CHAPTER 4

Multilinear algebra

4.1. Tensor products

A multiplication law in the widest possible sense is a map from a product $X \times Y$ of two sets into a third set $Z$. This notion is so general as to be practically useless. However, useful mathematics can be done in situations where the sets $X$, $Y$ and $Z$ are equipped with some extra structure and the multiplication law respects that structure. For instance, in the case of vector spaces $U$, $V$ and $W$ over a field $F$, we will always mean by a multiplication law $m: U \times V \rightarrow W$ a bilinear map, i.e. a map that is linear in each variable:

$$m(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 m(u_1, v) + \lambda_2 m(u_2, v),$$
$$m(u, \lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 m(u, v_1) + \lambda_2 m(u, v_2),$$

for all $u_1, u_2 \in U$, $v_1, v_2 \in V$ and $\lambda_1, \lambda_2 \in F$. We have seen many examples of such multiplication laws.

(i) Scalar multiplication in any vector space over any field: $F \times V \rightarrow V$.

(ii) The polynomial algebra $F[x]$ and the formal power series algebra $F[[x]]$.

(iii) Extension fields of $F$, i.e. fields containing $F$ as a subfield, such as the rational functions $F(x)$ or the formal Laurent series $F((x))$.

(iv) The function algebras $F^S$ and $F^{(S)}$ for any set $S$.

(v) The cross product on $F^3$. (See Exercise 4.13.)

(vi) Real division algebras, for example the quaternions $\mathbb{H}$. (See Exercises 1.19 and 3.7.)

(vii) Matrix multiplication: $M_{m,n}(F) \times M_{n,p}(F) \rightarrow M_{m,p}(F)$.

(viii) Bilinear forms: $V \times V \rightarrow F$. (See Section 1.9.)

(ix) Sesquilinear forms: $\bar{V} \times V \rightarrow F$. (Here $F = \mathbb{C}$; see Exercises 4.1 and 4.2.)

In Example (i) we have $U = F$ and $V = W$. In Examples (ii)–(vi) we have $U = V = W$. In Example (vii) the three spaces $U$, $V$ and $W$ are distinct (unless $m = n = p$). In Examples (viii)–(ix) the target space is $W = F$, the ground field.

In this section we will construct a “universal” target space from which all multiplication laws defined on $U \times V$ can be derived. It is called the tensor product of $U$ and $V$ and is denoted by $U \otimes V$ or $U \otimes_F V$. Elements of $U \otimes V$ are called tensors. The multiplication law $U \times V \rightarrow U \otimes V$ is known as tensor multiplication and is denoted by $\otimes$.

The idea behind the construction is to form a huge vector space $E$ consisting of all formal linear combinations of ordered pairs $(u, v)$ with $u \in U$ and $v \in V$, and
to pare this space down by imposing bilinear relations. Let
\[ E = \mathbb{F}(U \times V) \]
be the free \( \mathbb{F} \)-vector space spanned by the set \( U \times V \). From Exercise 1.8 we know that \( E \) has a basis \( \mathcal{B} \) consisting of all elements of the form \( e_{u,v} \),
\[ \mathcal{B} = \{ e_{u,v} \mid u \in U, v \in V \}. \]
Thus \( E \) is monstrously large: its dimension is equal to the cardinality of \( U \times V \). (Don’t confuse \( E \) with the product \( U \times V \), which is far smaller.) To cut it down to the right size, we must identify certain elements of \( E \) with one another, that is to say take a quotient by a subspace \( F \). An element of \( E \) can be written uniquely as an \( \mathbb{F} \)-linear combination
\[ \sum_{u \in U, v \in V} \lambda_{u,v} e_{u,v}, \]
where \( \lambda_{u,v} \in \mathbb{F} \) is 0 for all but finitely many \( u \) and \( v \). We would like to think of \( e_{u,v} \) as a product \( u \otimes v \) depending bilinearly on \( u \) and \( v \), and so we wish to impose the rules
\[
\begin{align*}
\quad \quad e_{\lambda_1 u_1 + \lambda_2 u_2, v} &= \lambda_1 e_{u_1,v} + \lambda_2 e_{u_2,v}, \\
\quad e_{u, \lambda_1 v_1 + \lambda_2 v_2} &= \lambda_1 e_{u,v_1} + \lambda_2 e_{u,v_2},
\end{align*}
\]  
for all \( u, u_1, u_2 \in U, v, v_1, v_2 \in V \) and \( \lambda_1, \lambda_2 \in \mathbb{F} \). These rules don’t hold because the \( e_{u,v} \)'s are all independent. But we can enforce the rules by identifying both sides in each of the identities (4.1) in the following manner. Define \( A \) to be the subset of \( E \) consisting of all elements of the form
\[
\begin{align*}
e_{\lambda_1 u_1 + \lambda_2 u_2, v} - \lambda_1 e_{u_1,v} - \lambda_2 e_{u_2,v}, \\
e_{u, \lambda_1 v_1 + \lambda_2 v_2} - \lambda_1 e_{u,v_1} - \lambda_2 e_{u,v_2},
\end{align*}
\]  
with \( u, u_1, u_2 \in U, v, v_1, v_2 \in V \) and \( \lambda_1, \lambda_2 \in \mathbb{F} \). Let \( F \) be the span of \( A \) and define \( U \otimes V = E/F \). Elements of \( U \otimes V \) are affine subspaces \( \tilde{x} = x + F \), where \( x \in E \). If \( x \) is a basis element \( e_{u,v} \), let us write \( u \otimes v \) instead of \( e_{u,v} \). If \( x \) is an element of the form (4.2), then \( x \in F \), so \( \tilde{x} = \tilde{0} \). In other words,
\[
\begin{align*}
\quad \quad [\lambda_1 u_1 + \lambda_2 u_2] \otimes v &= \lambda_1 [u_1 \otimes v] + \lambda_2 [u_2 \otimes v], \\
u \otimes [\lambda_1 v_1 + \lambda_2 v_2] &= \lambda_1 [u \otimes v_1] + \lambda_2 [u \otimes v_2].
\end{align*}
\]
Let \( \otimes : U \times V \to U \otimes V \) be the map that sends a pair \((u, v)\) to \( u \otimes v \). Then (4.3) can be restated as saying that \( \otimes \) is a bilinear multiplication law.

Our first result says that the dimension of the tensor product is no greater than the product of the dimensions of the two factors. (We will improve on this result in Theorem 4.4.)

4.1. Lemma. Let \( R \) be a spanning set of \( U \) and \( S \) a spanning set of \( V \). Then the set \( \mathcal{T} = \{ u \otimes v \mid u \in R, v \in S \} \) spans \( U \otimes V \). Hence \( \dim(U \otimes V) \leq \dim(U) \dim(V) \). In particular, if \( U \) and \( V \) are finite-dimensional, then so is \( U \otimes V \).
4.1. TENSOR PRODUCTS

**Proof.** Observe that, since the collection of all \( e_{u,v} \)'s form a basis of \( E \) and the quotient map \( E \to E/F = U \otimes V \) is surjective, the collection of all \( u \otimes v \) spans \( U \otimes V \). Therefore every element of the tensor product is a linear combination of the form \( \sum_{u \in U, v \in V} \lambda_{u,v} |u \otimes v| \), where only finitely many \( \lambda_{u,v} \) are nonzero. Expanding \( u = \sum_{i=1}^{m} \mu_{i} u_{i} \) and \( v = \sum_{j=1}^{n} \nu_{j} v_{j} \) in terms of the spanning vectors \( u_{i} \in U \) and \( v_{j} \in V \) we get

\[
\sum_{u \in U, v \in V} \lambda_{u,v} |u \otimes v| = \sum_{u \in U, v \in V} \lambda_{u,v} \sum_{i=1}^{m} \sum_{j=1}^{n} \mu_{i} \nu_{j} |u_{i} \otimes v_{j}|,
\]

which is in the span of \( Q \).

The product of two cardinals \( |X| \) and \( |Y| \) is by definition the cardinal \( |X \times Y| \).

If we choose \( R \) and \( S \) to be bases of \( U \), resp. \( V \), then we see that \( U \otimes V \) has a spanning set \( J \) of cardinality at most \( |R \times S| \). Hence \( \dim U \otimes V \leq |J| \leq |R \times S| = (\dim U)(\dim V) \).

QED

4.2. **Theorem (universal property of the tensor product).** Let \( m: U \times V \to W \) be any bilinear map. Then there exists a unique linear map \( \tilde{m}: U \otimes V \to W \) such that \( m(u,v) = \tilde{m}(u \otimes v) \) for all \( u \in U \) and \( v \in V \).

\[
\begin{array}{ccc}
U \times V & \xrightarrow{m} & W \\
\otimes & \quad \downarrow & \\
U \otimes V & \xrightarrow{\tilde{m}} & W
\end{array}
\]

**Proof.** Suppose \( \tilde{m}_{1} \) and \( \tilde{m}_{2} \) are two maps from \( U \otimes V \) to \( W \) satisfying the requisite properties. Since every element of \( U \otimes V \) is of the form \( \sum_{i=1}^{n} \lambda_{i} (u_{i} \otimes v_{i}) \), we have

\[
\tilde{m}_{1}(\xi) = \sum_{i=1}^{n} \lambda_{i} \tilde{m}_{1}(u_{i} \otimes v_{i}) = m(u,v) = \sum_{i=1}^{n} \lambda_{i} \tilde{m}_{2}(u_{i} \otimes v_{i}) = \tilde{m}_{2}(\xi),
\]

so \( \tilde{m}_{1} = \tilde{m}_{2} \), which proves uniqueness.

To prove existence, define a map \( M: B \to W \) by \( M(e_{u,v}) = m(u,v) \) for \( u \in U \) and \( v \in V \). Since \( B \) is a basis of \( E \), \( M \) extends in a unique way to a linear map \( E \to W \), which we shall also denote by \( M \). From the bilinearity of \( m \) and the linearity of \( M \) we get

\[
M(e_{\lambda_{1} u_{1} + \lambda_{2} u_{2}, v}) = M(\lambda_{1} u_{1} + \lambda_{2} u_{2}, v) = \lambda_{1} m(u_{1}, v) + \lambda_{2} m(u_{2}, v)
= \lambda_{1} M(e_{u_{1}, v}) + \lambda_{2} M(e_{u_{2}, v}) = M(\lambda_{1} e_{u_{1}, v} + \lambda_{2} e_{u_{2}, v}),
\]

and hence

\[
M(e_{\lambda_{1} u_{1} + \lambda_{2} u_{2}, v} - \lambda_{1} e_{u_{1}, v} - \lambda_{2} e_{u_{2}, v}) = 0.
\]
Similarly
\[ M(\mathbf{e}_u, \lambda_1 v_1 + \lambda_2 v_2) = -\lambda_1 e_{u,v_1} - \lambda_2 e_{u,v_2} = 0. \]
In other words, \( M(x) = 0 \) for all \( x \in \mathbb{F} \), or \( \ker M \) contains the subspace \( F \). This implies that the map \( m : E/F \to W \) defined by \( m(\bar{x}) = M(x) \) is a well-defined linear map. By construction \( m \) satisfies \( \hat{m}(\bar{e}_{u,v}) = \hat{m}(\bar{e}_{u,v}) = m(u,v) \). QED

The message of the next result is that it didn’t matter how we constructed the tensor product, because it is uniquely determined by the universal property. (Of course the construction is still important, because it establishes the existence of an object satisfying the universal property.) This style of theorem is called category theory or general abstract nonsense.

