of A onto itself such that $(ab)^t = b^t a^t$ for all a and b. (For example, $a \mapsto a^t$ could be the transpose mapping if $A = M_n(\mathbb{K})$.) Prove that the vector subspace of all a with $a^t = a$ is carried to itself by the Jordan product $a \cdot b$ and hence is a Jordan algebra.

51. Let V be a finite-dimensional vector space over $\mathbb{K}$, and let $\langle \cdot\,, \,\cdot \rangle$ be a symmetric bilinear form on V. Define $A = \mathbb{K}1 \oplus V$ as a vector space, and define a multiplication in A by $(c1, x) \cdot (d1, y) = \big((cd + \langle x, y\rangle)1,\ cy + dx\big)$. Prove that A is a Jordan algebra under this definition of multiplication.

Problems 52–56 deal with the algebra $\mathbb{O}$ of real **octonions**, sometimes known as the **Cayley numbers**. This is a certain 8-dimensional nonassociative algebra with identity over $\mathbb{R}$ with an inner product such that $\|ab\| = \|a\|\|b\|$ for all a and b and such that the left and right multiplications by any element $a \neq 0$ are always invertible.

52. Let A be an algebra over $\mathbb{R}$. Let $[a, b] = ab - ba$ and $[a, b, c] = (ab)c - a(bc)$.
    (a) The 3-multilinear function $(a, b, c) \mapsto [a, b, c]$ from $A \times A \times A$ to A is called the **associator** in A. Observe that it is 0 if and only if A is associative. Show that it is alternating if and only if A always satisfies the limited associativity laws

$$(aa)b = a(ab), \qquad (ab)a = a(ba), \qquad (ba)a = b(aa).$$

    In this case, A is said to be **alternative**.
    (b) Show that A is alternative if the first and third of the limited associativity laws in (a) are always satisfied.

53. **(Cayley–Dickson construction)** Suppose that A is an algebra over $\mathbb{R}$ with a two-sided identity 1, and suppose that there is an $\mathbb{R}$ linear function $*$ from A to itself (called "conjugation") such that $1^* = 1$, $a^{**} = a$, and $(ab)^* = b^* a^*$ for all a and b in A. Define an algebra B over $\mathbb{R}$ to have the underlying real vector-space structure of $A \oplus A$ and to have multiplication and conjugation given by

$$(a, b)(c, d) = (ac - db^*,\ a^*d + cb) \qquad \text{and} \qquad (a, b)^* = (a^*, -b).$$

    (a) Prove that $(1, 0)$ is a two-sided identity in B and that the operation $*$ in B satisfies the required properties of a conjugation.
    (b) Prove that if $a^* = a$ for all $a \in A$, then A is commutative.
    (c) Prove that if $a^* = a$ for all $a \in A$, then B is commutative.
    (d) Prove that if A is commutative and associative, then B is associative.
    (e) Verify the following outcomes of the above construction $A \to B$:
        (i) $A = \mathbb{R}$ yields $B = \mathbb{C}$,
        (ii) $A = \mathbb{C}$ yields $B = \mathbb{H}$, the algebra of quaternions.

54. Suppose that A is an algebra over $\mathbb{R}$ with an identity and a conjugation as in the previous problem. Say that A is **nicely normed** if
    (i) $a + a^*$ is always of the form $r1$ with r real and
    (ii) aa^* always equals a^*a and for $a \neq 0$, is of the form $r1$ with r real and positive.
    (a) Prove that if A is nicely normed, then so is the algebra B of the previous problem.
    (b) Prove that if A is nicely normed, then $(a, b) = \frac{1}{2}(ab^* + ba^*)$ is an inner product on A with norm $\|a\| = (aa^*)^{1/2} = (a^*a)^{1/2}$.
    (c) Prove that if A is associative and nicely normed, then the algebra B of the previous problem is alternative.
55. Starting from the real algebra $A = \mathbb{H}$, apply the construction of Problem 53, and let the resulting 8-dimensional real algebra be denoted by $\mathbb{O}$, the algebra of octonions.
    (a) Prove that $\mathbb{O}$ is an alternative algebra and is nicely normed.
    (b) Prove that $(xx^*)y = x(x^*y)$ and $x(yy^*) = (xy)y^*$ within $\mathbb{O}$.
    (c) Prove that $\|ab\|^2 a = \|a\|^2\|b\|^2 a$ within $\mathbb{O}$.
    (d) Conclude from (c) that the operations of left and right multiplication by any $a \neq 0$ within $\mathbb{O}$ are invertible.
    (e) Show that the inverse operators are left and right multiplication by $\|a\|^{-2}a^*$.
    (f) Denote the usual basis vectors of $\mathbb{H}$ by $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$. Write down a multiplication table for the eight basis vectors of $\mathbb{O}$ given by $(x, 0)$ and $(0, y)$ as x and y run through the basis vectors of $\mathbb{H}$.
56. What prevents the construction of Problem 53, when applied with $A = \mathbb{O}$, from yielding a 16-dimensional algebra B in which $\|ab\|^2 = \|a\|^2\|b\|^2$ and therefore in which the operations of left and right multiplication by any $a \neq 0$ within B are invertible?