4.3. Theorem. Let \( \hat{\otimes} : U \times V \to U \hat{\otimes} V \) be another bilinear map satisfying the universal property of Theorem 4.2. Then there exists a unique isomorphism \( f : U \otimes V \to U \hat{\otimes} V \) such that \( f(u \otimes v) = u \hat{\otimes} v \) for all \( u \in U \) and \( v \in V \).

Proof. Applying the universal property for \( U \otimes V \) to the bilinear map \( \hat{\otimes} \) and the space \( W = U \otimes V \),
\[
\begin{array}{ccc}
U \times V & \longrightarrow & U \otimes V \\
\otimes & \downarrow & \otimes \\
U \otimes V, & \quad f & \\
\end{array}
\]
we find a unique linear map \( f : U \otimes V \to U \hat{\otimes} V \) satisfying \( u \hat{\otimes} v = f(u \otimes v) \). Interchanging the roles of \( \otimes \) and \( \hat{\otimes} \) we find a linear map \( g : U \hat{\otimes} V \to U \otimes V \) satisfying \( u \otimes v = g(u \hat{\otimes} v) \). The composition \( h = g \circ f \) is a linear map from \( U \otimes V \) to itself satisfying \( h(u \otimes v) = u \otimes v \). The identity map \( \text{id}_{U \otimes V} \) also satisfies \( \text{id}_{U \otimes V}(u \otimes v) = u \otimes v \). Hence, by the uniqueness part of the universal property for \( U \otimes V \), applied to the bilinear map \( \hat{\otimes} \) and the space \( W = U \otimes V \),
\[
\begin{array}{ccc}
U \times V & \longrightarrow & U \otimes V \\
\otimes & \downarrow & \otimes \\
U \otimes V, & \quad h = \text{id}_{U \otimes V} & \\
\end{array}
\]
we get \( g \circ f = h = \text{id}_{U \otimes V} \). Similarly, \( f \circ g = \text{id}_{U \hat{\otimes} V} \), so \( f \) is an isomorphism. QED

Here is an application of the universal property: let \( k \in U^* \) and \( l \in V^* \) be covectors. Then for each \( u \in U \) and \( v \in V \) we have scalars \( k(u) \) and \( l(v) \), so using multiplication \( F \) we can define a map \( m : U \times V \to F \) by \( m(u,v) = k(u)l(v) \). This map is clearly bilinear, so there exists a unique linear map \( \hat{m} : U \otimes V \to F \) satisfying \( \hat{m}(u \otimes v) = k(u)l(v) \). We will denote the map \( \hat{m} \), which depends on \( k \) and \( l \), by \( T(k,l) \). By construction it is a covector on \( U \otimes V \text{ : } T(k,l) \in (U \otimes V)^* \).

This observation enables us to construct a basis of \( U \otimes V \) starting from bases \( A = \{ u_i \mid i \in I \} \) of \( U \) and \( B = \{ v_j \mid j \in J \} \) of \( V \). As in Section 1.8 define the coordinate functionals \( u_i^* \in U^* \) by \( u_i^*(u_{i'}) = \delta_{ii'} \), and \( v_j^* \in V^* \) by \( v_j^*(v_{j'}) = \delta_{jj'} \). Then
\[
T(u_i^*, v_j^*) (u_{i'} \otimes v_{j'}) = u_i^*(u_{i'}) v_j^*(v_{j'}) = \delta_{ii'} \delta_{jj'}. \tag{4.4}
\]
4.4. Theorem. Let \( A = \{ u_i \mid i \in I \} \) be a basis of \( U \) and \( B = \{ v_j \mid j \in J \} \) a basis of \( V \). Then \( C = \{ u_i \otimes v_j \mid (i, j) \in I \times J \} \) is a basis of \( U \otimes V \). In particular, \( \dim U \otimes V = (\dim U)(\dim V) \).

Proof. From Lemma 4.1 we know that \( C \) spans the tensor product, so it remains to show that \( C \) is independent. Suppose we have a dependence relation
\[
\sum_{i', j' \in J} \lambda_{i', j'} (u_{i'} \otimes v_{j'}) = 0.
\]
Applying the functional \( \Gamma [u^*_i, v^*_j] \) to both sides and using (4.4) we find
\[
0 = \sum_{i', j' \in J} \lambda_{i', j'} \Gamma [u^*_i, v^*_j] (u_{i'} \otimes v_{j'}) - \sum_{i', j' \in J} \lambda_{i', j'} \delta_{i', i} \delta_{j', j} - \lambda_{i, j}
\]
for all \( i \) and \( j \), so \( C \) is independent. QED

There is an obvious generalization of this theory to tensor products of \( q \) factors. Given vector spaces \( V_1, V_2, \ldots, V_q, W \) over \( F \), a map
\[
m: \prod_{p=1}^{q} V_p = V_1 \times V_2 \times \cdots \times V_q \to W
\]
is \( q \)-multilinear or simply multilinear if
\[
m(v_1, v_2, \ldots, \lambda v_p + \lambda' v'_p, \ldots, v_q) = \lambda m(v_1, v_2, \ldots, v_p, \ldots, v_q) + \lambda' m(v_1, v_2, \ldots, v'_p, \ldots, v_q)
\]
for all \( \lambda, \lambda' \in F \) and \( v_1 \in V_1, v_2 \in V_2, \ldots, v_p, v'_p \in V_p, \ldots, v_q \in V_q \). (For \( q = 1 \) this is the familiar notion of a linear map. For \( q = 0 \) we define \( \prod_{p=1}^{0} V_p = F \).

A 0-multilinear map is by definition the zero map \( F \to W \). The tensor product \( \otimes_{p=1}^{q} V_p = V_1 \otimes V_2 \otimes \cdots \otimes V_q \) is the vector space \( E/F \), where
\[
E = F(\prod_{p=1}^{q} V_p) = F(V_1 \times V_2 \times \cdots \times V_q),
\]
and where \( F \) is the subspace of all "multilinear relations", i.e. the subspace of \( E \) spanned by all expressions of the form
\[
e_{v_1, v_2, \ldots, \lambda v_p + \lambda' v'_p, \ldots, v_q} = \lambda e_{v_1, v_2, \ldots, v_p, \ldots, v_q} - \lambda_2 e_{v_1, v_2, \ldots, v'_p, \ldots, v_q}.
\]
As before we write \( v_1 \otimes v_2 \otimes \cdots \otimes v_q \) for the element \( e_{v_1, v_2, \ldots, v_q} \) of the tensor product, and we define the multilinear map \( \otimes: \prod_{p=1}^{q} V_p \to \otimes_{p=1}^{q} V_p \) by
\[
\otimes[v_1, v_2, \ldots, v_p, \ldots, v_q] = v_1 \otimes v_2 \otimes \cdots \otimes v_q.
\]

Here are the multilinear analogues of Theorems 4.2–4.4.

4.5. Theorem. Let \( m: \prod_{p=1}^{q} V_p \to W \) be any multilinear map. Then there exists a unique linear map \( \tilde{m}: \otimes_{p=1}^{q} V_p \to W \) such that
\[
m(v_1, v_2, \ldots, v_q) = \tilde{m}(v_1 \otimes v_2 \otimes \cdots \otimes v_q)
\]
for all \( v_p \in V_p \) and \( p = 1, 2, \ldots, q \).

\[
\begin{array}{c}
\prod_{p=1}^{q} V_p \\
\otimes \\
\otimes_{p=1}^{q} V_p
\end{array}
\]

4.6. Theorem. Let \( \hat{\otimes} : \prod_{p=1}^{q} V_p \to \hat{\otimes}_{p=1}^{q} V_p \) be another multilinear map satisfying the universal property of Theorem 4.5. Then there exists a unique isomorphism \( f : \hat{\otimes}_{p=1}^{q} V_p \to \prod_{p=1}^{q} V_p \) such that

\[
f(v_1 \otimes v_2 \otimes \cdots \otimes v_q) = v_1 \hat{\otimes} v_2 \hat{\otimes} \cdots \hat{\otimes} v_q
\]

for all \( v_p \in V_p \) and \( p = 1, 2, \ldots, q \).

4.7. Theorem. Let \( B(p) = \{v_i | i \in I(p)\} \) be a basis of \( V_p \) for \( p = 1, 2, \ldots, q \). Then

\[
\hat{e} = \{v_{i(1)} \otimes v_{i(2)} \otimes \cdots \otimes v_{i(q)} | i(1) \in I(1), i(2) \in I(2), \ldots, i(q) \in I(q)\}
\]

is a basis of \( \hat{\otimes}_{p=1}^{q} V_p \). In particular, \( \dim \hat{\otimes}_{p=1}^{q} V_p = \prod_{p=1}^{q} \dim V_p \).

A useful side benefit of Theorem 4.6 is the following result.

4.8. Corollary (associativity of tensor products). For all \( k \geq 0 \) and \( l \geq 0 \) such that \( k + l = q \) there exist unique isomorphisms

\[
f_{k,l} : \left( \bigotimes_{p=1}^{k} V_p \right) \otimes \left( \bigotimes_{p=k+1}^{q} V_p \right) \to \bigotimes_{p=1}^{q} V_p
\]

satisfying

\[
f_{k,l}((v_1 \otimes v_2 \otimes \cdots \otimes v_k) \otimes (v_{k+1} \otimes v_{k+2} \otimes \cdots \otimes v_q)) = v_1 \otimes v_2 \otimes \cdots \otimes v_k \otimes v_{k+1} \otimes v_{k+2} \otimes \cdots \otimes v_q
\]

for all \( v_1 \in V_1, v_2 \in V_2, \ldots, v_q \in V_q \).

Proof. See Exercise 4.8.

QED

4.2. Algebras

Among the most interesting multiplication laws \( U \times V \to W \) are those for which \( U = V = W \). These deserve a special name: a vector space \( V \) over \( F \) equipped with a bilinear map \( m: V \times V \to V \) is called an algebra over \( F \). In an algebra one frequently abbreviates \( m(x, y) \) to \( xy \) and calls this the product of \( x \) and \( y \). (Notations such as \( x \otimes y \), \( x \times y \), \( x \wedge y \) or \( [x, y] \) are also commonly used in special types of algebras.) Examples (ii)–(vi) in Section 4.1 are algebras, and so is Example (vii) if \( m = n = p \). An algebra \( A \) is associative if \( (xy)z = x(yz) \) for all \( x, y \) and \( z \in A \). It is commutative if \( xy = yx \) for all \( x \) and \( y \in A \). It is unital if
it possesses a two-sided unit 1 satisfying 1x = x1 = x for all x ∈ A. The following
table shows which of the Examples (ii)–(vii) in Section 4.1 enjoy these properties.

<table>
<thead>
<tr>
<th>algebra</th>
<th>associative</th>
<th>commutative</th>
<th>unital</th>
</tr>
</thead>
<tbody>
<tr>
<td>F[x], F[x], F[x], F[x]</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>F^S</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>F(S)</td>
<td>yes</td>
<td>yes</td>
<td>no (unless S finite)</td>
</tr>
<tr>
<td>(F^3, ×)</td>
<td>no</td>
<td>no (unless char F = 2)</td>
<td>no</td>
</tr>
<tr>
<td>H</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>M_n(F)</td>
<td>yes</td>
<td>no (unless n = 1)</td>
<td>yes</td>
</tr>
</tbody>
</table>

4.3. Tensor algebras

Clearly the tensor product U × V → U ⊗ V is not an algebra, but by letting
U = V and taking repeated tensor products we can manufacture an algebra out
of any vector space V. The first step is to define, for any integer k ≥ 0, the k-th
tensor power of a vector space V. This is the vector space

\[ T^k(V) = \bigotimes_{j=1}^{k} V = V \otimes V \otimes \cdots \otimes V. \]

We have T^0(V) = F and T^1(V) = V. An element of T^k(V) is called a tensor of
degree k. We can multiply degree k tensors by degree l tensors to get degree k+l
tensors.

4.9. **Lemma.** For k ≥ 0 there exists a unique bilinear map \( m_{k,1} : T^k(V) \times T^1(V) \to T^{k+1}(V) \) satisfying

\[ m_{k,1}(v_1 \otimes v_2 \otimes \cdots \otimes v_k, w_1 \otimes w_2 \otimes \cdots \otimes w_l) = v_1 \otimes v_2 \otimes \cdots \otimes v_k \otimes w_1 \otimes w_2 \otimes \cdots \otimes w_l \]

for all \( v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_l \in V \).

**Proof.** Let \( E = T^k(V) \) and \( F = T^1(V) \). Then, by Corollary 4.8, there is a
unique isomorphism \( f_{k,1} : E \otimes F \to T^{k+1}(V) \) which sends

\[ \{ v_1 \otimes v_2 \otimes \cdots \otimes v_k \} \otimes \{ w_1 \otimes w_2 \otimes \cdots \otimes w_l \} \]

to \( v_1 \otimes v_2 \otimes \cdots \otimes v_k \otimes w_1 \otimes w_2 \otimes \cdots \otimes w_l \). By composing the linear map \( f_{k,1} \) with
the bilinear map \( \otimes : E \times F \to E \otimes F \) we find the desired bilinear map \( m_{k,1} \). The
elements of the form \( v_1 \otimes v_2 \otimes \cdots \otimes v_k \) span \( T^k(V) \), so \( m_{k,1} \) is uniquely determined
by the conditions imposed on it. QED

4.10. **Remark.** For \( k = 0 \) we have \( T^0(V) = F \). The bilinear maps

\( m_{0,1} : F \times T^1(V) = T^0(V) \times T^1(V) \to T^{0+1}(V) = T^1(V) \)

are given by scalar multiplication:

\( m_{0,1}(\lambda, w_1 \otimes w_2 \otimes \cdots \otimes w_l) = \lambda(w_1 \otimes w_2 \otimes \cdots \otimes w_l) \).
The second step in building an algebra out of $V$ is to put all of the multiplication laws $m_{k,l}$ together in one big space. The tensor algebra $T(V)$ of $V$ is defined as the direct sum

$$T(V) = \bigoplus_{k=0}^{\infty} T^k(V) = F \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots.$$ 

Elements of the tensor algebra are called tensors. They can be written as infinite tuples

$$\nu = (\nu^{(k)})_{k=0}^{\infty} = (\nu^{[0]}, \nu^{[1]}, \nu^{[2]}, \ldots, \nu^{[k]}, \ldots)$$

with $\nu^{[k]} \in T^k(V)$ and $\nu^{[k]} \neq 0$ for only finitely many $k$. Following the convention explained at the end of Exercise 3.4, we shall write $\nu = \sum_{k=0}^{\infty} \nu^{[k]}$. What is the multiplication law $m$ in the tensor algebra? We want $m$ to be bilinear, so for $\nu = \sum_{k=0}^{\infty} \nu^{[k]}$ and $\omega = \sum_{l=0}^{\infty} \omega^{[l]}$ we must define

$$m(\nu, \omega) = m\left( \sum_{k=0}^{\infty} \nu^{[k]}, \sum_{l=0}^{\infty} \omega^{[l]} \right) = \sum_{k,l=0}^{\infty} m(\nu^{[k]}, \omega^{[l]}).$$

To define the product of tensors of degree $k$ and of degree $l$ we resort to the multiplication law $m_{k,l}$ defined in Lemma 4.9: $m(\nu^{[k]}, \omega^{[l]}) = m_{k,l}(\nu^{[k]}, \omega^{[l]})$. This determines our multiplication law $m_1: T(V) \times T(V) \to T(V)$ and so turns $T(V)$ into an algebra. This multiplication is called tensor multiplication and denoted, again, by the symbol $\otimes$. (This notation may be confusing initially, but I assure you it makes sense and it helps avoid a surfeit of new symbols.) Vectors in $V = T^0(V)$ are tensors of degree $1$ and so can be regarded as elements of the tensor algebra. In this sense, the tensor algebra contains $V$ itself, and so we have enlarged the vector space $V$ to an algebra $T(V)$.

Let us now write an explicit formula for the multiplication law in $T(V)$. We start by choosing a basis $B = \{v_i \mid i \in I\}$ of $V$. (Here $I$ is some unspecified, not necessarily finite, set indexing the basis elements.) It follows from Theorem 4.7 that every tensor $\nu$ of degree $k$ can be written in a unique way as a linear combination

$$\nu = \sum_{i_1 \in I} \cdots \sum_{i_k \in I} \lambda_{i_1,i_2,\ldots,i_k} v_{i_1} \otimes v_{i_2} \otimes \cdots \otimes v_{i_k},$$

where $\lambda_{i_1,i_2,\ldots,i_k} \in F$ and $i_1, i_2, \ldots, i_k \in I$. To tidy up this burdensome notation, we introduce the notion of a multi-index of degree $k$, which is a $k$-tuple $i = (i_1, i_2, \ldots, i_k) \in I^k$ of elements of the index set $I$. We abbreviate $\lambda_{i_1,i_2,\ldots,i_k}$ to $\lambda_i$ and $v_{i_1} \otimes v_{i_2} \otimes \cdots \otimes v_{i_k}$ to $v_i^\otimes$. We can now present degree $k$ tensors more efficiently as

$$\nu = \sum_{i \in I^k} \lambda_i v_i^\otimes,$$

As always, for this sum to make sense only finitely many of the scalars $\lambda_i$ may be nonzero. These remarks can be summarized as follows.

4.11. Lemma. Let $B = \{v_i \mid i \in I\}$ be a basis of $V$. Then the collection

$$B^{\otimes k} = \{v_i^\otimes \mid i \in I^k\}$$

is a basis of $T^k(V)$. (Here we define $B^{\otimes 0} = \{1\}$ and $B^{\otimes 1} = B$.) In particular, $\dim T^k(V) = (\dim V)^k$. 


4.3. TENSOR ALGEBRAS

From Lemma 4.11 and Exercise 1.10 we now get a basis for the tensor algebra. Note that \( T(V) \) is not finite-dimensional (unless \( V = \{0\} \), in which case \( T(V) = F \).

4.12. Lemma. Let \( B = \{v_i \mid i \in I\} \) be a basis of \( V \). Then the collection
\[
\{1\} \cup B \cup B^2 \cup \cdots = \bigcup_{k=0}^{\infty} B^k = \bigcup_{k=0}^{\infty} \{v_i^k \mid i \in I^k\}
\]
is a basis of \( T(V) \).

Employing the notation (4.6), we can now write every tensor \( v \) in a unique way as a linear combination
\[
v = \sum_{k=0}^{\infty} \sum_{i \in I^k} \lambda_i v_i^k = \lambda_1 1 + \sum_{i \in I} \lambda_i v_i + \sum_{i_1, i_2 \in I} \lambda_{i_1, i_2} v_{i_1} \otimes v_{i_2} + \cdots
\]
(with finitely many nonzero terms). The rule for multiplying such a sum by another sum \( w = \sum_{j=0}^{\infty} \sum_{i_j \in I_j} \mu_j w_j^j \) is
\[
v \otimes w = \sum_{k=0}^{\infty} \sum_{i \in I^k} \sum_{j \in I^j} \lambda_i \mu_j (v_i^k \otimes w_j^j).
\]
In other words, the product is determined by bilinearity plus the “concatenation” rule
\[
[v_1 \otimes v_2 \otimes \cdots \otimes v_k] \otimes [w_1 \otimes w_2 \otimes \cdots \otimes w_l] = v_1 \otimes v_2 \otimes \cdots \otimes v_k \otimes w_1 \otimes w_2 \otimes \cdots \otimes w_l.
\]
Thus tensors are like polynomials in the “variables” \( v_i \in B \) with coefficients \( \lambda_i \in F \), except that we are not allowed to interchange two variables \( v_i \) and \( v_j \) if \( i \neq j \). For this reason \( T(V) \) is called the noncommutative polynomial algebra in the variables \( v_i \).

A homomorphism of algebras \( A \) and \( B \) over \( F \) is a map \( f: A \to B \) which is linear and multiplicative (\( f(xy) = f(x)f(y) \) for all \( x, y \in A \)). If \( A \) and \( B \) are unital with units \( 1_A \), resp. \( 1_B \), a homomorphism \( f: A \to B \) is unital if \( f(1_A) = 1_B \). The next result says that any linear map from \( V \) into an associative unital algebra can be extended to a unital algebra homomorphism defined on all of \( T(V) \).

4.13. Theorem (universal property of the tensor algebra). \( T(V) \) is an associative unital algebra. For every associative unital algebra \( A \) and every linear map \( f: V \to A \) there exists a unique unital homomorphism \( \tilde{f}: T(V) \to A \) satisfying \( \tilde{f}(V) = f \), i.e. \( \tilde{f}(v) = f(v) \) for \( v \in V \).

\[
\begin{array}{ccc}
V & \xrightarrow{f} & A \\
\cap & \downarrow & \downarrow \\
T(V) & \xrightarrow{\tilde{f}} & A
\end{array}
\]

Proof. The tensor algebra contains the scalars \( F = T^0(V) \). According to Remark 4.10, the unit \( 1 \in F \) also acts as a unit in \( T(V) \). Associativity of \( T(V) \) follows from the associativity of the tensor product, Corollary 4.8. Consider an
element \( v \) of \( T(V) \), written in the form (4.7). Suppose \( f : V \to A \) is linear and suppose \( \tilde{f} \) is a unital homomorphism \( \tilde{f} : T(V) \to A \) extending \( f \). Then
\[
\tilde{f}(v) = \tilde{f}\left( \sum_{i \in I} \lambda_i v_i + \sum_{i_1, i_2 \in I} \lambda_{i_1, i_2} v_{i_1} \otimes v_{i_2} + \cdots \right)
\]
\[
= \lambda_1 \tilde{f}(1) + \sum_{i \in I} \lambda_i \tilde{f}(v_i) + \sum_{i_1, i_2 \in I} \lambda_{i_1, i_2} \tilde{f}(v_{i_1} \otimes v_{i_2}) + \cdots
\]
\[
= \lambda_1 f(1) + \sum_{i \in I} \lambda_i f(v_i) + \sum_{i_1, i_2 \in I} \lambda_{i_1, i_2} f(v_{i_1}) f(v_{i_2}) + \cdots
\]
\[
= \sum_{k=0}^{\infty} \sum_{i_1, \ldots, i_k} \lambda_{i_1} \cdots \lambda_{i_k} f(v_{i_1} \otimes \cdots \otimes v_{i_k}), \quad (4.9)
\]
so \( \tilde{f} \) is determined by \( f \). To show that \( \tilde{f} \) exists, one defines it by the formula (4.9) and checks that it has the requisite properties.

**4.4. Ideals and quotients**

Just as we can create a new vector space out of a vector space \( E \) by taking a quotient \( E/F \) by a subspace \( F \), we can form a new algebra out of an algebra \( A \) by taking a quotient by a subspace \( J \). However, to ensure that the multiplication law on \( A \) engenders a new multiplication law on \( A/J \), we must impose certain conditions on \( J \). Let \( A \) be an algebra over \( F \). A (two-sided) ideal of \( A \) is a linear subspace \( J \) such that \( xy \in J \) and \( yx \in J \) for all \( x \in A \) and \( y \in J \).

4.14. **Example.** The annihilator \( I_f \) of a linear map \( f : V \to V \) is an ideal in the polynomial algebra \( F[x] \). (See Lemma 3.25.)

4.15. **Example.** Let \( X \) be any nonempty subset of an algebra \( A \). Define
\[
(X) = \text{span}\{ axb | x \in X, a, b \in A \}.
\]
Then \( (X) \) is an ideal of \( A \), called the ideal generated by \( X \). (See Exercise 4.14.) If \( A \) is commutative, we can interchange \( x \) and \( b \) to get
\[
(X) = \text{span}\{ ax | x \in X, a \in A \}.
\]
If \( X \) consists of a single point \( x \), we write \( (x) \) instead of \( (X) \). For instance, Lemma 3.25 says that \( I_f = (m_f) \), the ideal generated by the minimal polynomial \( m_f \) of \( f \).

4.16. **Theorem.** Let \( A \) be an algebra with multiplication law \( m : A \times A \to A \). Let \( J \) be an ideal in \( A \). Let \( B \) be the quotient space \( A/J \) and let \( pr : A \to B \) be the projection map. Then there exists a unique multiplication law \( \tilde{m} : B \times B \to B \) such that \( \tilde{m}(pr(x), pr(y)) = pr(m(x, y)) \) for all \( x, y \in A \). If \( m \) is associative, unital, or commutative, then \( \tilde{m} \) has the same property.

**Proof.** For convenience let us write \( \bar{x} \) instead of \( pr(x) \). By Lemma 1.2, \( \bar{x} = \bar{x}' \) if and only if \( x - x' \in J \). The only way to meet the requirements on \( \tilde{m} \) is to define \( \tilde{m}(\bar{x}, \bar{y}) = m(x, y) \). We need to check that \( \tilde{m} \) is well-defined and bilinear. Suppose
\[ \tilde{x} = \tilde{x}' \text{ and } \tilde{y} = \tilde{y}'. \] Is \( \tilde{m}(x, y) = \tilde{m}(x', y') \)? Write \( x' = x + x'' \) and \( y' = y + y'' \) with \( x'', y'' \in J \). By bilinearity of \( m \),

\[
m(x', y') = m(x + x'', y + y'') = m(x, y) + m(x'', y) + m(x', y'')
\]

with \( z = m(x, y'') + m(x'', y) + m(x', y'') \). Since \( J \) is an ideal, \( z \in J \) and so \( \tilde{m}(x, y) = \tilde{m}(x', y') \). We conclude that \( \tilde{m}(\tilde{x}, \tilde{y}) = \tilde{m}(\tilde{x}', \tilde{y}') \), so \( \tilde{m} \) is well-defined. The bilinearity of \( m \) also follows from the bilinearity of \( m \). (See Exercise 4.15.) The last assertion of the theorem is easy to check by using the definition of \( \tilde{m} \). QED

We call the algebra \( B \) equipped with the multiplication law \( \tilde{m} \) the quotient algebra of \( A \) by \( J \). As usual, we will frequently write \( \tilde{x} \tilde{y} \) instead of \( \tilde{m}(\tilde{x}, \tilde{y}) \).

Using ideals and quotient algebras, in Section 4.5 we shall construct an algebra of ordinary, i.e. commutative, polynomials, and in Section 4.6 we shall construct an algebra of anticommutative polynomials.

### 4.5. Symmetric algebras

The symmetric algebra \( S(V) \) of a vector space \( V \) is a commutative analogue of the tensor algebra. It is obtained from the tensor algebra by “making the multiplication commutative”, in other words by imposing the rule \( v \otimes w = w \otimes v \) for all \( v, w \in V \). This rule does not hold in the tensor algebra, but we can manufacture an algebra in which it holds by taking a quotient of \( T(V) \) by an appropriate ideal. Let \( X \) be the subset of \( T(V) \) consisting of all degree 2 tensors of the form

\[
v \otimes w - w \otimes v,
\]

where \( v \) and \( w \) are arbitrary vectors in \( V \), and let \( J = \{ X \} \) be the ideal generated by \( X \). The symmetric algebra \( S(V) \) of \( V \) is declared to be the quotient algebra \( T(V)/J \). Elements of \( S(V) \) are symmetric tensors. The product of symmetric tensors \( x \) and \( y \) is written as \( x \circ y \).

Just like the tensor algebra, the symmetric algebra breaks up into pieces of different degrees. To see how, let us put \( J_k = J \cap T^k(V) \). Since \( X \subseteq T^2(V) \), we have \( J_0 = J_1 = \emptyset \).

#### 4.17. Lemma

The ideal \( J = \bigoplus_{k=2}^{\infty} J_k \) is generated by \( J' \). By definition, every \( v \) in \( J' \) is a finite sum \( \sum_{k=2}^{\infty} v^{(k)} \) with \( v^{(k)} \) in \( J_k \). Since \( J_k \subseteq J \) and \( J \) is a subspace, we get \( v \in J \). Now let \( v \in J \). By the definition of \( J \) we can write

\[
v = \sum_{p=1}^{\infty} \lambda_p a_p \otimes (v_p \otimes w_p - w_p \otimes v_p) \otimes b_p
\]

for certain \( \lambda_p \in F, v_p, w_p \in V, \) and \( a_p, b_p \in T(V) \). The \( p \)-th term in this expression is of degree \( k \) if and only if \( \deg a_p + \deg b_p = k - 2 \). Thus the component of \( v \) of degree \( k \) is

\[
\sum_{p=1,2,\ldots,q} \lambda_p a_p \otimes (v_p \otimes w_p - w_p \otimes v_p) \otimes b_p,
\]

where \( a_p, b_p \) have degree \( k - 2 \).
which is an element of $J_k$. Hence $v \in J'$.

QED

It follows that the symmetric algebra is a direct sum,

$$S(V) = T(V)/J = \left( \bigoplus_{k=0}^{\infty} T^k(V) \right)/\left( \bigoplus_{k=0}^{\infty} J_k \right) = \bigoplus_{k=0}^{\infty} T^k(V)/J_k$$

$$= F \oplus V \oplus T^2(V)/J_2 \oplus T^3(V)/J_3 \oplus \cdots.$$

The $k$-th summand $T^k(V)/J_k$ of the symmetric algebra is called the $k$-th symmetric power of $V$ and denoted by $S^k(V)$. In particular, $S^0(V) = T^0(V) = F$ and $S^1(V) = T^1(V) = V$, so $S(V)$ contains a copy of the scalars $F$ and of the vector space $V$. Moreover, since $T^k(V)$ is spanned by elements of the form $v_1 \otimes v_2 \otimes \cdots \otimes v_k$ with $v_1, v_2, \ldots, v_k \in V$, and the projection map $\text{pr}: T^k(V) \to S^k(V)$ is surjective for each $k$, we see that $S^k(V)$ is spanned by elements of the form

$$\text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \text{pr}(v_1) \cdot \text{pr}(v_2) \cdots \text{pr}(v_k) = v_1 v_2 \cdots v_k,$$

in other words,

$$S^k(V) = \text{span}\{v_1 v_2 \cdots v_k \mid v_1, v_2, \ldots, v_k \in V\}. \quad (4.10)$$

4.18. **Theorem** (universal property of the symmetric algebra). $S(V)$ is an associative commutative unital algebra. For every associative commutative unital algebra $A$ and every linear map $f: V \to A$ there exists a unique unital homomorphism $\tilde{f}: S(V) \to A$ satisfying $\tilde{f}(V) = f$.

**Proof.** Associativity and existence of a unit follow from the corresponding properties of the tensor algebra and from Theorem 4.16. If $v_1, v_2, \ldots, v_k \in V$, then $\text{pr}(v_1 \otimes v_j - v_i \otimes v_j) = 0$, where $\text{pr}: T(V) \to S(V)$ is the projection map, so

$$\text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \text{pr}(v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(k)})$$

for every permutation $\sigma$ of $k$ elements. In particular,

$$\text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) \cdot \text{pr}(w_1 \otimes w_2 \otimes \cdots \otimes w_l)$$

$$= \text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k \otimes w_1 \otimes w_2 \otimes \cdots \otimes w_l)$$

$$= \text{pr}(w_1 \otimes w_2 \otimes \cdots \otimes w_l) \cdot \text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k).$$

By (4.10), the elements of the form $v_1 v_2 \cdots v_k$ span $S(V)$.

Hence, by bilinearity, $S(V)$ is commutative.

The universal property of the tensor product gives, for every linear map $f: V \to A$, a unital algebra homomorphism $\tilde{f}: T(V) \to A$ extending $f$. The algebra $A$ being commutative, the kernel of $\tilde{f}$ contains the ideal $J$. Hence, by Exercise 1.15, $\tilde{f}$ descends to a unique linear map $\tilde{f}: S(V) = T(V)/J \to A$. From the fact that $\tilde{f}$
is a unital algebra homomorphism one deduces easily that \( \tilde{f} \) is a unital algebra homomorphism as well. It follows from

\[
\tilde{f}(v_1 v_2 \cdots v_k) = \tilde{f}(v_1) \tilde{f}(v_2) \cdots \tilde{f}(v_k) = f(v_1) f(v_2) \cdots f(v_k)
\]

and (4.10) that \( \tilde{f} \) is uniquely determined by \( f \).  

QED

Now put

\[
V^k = \prod_{j=1}^{k} V = V \times V \times \cdots \times V, \quad \text{k times}
\]

A multilinear map \( f : V^k \to W \) is symmetric if

\[
f(v_1, v_2, \ldots, v_k) = f(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)})
\]

for every permutation \( \sigma \) of \( k \) elements. Let \( m : V^k \to S^k(V) \) be the multiplication map defined by \( m(v_1, v_2, \ldots, v_k) = v_1 v_2 \cdots v_k \). The following result can be deduced from Theorem 4.5 in the same way that we deduced Theorem 4.18 from Theorem 4.13.

4.19. **Theorem** (universal property of the k-th symmetric power). \( m \) is symmetric multilinear for all \( k \geq 0 \). For every vector space \( W \) over \( F \) and every symmetric multilinear map \( f : V^k \to W \) there exists a unique linear map \( \tilde{f} : S^k(V) \to W \) such that \( \tilde{f} \circ m = f \).

**Proof.** See Exercise 4.16.  

QED

We can use this to describe in more detail each of the symmetric powers \( S^k(V) = T^k(V)/\mathcal{I}_k \) in terms of a basis \( \mathcal{B} = \{ v_i \mid i \in I \} \) of \( V \). By means of (4.10) we can express any symmetric tensor \( \nu \) of degree \( k \) as

\[
 \nu = \sum_{i_1 \in I} \sum_{i_2 \in I} \cdots \sum_{i_k \in I} \lambda_{i_{1}, i_{2}, \ldots, i_{k}} v_{i_{1}} v_{i_{2}} \cdots v_{i_{k}}. \tag{4.11}
\]

Because \( S(V) \) is commutative, in each of the terms of this expression we can lump repeating factors together. Thus, if a basis vector \( v_i \) occurs \( d_i \) times in a product \( v_{i_{1}} v_{i_{2}} \cdots v_{i_{k}} \), we can take out a factor \( v_i^{d_i} \). Doing this for each \( i \) in the index set \( I \), we can rewrite \( v_{i_{1}} v_{i_{2}} \cdots v_{i_{k}} \) in the form \( \prod_{i \in I} v_i^{d_i} \). Such a product \( \prod_{i \in I} v_i^{d_i} \) is called a *monomial* in the basis elements and is abbreviated to \( \nu^d \). Here \( d = (d_i)_{i \in I} \) stands for the \( I \)-tuple of exponents occurring in the product, which is known as the **multidegree** of the monomial. Formally speaking, \( d \) is a function \( I \to \mathbb{N} \) such that \( d_i = 0 \) for all but finitely many \( i \in I \); in other words \( d \in \mathbb{N}^{|I|} \). The **total degree** \( k \), i.e. the total number of factors occurring in \( \nu^d \), is equal to the sum of the
exponents $\sum_{i \in I} d_i$, and is often denoted by $|d|$. Thus we can write a symmetric tensor $v$ in $S^k(V)$ as a \textit{homogeneous polynomial of degree} $k$,

$$v = \sum_{d \in \mathbb{N}^{[1]}} \frac{\lambda_d v^d}{|d| - k}$$

(4.12)

with \textit{finitely many nonzero coefficients} $\lambda_d \in \mathbb{F}$. Likewise, an \textit{arbitrary symmetric tensor} $v$ in $S(V)$ is a \textit{polynomial},

$$v = \sum_{d \in \mathbb{N}^{[1]}} \lambda_d v^d$$

(4.13)

with $\lambda_d \in \mathbb{F}$. We will now show that the coefficients $\lambda_d$ in these expressions are uniquely determined.

\textbf{4.20. THEOREM.} Let $\mathcal{B} = \{v_i \mid i \in I\}$ be a basis of $V$.

(i) For all $k \in \mathbb{N}$ the collection of monomials of degree $k$,

$$\mathcal{B}^{(k)} = \{v^d \mid d \in \mathbb{N}^{[1]}, |d| = k\},$$

is a basis of $S^k(V)$. Hence $\dim S^k(V) = \binom{n + k - 1}{n - 1}$, where $n = |I| = \dim V$. (Here we put $\mathcal{B}^{(0)} = \{1\}$ and $\mathcal{B}^{(1)} = \mathcal{B}$. We put $\binom{n + k - 1}{n - 1} = n$ if $n$ is infinite.)

(ii) The collection of all monomials,

$$\bigcup_{k=0}^{\infty} \mathcal{B}^{(k)} = \{v^d \mid d \in \mathbb{N}^{[1]}\},$$

is a basis of $S(V)$. Hence $\dim S(V)$ is countable if $V$ is finite-dimensional and $\dim S(V) = \dim V$ if $V$ is infinite-dimensional.

\textbf{PROOF.} The proof follows the pattern of the proof of Theorem 4.4. It follows from (4.12) that $\mathcal{B}^{(k)}$ spans $S^k(V)$. It remains to show that $\mathcal{B}^{(k)}$ is independent. Let $l_1, l_2, \ldots, l_q \in V^*$ be covectors on $V$ and, for each $p = 1, 2, \ldots, q$, let $d_p$ be a nonnegative integer such that $d_1 + d_2 + \cdots + d_q = k$. From these inputs we can manufacture a multilinear function $m' : V^k \to \mathbb{F}$, namely

$$m'(v_1, v_2, \ldots, v_k) = \prod_{p=1}^{q} \prod_{i=0}^{d_p} l_p(v_i)$$

$$= l_1(v_1) l_1(v_2) \cdots l_1(v_{d_1}) l_2(v_{d_1} + 1) l_2(v_{d_1 + 2}) \cdots l_2(v_{d_1 + d_2}) \cdots l_q(v_{d_1 + d_2 + \cdots + d_{q-1} + 1}) l_q(v_{d_1 + d_2 + \cdots + d_q}).$$

We want to have a symmetric function $m : V^k \to \mathbb{F}$ to which we can apply the universal property of $S^k(V)$. However, $m'$ is not symmetric under all permutations of the vectors $v_1, v_2, \ldots, v_k$. It is only symmetric under permutations of the vectors $v_1, v_2, \ldots, v_{d_1}$, and under permutations of the vectors $v_{d_1 + 1}, v_{d_1 + 2}, \ldots, v_{d_1 + d_2}$, etc. For this reason we \textit{symmetrize} $m'$:

$$m(v_1, v_2, \ldots, v_k) = \sum_{\sigma \in S_k / ((\prod_{p=1}^{q} S_{d_p}) \prod_{i=0}^{d_p} l_p(v_{\sigma(i)}))} \prod_{p=1}^{q} \prod_{i=0}^{d_p} l_p(v_{\sigma(i)}).$$

(4.14)
In order to avoid superfluous summands, we let this sum range not over all permutations, but only over the quotient $S_k/\langle S_{d_1} \times S_{d_2} \times \cdots \times S_{d_q} \rangle$. For instance, if $k$ is arbitrary, $q = k$, and $d_1 = d_2 = \cdots = d_q = 1$, we get a sum of $k!$ terms,

$$m(v_1, v_2, \ldots, v_k) = \sum_{\sigma \in S_k} l_1 (v_{\sigma(1)}) l_2 (v_{\sigma(2)}) \cdots l_k (v_{\sigma(k)}),$$

but if $k = 3$, $q = 2$, and $d_1 = 2$, $d_2 = 1$, we get a sum of $3/2! = 3$ terms,

$$m(v_1, v_2, v_3) = l_1 (v_1) l_1 (v_2) l_2 (v_3) + l_1 (v_1) l_1 (v_3) l_2 (v_2) + l_1 (v_2) l_1 (v_3) l_2 (v_1).$$

By construction, $m$ is symmetric and multilinear, so by the universal property of symmetric powers, Theorem 4.19, $m$ descends to a linear map $m: S^k(V) \to F$. The map $\hat{m}$ depends on the covectors $l_i \in V^*$ and on the integers $d_i \in \mathbb{N}$, and therefore we will denote it by $S(l_1, d_1; l_2, d_2; \ldots; l_q, d_q)$.

Now let $d \in \mathbb{N}^{(1)}$ be a tuple of degree $|d| = k$. Let $i_1, i_2, \ldots, i_q$ be the collection of all $i \in I$ such that $d_i \neq 0$. Then $k = |d| = d_{i_1} + d_{i_2} + \cdots + d_{i_q}$. Define a linear map $S_d: S^k(V) \to F$, i.e. a covector on $S^k(V)$, by

$$S_d = S(v_{j_1}^*, d_{i_1}; v_{j_2}^*, d_{i_2}; \ldots; v_{j_q}^*, d_{i_q}).$$

If $d' \in \mathbb{N}^{(1)}$ is another tuple of degree $|d'| = k$, then one sees from (4.14) that $S_d(v^{d'}) = 0$ unless $d = d'$, and $S_d(v^d) = 1$. In short,

$$S_d(v^d) = \delta_{dd'}, \quad (4.15)$$

where $\delta_{dd'}$ is a generalized Kronecker delta, $\delta_{dd'} = 1$ if $d = d'$ and $\delta_{dd'} = 0$ if $d \neq d'$. ((4.15) is the symmetric analogue of (4.4).) Thus, if $\sum_{d' \in \mathbb{N}^{(1)}} \lambda_{d'} v^{d'} = 0$ is a dependence relation among the monomials, after applying $S_d$ to both sides we find

$$0 = S_d \left( \sum_{|d'|-k} \lambda_{d'} v^{d'} \right) = \sum_{|d'|-k} \lambda_{d'} S_d(v^{d'}) = \sum_{|d'|-k} \lambda_{d} \delta_{dd'}, \quad so \lambda_d = 0, \quad and \ hence \ B^{(k)} \ is \ independent. \ See \ Exercise \ 4.18 \ for \ counting \ the \ dimension \ of \ S^k(V). \ Assertion \ (ii) \ follows \ from \ (i) \ and \ the \ fact \ that \ S(V) \ is \ the \ direct \ sum \ of \ all \ S^k(V).$$

4.21. COROLLARY. If $V$ is finite-dimensional, then so is $S^k(V)$ for all $k \geq 0$. The functionals $S_d \in (S^k(V))^*$, where $d \in \mathbb{N}^{(1)}$ ranges over all tuples of degree $|d| = k$, form a basis of $(S(V)^k)^*$ which is dual to the basis $v^d$ of $S^k(V)$.

PROOF. The first assertion follows from Theorem 4.20(i) and the second assertion follows from (4.15). QED

Like the tensor algebra, the symmetric algebra is infinite-dimensional (unless $V = \{0\}$, in which case we have $S(V) = T(V) = F$).

As a second application of the universal property, let us replace the vector space $V$ in Theorem 4.18 by its dual, and for $A$ let us take the algebra $F^V$ consisting of all $F$-valued functions on $V$. Elements of $V^*$ are (linear) functions on $V$, so $V^*$ is
a subspace of $F^V$. By Theorem 4.18 the inclusion map $V^* \to F^V$ extends to an algebra homomorphism $S(V^*) \to F^V$.

$$
\begin{array}{ccc}
V^* & \to & F^V \\
\downarrow & & \downarrow \\
S(V^*)
\end{array}
$$

Let us denote the image of an element $p \in S(V^*)$ by $\tilde{p}$. The function $\tilde{p} : V \to F$ is called a polynomial function on $V$.

To understand this a little better, let us suppose $V = F^n$, where $n$ is finite, and let us denote the dual basis element $e_i^*$ by $x_i$. By Theorem 4.20, every element of $S([F^n]^*)$ can be written uniquely as a polynomial $p(x) = \sum_{a \in \mathbb{N}^n} \lambda_a x^a$. For this reason we usually write

$$
S([F^n]^*) = F[x_1, x_2, \ldots, x_n]
$$

by analogy with the polynomial algebra $F[x]$ in one variable. The value of the function $\tilde{p}$ at $a = (a_1, a_2, \ldots, a_n)^T \in F^n$ is given by

$$
\tilde{p}(a) = \sum_{d \in \mathbb{N}^n} \lambda_d x^d(a) = \sum_{d \in \mathbb{N}^n} \lambda_d x_1(a)^{d_1} x_2(a)^{d_2} \cdots x_n(a)^{d_n} = \sum_{d \in \mathbb{N}^n} \lambda_d d_1^{a_1} d_2^{a_2} \cdots d_n^{a_n}.
$$

This is called evaluating the polynomial at the vector $a$. See Exercise 1.4 for the one-dimensional case. As in the one-dimensional case, the map $S(V^*) \to F^V$ is not injective if the field is finite, so the polynomial function $\tilde{p}$ does not determine the polynomial $p$.

### 4.6. Alternating algebras

The alternating algebra $A(V)$ of a vector space $V$ is the anticommutative twin of the symmetric algebra. It is obtained from the tensor algebra by “making the multiplication alternating”, in other words by imposing the rule $\nu \otimes \nu = 0$ for all $\nu \in V$. Formally, the alternating algebra is very similar to the symmetric algebra, so we shall develop its properties in parallel to the symmetric case. However, let us immediately point out one fundamental difference: if $V$ is $n$-dimensional, then $A(V)$ is $2^n$-dimensional. In particular, if $V$ is finite-dimensional, then so is $A(V)$, unlike $S(V)$ or $T(V)!$ Let $Y$ be the subset of $T(V)$ consisting of all degree 2 tensors of the form

$$
\nu \otimes \nu,
$$

where $\nu$ is an arbitrary vector in $V$, and let $K = \langle Y \rangle$ be the ideal generated by $Y$. The alternating algebra $A(V)$ of $V$ (also denoted by $A(V)$) is declared to be the quotient algebra $T(V)/K$. Elements of $A(V)$ are alternating tensors. The product of alternating tensors $x$ and $y$ is written as $x \wedge y$. Put $K_k = K \cap T^k(V)$. Since $Y \subseteq T^2(V)$, we have $K_0 = K_1 = \{0\}$.

4.22. **Lemma.** $K = \bigoplus_{k=2}^\infty K_k = K_2 \oplus K_3 \oplus K_4 \oplus \cdots$. 
4.6. ALTERNATING ALGEBRAS

PROOF. Parallel to the proof of Lemma 4.17. Put \( K' = \bigoplus_{k=0}^{\infty} K_k \). The inclusion \( K' \subseteq K \) is proved as in the symmetric case. Now let \( \nu \in k \). By the definition of \( K \) we can write

\[
\nu = \sum_{p=1}^{q} \lambda_p a_p \otimes [v_p \otimes v_p] \otimes b_p
\]

for certain \( \lambda_p \in F \), \( v_p \in V \), and \( a_p \), \( b_p \in T(V) \). The \( p \)-th term in this expression is of degree \( k \) if and only if \( \deg a_p = \deg b_p = k - 2 \). Thus the component of \( \nu \) of degree \( k \) is

\[
\sum_{\deg a_p + \deg b_p = k - 2} \lambda_p a_p \otimes [v_p \otimes v_p] \otimes b_p,
\]

which is an element of \( K_k \). Hence \( \nu \in K' \). QED

It follows that the alternating algebra is a direct sum,

\[
A(V) = T(V)/K = \left( \bigoplus_{k=0}^{\infty} T^k(V) \right)/\left( \bigoplus_{k=0}^{\infty} K_k \right) = \bigoplus_{k=0}^{\infty} T^k(V)/K_k = F \oplus V \oplus T^2(V)/K_2 \oplus T^3(V)/K_3 \oplus \ldots.
\]

The \( k \)-th summand \( T^k(V)/K_k \) of the symmetric algebra is called the \( k \)-th alternating power of \( V \) and denoted by \( A^k(V) \). In particular, \( A^0(V) = T^0(V) = F \) and \( A^1(V) = T^1(V) = V \), so \( A(V) \) contains a copy of the scalars \( F \) and of the vector space \( V \). Moreover, since \( T^k(V) \) is spanned by elements of the form \( v_1 \otimes v_2 \otimes \cdots \otimes v_k \) with \( v_1, v_2, \ldots, v_k \in V \), and the projection map \( \text{pr}: T^k(V) \rightarrow A^k(V) \) is surjective for each \( k \), we see that \( A^k(V) \) is spanned by elements of the form

\[
\text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \text{pr}(v_1) \wedge \text{pr}(v_2) \wedge \cdots \wedge \text{pr}(v_k) = v_1 \wedge v_2 \wedge \cdots \wedge v_k,
\]

in other words,

\[
A^k(V) = \text{span}(v_1 \wedge v_2 \wedge \cdots \wedge v_k \mid v_1, v_2, \ldots, v_k \in V). \quad (4.16)
\]

4.23. LEMMA (graded commutativity). Let \( x \in A^k(V) \) and \( y \in A^l(V) \). Then

\[
y \wedge x = (-1)^{kl} x \wedge y.
\]

PROOF. If \( v_1, v_2, \ldots, v_k \in V \), then

\[
\text{pr}(v_i \otimes v_i) = 0, \quad \text{pr}(v_i + v_j \otimes v_i + v_j) = 0
\]

for all \( i \) and \( j \), where \( \text{pr}: T(V) \rightarrow A(V) \) is the projection map. Bilinearity gives

\[
\text{pr}(v_i \otimes v_i) = -\text{pr}(v_i \otimes v_i), \quad \text{and hence}
\]

\[
\text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = [-1]^{|\text{sgn}(e)} \text{pr}(v_{e(1)} \otimes v_{e(2)} \otimes \cdots \otimes v_{e(k)}) \quad (4.17)
\]

for every permutation \( e \) of \( k \) elements. In particular,

\[
\text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) \wedge \text{pr}(w_1 \otimes w_2 \otimes \cdots \otimes w_l)
\]

\[
= \text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k \wedge w_1 \otimes w_2 \otimes \cdots \otimes w_l)
\]

\[
= (-1)^{kl} \text{pr}(w_1 \otimes w_2 \otimes \cdots \otimes w_l \otimes v_1 \otimes v_2 \otimes \cdots \otimes v_k)
\]

\[
= (-1)^{kl} \text{pr}(w_1 \otimes w_2 \otimes \cdots \otimes w_l) \wedge \text{pr}(v_1 \otimes v_2 \otimes \cdots \otimes v_k).
\]
By (4.10), the elements of the form \( v_1 \wedge v_2 \wedge \cdots \wedge v_k = \text{pr}\{v_1 \otimes v_2 \otimes \cdots \otimes v_k\} \) span \( A(V) \). Hence, by bilinearity, \( A(V) \) is commutative. 

QED

An algebra \( A \) over \( F \) is graded if there exist linear subspaces \( A_0, A_1, A_2, \ldots \) such that \( A = \bigoplus_{k=0}^\infty A_k \) and if \( xy \in A_{k+1} \) for all \( x \in A_k \) and \( y \in A_1 \). For instance, the algebras \( T(V), S(V) \) and \( A(V) \) are graded. A graded algebra is graded commutative if \( y \wedge x = (-1)^{k}x \wedge y \) for all \( x \in A_k \) and \( y \in A_1 \). For instance, \( A(V) \) is graded commutative by Lemma 4.23, but \( T(V) \) and \( S(V) \) are not. \( S(V) \) is an

4.24. Theorem (universal property of the alternating algebra). \( A(V) \) is an associative graded commutative unital algebra. For every associative unital algebra \( A \) and every linear map \( f: V \to A \) satisfying \( f(v)^2 = 0 \) for all \( v \in V \) there exists a unique unital homomorphism \( \tilde{f}: A(V) \to A \) satisfying \( \tilde{f}|V = f \).

\[ \begin{array}{c}
V \\
\downarrow f \downarrow \tilde{f} \\
\subseteq \downarrow
A(V)
\end{array} \]

Proof. Associativity and existence of a unit follow from the corresponding properties of the tensor algebra and from Theorem 4.16. Lemma 4.23 says that \( A(V) \) is graded commutativity. The universal property of the tensor product gives, for every linear map \( f: V \to A \) a unital algebra homomorphism \( \tilde{f}: T[V] \to A \) extending \( f \). Since \( f(v)^2 = 0 \) for all \( v \in V \), the kernel of \( \tilde{f} \) contains the ideal \( K \). Hence, by Exercise 1.15, \( \tilde{f} \) descends to a unique linear map \( \tilde{f}: A(V) = T(V)/K \to A \). From the fact that \( \tilde{f} \) is a unital algebra homomorphism one deduces easily that \( \tilde{f} \) is a unital algebra homomorphism as well. It follows from

\[ \tilde{f}(v_1 \wedge v_2 \wedge \cdots \wedge v_k) = \tilde{f}(v_1) \wedge \tilde{f}(v_2) \wedge \cdots \wedge \tilde{f}(v_k) = f(v_1) \wedge f(v_2) \wedge \cdots \wedge f(v_k) \]

and (4.16) that \( \tilde{f} \) is uniquely determined by \( f \). QED

Recall that a multilinear map \( f: V^k \to W \) is alternating if

\[ f(v_1, v_2, \ldots, v_k) = 0 \]

whenever \( v_i = v_j \) for some pair of distinct indices \( i \neq j \). (See Definition 2.1.) Let \( m: V^k \to A^k(V) \) be the multiplication map defined by \( m(v_1, v_2, \ldots, v_k) = v_1 \wedge v_2 \wedge \cdots \wedge v_k \). The following result can be deduced from Theorem 4.5 in the same way that we deduced Theorem 4.24 from Theorem 4.13.

4.25. Theorem (universal property of the k-th alternating power). \( m \) is alternating multilinear for all \( k \geq 0 \). For every vector space \( W \) over \( F \) and every alternating multilinear map \( f: V^k \to W \) there exists a unique linear map \( \tilde{f}: A^k(V) \to W \) such that \( \tilde{f} \circ m = f \).

\[ \begin{array}{c}
V^k \\
\downarrow f \downarrow m \downarrow \tilde{f} \\
A^k(V) \to W
\end{array} \]
4.6. ALTERNATING ALGEBRAS

PROOF. See Exercise 4.16. QED

Let us describe a basis of each of the alternating powers $\Lambda^k(V) = T^k(V)/K_k$ in terms of a basis $B = \{v_i \mid i \in I\}$ of $V$. To facilitate this, let us introduce a total ordering on the index set $I$. (For example, if $\dim V = n$ is finite, we can let $I = \{1, 2, \ldots, n\}$, or if $\dim V$ is countable, we can let $I = \{1, 2, 3, \ldots\}$ or $I = \mathbb{N} = \{0, 1, 2, \ldots\}$.) By means of (4.16) we can express any alternating tensor $v$ of degree $k$ as

$$v = \sum_{i_1 \in I} \ldots \sum_{i_k \in I} \cdot \cdot \cdot \sum_{i_k \in I} \lambda_{i_1, i_2, \ldots, i_k} v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_k}.$$  \hspace{1cm} (4.18)

By graded commutativity, if two or more of the indices $i_1, i_2, \ldots, i_k$ are the same, then $v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_k} = 0$. This has the following important consequence.

4.26. LEMMA. Suppose $\dim V = n < \infty$. Then $\Lambda^k(V) = \{0\}$ for $k > n$.

PROOF. The dimension of $V$ is the number of elements of $I$, so if $k > n = |I|$ and $i \in I^k$ is a multi-index with $k$ entries, at least two of the indices $i_1, i_2, \ldots, i_k$ are the same. Hence, by writing $v \in \Lambda^k(V)$ as in (4.18), we find $v = 0$. QED

If a multi-index $i = (i_1, i_2, \ldots, i_k)$ has distinct entries, we can rearrange them in increasing order by means of a permutation. A multi-index $i = (i_1, i_2, \ldots, i_k) \in I^k$ is increasing if $i_1 < i_2 < \cdots < i_k$. Let us denote the collection of all increasing multi-indices with $k$ entries by $I^k_+$. Let us also put $\lambda_i = \lambda_{i_1, i_2, \ldots, i_k}$ and $v^\wedge = v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_k}$. Then by using (4.17) we can rewrite the expansion (4.18) as a sum over all multi-indices $i \in I^k_+$,

$$v = \sum_{i \in I^k_+} \lambda_i v^\wedge$$ \hspace{1cm} (4.19)

with finitely many nonzero coefficients $\lambda_i \in F$. Likewise, an arbitrary symmetric tensor $v$ in $A(V)$ is of the form

$$v = \sum_{k=0}^{\infty} \sum_{i \in I^k_+} \lambda_i v^\wedge$$ \hspace{1cm} (4.20)

with finitely many nonzero $\lambda_i \in F$. The following results says that the coefficients $\lambda_i$ in these expressions are uniquely determined.

4.27. THEOREM. Let $B = \{v_i \mid i \in I\}$ be a basis of $V$, where $I$ is a totally ordered index set. Let $I^k_+$ denote the collection of all increasing multi-indices in $I^k$.

(i) For all $k \in \mathbb{N}$ the collection

$$\mathcal{B}^k_+ = \{v^\wedge \mid i \in I^k_+\}$$

is a basis of $\Lambda^k(V)$. Hence $\dim \Lambda^k(V) = \binom{n}{k}$, where $n = |I| = \dim V$.

(Here we put $\mathcal{B}^0_+ = \{1\}$ and $\mathcal{B}^1_+ = B$. We put $\binom{n}{k} = n$ if $n$ is infinite.)

(ii) The collection

$$\bigcup_{k=0}^{\infty} \mathcal{B}^k_+ = \{v^\wedge \mid i \in I^k_+\}$$

...
is a basis of $A(V)$. Hence $\dim A(V) = 2^n$ if $n$ is finite and $\dim A(V) = \infty$ if $n$ is infinite.

**Proof.** It follows from (4.19) that $B^k_+$ spans $A^k(V)$. It remains to show that $B^k_+$ is independent. Let $l_1, l_2, \ldots, l_k \in V^*$ be covectors on $V$. Then the function $m: V^k \to F$ defined by

$$m(v_1, v_2, \ldots, v_k) = \det(l_p(v_q))_{1 \leq p, q \leq k} = \sum_{\sigma \in S_k} \text{sign}(\sigma) l_1(v_{\sigma(1)}) l_2(v_{\sigma(2)}) \cdots l_k(v_{\sigma(k)}).$$

is alternating multilinear, because $\det: M_k(F) \to F$ is alternating multilinear. By the universal property of alternating powers, Theorem 4.25, $m$ descends to a linear map $\hat{m}: A^k(V) \to F$. The map $\hat{m}$ depends on the covectors $l_p \in V^*$ and therefore we will denote it by $A(l_1, l_2, \ldots, l_k)$.

Now let $i \in I^k_+$ be an increasing multi-index with $k$ entries. Define a linear map $A_i: A^k(V) \to F$, i.e. a covector on $A^k(V)$, by

$$A_i = A(v^*_{i_1}, v^*_{i_2}, \ldots, v^*_{i_k}).$$

If $j \in I^k_+$ is another increasing multi-index, then

$$A_i(v^*_j) = \det(v^*_{i_p}(v_{j_q})) = \det(\delta_{i_p j_q})_{1 \leq p, q \leq k}. $$

Let $B$ be the $k \times k$-matrix $(\delta_{i_p j_q})$. If $i = j$ then $B = I$, so $\det B = 1$. Suppose $i \neq j$. Then $i_q \neq j_q$ for some $q = 1, 2, \ldots, k$. Choose $q$ as small as possible, so that $i_p = j_p$ for $p < q$. There are two cases: $i_q < j_q$ and $i_q > j_q$. If $i_q < j_q$, then $i_p < j_p < j_{p+1} < \cdots < j_k$, because $j$ is increasing, so all entries $\delta_{i_p j_q}$ in $B$ with $p < q$ are 0. For $p < q$ we have $j_p = i_p < i_q$ because $l$ is increasing, so $\delta_{i_p j_q} = 0$ for $p < q$. In other words, the $q$-th row of $B$ is 0 and hence $\det B = 0$. If $i_q > j_q$ we find that the $q$-th column of $B$ is 0 and therefore again $\det B = 0$. The upshot is

$$A_i(v^*_j) = \delta_{ij}, \quad (4.21)$$

which is the alternating analogue of (4.4). Therefore, if $\sum_{i \in I^k_+} \lambda_i v^*_i = 0$ is a dependence relation among the $v^*_i$, after applying $A_i$ to both sides we find

$$0 = A_i \left( \sum_{j \in I^k_+} \lambda_j v^*_j \right) = \sum_{j \in I^k_+} \lambda_j A_i(v^*_j) = \sum_{j \in I^k_+} \lambda_j \delta_{ij} = \lambda_i,$$

so $\lambda_i = 0$, and hence $B^k_+$ is independent. Hence the dimension of $A^k(V)$ is equal to the cardinality of $I^k_+$, which is equal to the number of $k$-element subsets of $I$. Thus $\dim A^k(V) = \binom{n}{k}$, where $n = |I|$. Assertion (ii) follows from (i) and the fact that $A(V)$ is the direct sum of all $A^k(V)$.

QED
4.28. COROLLARY. If $V$ is finite-dimensional, then so is $A^k(V)$ for all $k \geq 0$. The functionals $\lambda_i \in (A^k(V))^*$, where $i \in I^k$, ranges over all increasing multi-indices with $k$ entries, form a basis of $(A^k(V))^*$ which is dual to the basis $v_i^\wedge$ of $A^k(V)$.

PROOF. The first assertion follows from Theorem 4.27(i) and the second assertion follows from (4.21). QED

You may be wondering where the alternating algebra comes from. Let us suppose $V = \mathbb{F}^n$, where $n$ is finite, and let us denote the dual basis element $e^*_i$ by $dx_i$. This notation is meant to suggest an "infinitesimal increment" along the $i$-th coordinate axis. A wedge product $dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}$ is to be thought of as an infinitesimal $k$-dimensional oriented volume element with edges $dx_{i_1}, dx_{i_2}, \ldots, dx_{i_k}$. Swapping two of the edges has the effect of reversing the orientation, which explains the alternating rule. Formally adding and multiplying volume elements then leads naturally to the alternating algebra. By Theorem 4.27, every element of $A^k((\mathbb{F}^n)^*)$ can be written uniquely in the form $\sum_{i \in I^k} \lambda_i dx_i^\wedge$. Elements of the alternating algebra $A((\mathbb{F}^n)^*)$ are often called differential forms (with constant coefficients) on $\mathbb{F}^n$ and they play an important role in differential and algebraic geometry. For instance, to compute the volume or the centroid of a $k$-dimensional manifold in $\mathbb{F}^n$, one has to integrate a differential form of degree $k$ over the manifold. See Exercise 4.23 for a geometric application of alternating tensors.

Exercises

4.1. Let $V$ be a complex vector space. Define a new scalar multiplication law on $V$ by \[ \lambda \cdot v = \overline{\lambda} v \] for $v \in V$ and $\lambda \in \mathbb{C}$. Here $\overline{\lambda}$ denotes the complex conjugate of $\lambda$.

(i) $V$ equipped with the old addition law and the new scalar multiplication law is a complex vector space (called the vector space conjugate to $V$ and denoted by $\bar{V}$).

(ii) The identity map $g : V \rightarrow \bar{V}$ is conjugate linear or semilinear, i.e. $g(v_1 + v_2) = g(v_1) + g(v_2)$ and $g(\lambda v) = \overline{\lambda} g(v)$ for all $v_1, v_2, v \in V$ and $\lambda \in \mathbb{C}$.

(iii) Let $W$ be a second complex vector space and let $f : V \rightarrow W$ be linear. Let $g_V : V \rightarrow \bar{V}$ and $g_W : W \rightarrow \bar{W}$ be the respective identity maps. Let $\bar{f} : \bar{V} \rightarrow \bar{W}$ be the map defined by $\bar{f} = g_W \circ f \circ g_V^{-1}$. Then $\bar{f}$ is linear.

(iv) A basis of $V$ is also a basis of $\bar{V}$.

(v) Let $A$ be the matrix of $f$ relative to ordered bases $\mathcal{B}$ and $\mathcal{C}$ of $V$, resp. $W$. What is the matrix of $\bar{f}$ relative to $\mathcal{B}$ and $\mathcal{C}$?

(Note that (iv) implies that $\bar{V}$ is isomorphic to $V$ as a complex vector space. But the isomorphism is not $g$!)

For two vector spaces $V$ and $W$ over the same field $F$, $\text{Bil}(V,W)$ denotes the collection of all bilinear maps $F : V \times W \rightarrow F$. (See Section 1.9.) For $F = \mathbb{C}$, $\text{Ses}(V,W)$ denotes the collection of all sesquilinear ("$\frac{1}{2}$-linear") maps $F : V \times W \rightarrow F$. Here we will follow the convention of the book and call $F$ sesquilinear if it is semilinear in the first and linear in the second variable. We also write $\text{Bil}(V) = \text{Bil}(V,V)$ and $\text{Ses}(V) = \text{Ses}(V,V)$.

4.2. Let $V$ and $W$ be complex vector spaces.

(i) $\text{Ses}(V,W)$ is a linear subspace of $\mathbb{C}^{V \times W}$ (and hence a complex vector space in its own right).
(ii) Let $F: V \times W \to \mathbb{C}$ be a sesquilinear form. Define a map $\tilde{F}: \overline{V} \times W \to \mathbb{C}$ by $\tilde{F}(v, w) = F(g(v), w)$, where $\overline{V}$ is the vector space conjugate to $V$ and $g: V \to \overline{V}$ is the identity map as in Exercise 4.1. Then $\tilde{F}$ is bilinear.

(iii) Define $\mathcal{J}: \text{Ses}(V, W) \to \text{Bil}(\overline{V}, W)$ by $\mathcal{J}(F) = \tilde{F}$. Then $\mathcal{J}$ is an isomorphism of vector spaces.

(iv) $\dim \text{Ses}(V, W) = (\dim V)(\dim W)$.

The upshot of this exercise is that we may regard sesquilinear forms either as bilinear forms $\overline{V} \times W \to \mathbb{C}$ or alternatively as linear maps $\overline{V} \to W^*$.

4.3 (base-change for sesquilinear forms). Let $f: \overline{V} \to V^*$ be linear and let $F_1: V \times V \to F$ be the associated sesquilinear form on $V$. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be an ordered basis of $V$. Let $\mathcal{B}^* = \{v_1^*, v_2^*, \ldots, v_n^*\}$ be the dual basis of $V^*$. Let $A$ be the matrix of $f$ relative to $\mathcal{B}$ and $\mathcal{B}^*$ and let $M$ be the matrix of $F_1$ relative to $\mathcal{B}$ (which is defined by $m_{ij} = F_1(v_i, v_j)$).

(i) $A = M^T$.

(ii) If $C = \{w_1, w_2, \ldots, w_n\}$ is a new ordered basis of $V$ and $N$ is the matrix of $F_1$ relative to the new basis, then $N = \overline{P}^T M P$, where $P$ is the base-change matrix from $\mathcal{B}$ to $\mathcal{C}$.

In the following exercises, with the sole exception of Exercise 4.9, $U$, $V$, $W$, $Z$, $E$ and $F$ denote vector spaces over the same field $F$.

4.4. If $m: U \times V \to W$ is bilinear and $f: W \to Z$ is linear, then $f \circ m: U \times V \to Z$ is bilinear.

4.5. $\lambda(u \otimes v) = (\lambda u) \otimes v = u \otimes (\lambda v)$ for all $\lambda \in F$, $u \in U$ and $v \in V$. (For this reason we usually simply write $\lambda u \otimes v$ for any of these three expressions.)

4.6. $U \otimes V$ is isomorphic to $V \otimes U$. In fact, there is a unique isomorphism $f: U \otimes V \to V \otimes U$ satisfying $f(u \otimes v) = v \otimes u$ for all $u \in U$ and $v \in V$. (You can do this directly from the definition of tensor products, but it is easier to use the universal property. Define a map $F: U \times V \to V \otimes U$ by $F(u, v) = (v, u)$. Check that $\otimes f: U \times V \to V \otimes U$ is bilinear and use this to get the map $f: U \otimes V \to V \otimes U$.)

4.7. $F \otimes V$ is isomorphic to $V$. In fact, there is a unique isomorphism $f: F \otimes V \to V$ satisfying $f(\lambda \otimes v) = \lambda v$ for all $\lambda \in F$ and $v \in V$. (Similar comment as for Exercise 4.6.)

4.8. Prove Corollary 4.8. (Use Theorem 4.6.) Explain why this is called the associativity property of the tensor product.

4.9 (extension of scalars done right). Let $F'$ be a subfield of $F$ and let $V$ be a vector space over $F'$. Regard $F$ as an $F'$-vector space and define $V_F = F \otimes_{F'} V$.

(i) There exists a unique $F'$-linear map $i: V \to V_F$ satisfying $i(v) = 1_F \otimes v$ for all $v \in V$. The map $i$ is injective.

(ii) There exists a unique bilinear map $s: F \times V_F \to V_F$ satisfying $s(\lambda, v) = \lambda v$ for all $\lambda \in F$ and $v \in V_F$.

(iii) $V_F$ equipped with the addition law coming from its $F'$-vector space structure and the scalar multiplication law defined in (ii), is a vector space over $F$.

(iv) For all $F$-vector spaces $Z$ and all $F'$-linear maps $g: V \to Z$, there exists a unique $F$-linear map $\tilde{g}: V_F \to Z$ such that $\tilde{g} \circ i = g$.
4.10. $\text{Bil}(U, V)$ is isomorphic to $(U \otimes V)^*$. (This does not require finite-dimensionality of $U$ or $V$. Use the universal property of tensor products to associate to any bilinear map in $\text{Bil}(U, V)$ an element of $(U \otimes V)^*$.)

4.11. (i) The map

$$T: U^* \times V^* \to (U \otimes V)^*$$

which sends a pair of covectors $(k, l)$ to $T[k, l]$ is bilinear.

(ii) There is a unique linear map $t: U^* \otimes V^* \to (U \otimes V)^*$ which satisfies $t(k \otimes l) = T[k, l]$ for all $k \in U^*$ and $l \in V^*$.

(iii) Select bases $\mathcal{A} = \{u_i\}_{i \in I}$ of $U$ and $\mathcal{B} = \{v_j\}_{j \in J}$ of $V$. Define the coordinate functionals $u^*_i \in U^*$ by $u^*_i(u_{i'}) = \delta_{i'i}$ and $v^*_j \in V^*$ by $v^*_j(v_{j'}) = \delta_{jj'}$. Then

$$t(u^*_i \otimes v^*_j) = \delta_{i'j}.$$  

(iv) If $U$ and $V$ are finite-dimensional, then $t: U^* \otimes V^* \to (U \otimes V)^*$ is an isomorphism. (Use part (iii), the dual basis theorem, Proposition 1.11, and Theorem 4.4 to conclude that the elements $t(u^*_i \otimes v^*_j)$ form a basis of $(U \otimes V)^*$.)

4.12. Let $f: U \to E$ and $g: V \to F$ be linear maps.

(i) There is a unique linear map $h: U \otimes E \to F \otimes V$ satisfying $h(\mathbf{u} \otimes \mathbf{v}) = f(\mathbf{u}) \otimes g(\mathbf{v})$ for all $\mathbf{u} \in U$ and $\mathbf{v} \in V$. We denote $h$ by $f \otimes g$.

(ii) Suppose $U, E, V, F$ are of finite dimension $n, m, q, p$, respectively. Choose respective ordered bases $\mathcal{A} = \{u_1, u_2, \ldots, u_n\}$, $\mathcal{E} = \{e_1, e_2, \ldots, e_m\}$, $\mathcal{B} = \{v_1, v_2, \ldots, v_q\}$, $\mathcal{D} = \{f_1, f_2, \ldots, f_p\}$ of these four spaces. Let $A \in M_{m,n}(F)$ be the matrix of $f$ with respect to $\mathcal{A}$ and $\mathcal{E}$ and let $B \in M_{p,q}(F)$ be the matrix of $g$ with respect to $\mathcal{B}$ and $\mathcal{D}$. Then the matrix of $f \otimes g$ with respect to the ordered bases

$$\{u_1 \otimes v_1, u_1 \otimes v_2, \ldots, u_1 \otimes v_q, u_2 \otimes v_1, u_2 \otimes v_2, \ldots, u_2 \otimes v_q, \ldots, u_n \otimes v_1, u_n \otimes v_2, \ldots, u_n \otimes v_q\}$$

of $U \otimes V$ and

$$\{e_1 \otimes f_1, e_1 \otimes f_2, \ldots, e_1 \otimes f_p, e_2 \otimes f_1, e_2 \otimes f_2, \ldots, e_2 \otimes f_p, \ldots, e_m \otimes f_1, e_m \otimes f_2, \ldots, e_m \otimes f_p\}$$

of $E \otimes F$ is the $mp \times nq$-matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1n}B \\ a_{21}B & a_{22}B & \ldots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \ldots & a_{mn}B \end{pmatrix},$$

known as the Kronecker or tensor product of $A$ and $B$.

(iii) Find invertible matrices $P \in M_{m,n}(F)$ and $Q \in M_{mp,q}(F)$ such that $B \otimes A = Q^{-1}(A \otimes B)P$.

4.13. (i) Define a cross product $\times$ on $\mathbb{R}^3$ analogous to the one on $\mathbb{R}^3$.

(ii) $F^3$ equipped with the cross product is an algebra.
4. MULTILINEAR ALGEBRA

(iii) This algebra is neither associative nor unital.
(iv) This algebra is anti-commutative: \( x \times y = -y \times x \) for all \( x \) and \( y \) in \( F^3 \).
(v) This algebra satisfies the Jacobi identity: \( (x \times y) \times z + (z \times x) \times y + (y \times z) \times x = 0 \)
for all \( x, y, \) and \( z \) in \( F^3 \).

An anti-commutative algebra satisfying the Jacobi identity is called a Lie algebra after the Norwegian mathematician Sophus Lie (pronounced as "lee", not "lye").

4.14. Let \( X \) be any nonempty subset of an algebra \( A \) over \( F \).
(i) The set \( \{ X \} \) defined in Example 4.15 is an ideal of \( A \).
(ii) If \( J \) is an ideal of \( A \) containing \( X \), then \( X \) contains \( \{ X \} \).

4.15. Check that the multiplication law \( m \) of Theorem 4.16 is bilinear.


4.17. Check that the map \( m: V^k \to F \) defined in (4.14) is symmetric multilinear.

4.18. (i) Let \( n \) and \( k \) be natural numbers. There exist \( \binom{n+k-1}{n-1} \) monomials
of degree \( k \) in the \( n \) variables \( x_1, x_2, \ldots, x_n \). (Let \( x^k = x_1^{i_1}x_2^{i_2} \cdots x_n^{i_n} \) be a
monomial of degree \( k \). Counting the number of these is complicated by the fact
that some of the exponents \( i_1 \) may be 0. Therefore multiply \( x^k \) by \( x_1 \) \( x_2 \) \( \cdots \) \( x_n \)
to get the monomial \( x_1^{i_1+1}x_2^{i_2+1} \cdots x_n^{i_n+1} \) of degree \( n+k \) with all exponents
positive. To count these, note that specifying such a monomial is equivalent
to partitioning a row of \( n+k \) things into \( n \) subrows.)
(ii) Let \( \{ x_i | i \in I \} \) be a collection of variables of infinite cardinality \( n = |I| \). Then
there exist \( n \) monomials of degree \( k \) in the variables \( x_i \).

4.19. If \( f: V \to V \) is linear and \( \{ v_1, v_2, \ldots, v_n \} \) is a basis of \( V \), then
\[ f(v_1) \wedge f(v_2) \wedge \cdots \wedge f(v_n) = \det f \wedge v_2 \wedge \cdots \wedge v_n. \]

4.20. Let \( f: V \to W \) be a linear map.
(i) There is a unique unital algebra homomorphism \( g: \Lambda(V) \to \Lambda(W) \) satisfying
\( g(V) = f \). (We denote \( g \) by \( A(f) \).)
(ii) \( A(f) \) maps \( \Lambda^k(V) \) to \( \Lambda^k(W) \) for all \( k \geq 0 \). (We denote the restriction of \( A(f) \)
to \( \Lambda^k(V) \) by \( \Lambda^k(f) \).)

4.21. Let \( f: V \to W \) be a linear map. Suppose \( V \) and \( W \) are finite-dimensional
and select bases \( \{ v_1, v_2, \ldots, v_n \} \) of \( V \) and \( \{ w_1, w_2, \ldots, w_m \} \) of \( W \). Let \( M = \{ m_{ij} \} \) be
the matrix of \( f \) relative to these bases. Find the matrix \( \Lambda^k(M) \) of \( \Lambda^k(f) \) relative to the
bases \( \{ w_1 \wedge \cdots \wedge w_i \} \) of \( \Lambda^k(V) \) and \( \{ w_1 \wedge \cdots \wedge w_j \} \) of \( \Lambda^k(W) \).
Here \( I = \{ 1, 2, \ldots, n \} \) and \( J = \{ 1, 2, \ldots, m \} \). (To facilitate this calculation, recall from Proposition 1.15 that
\( m_{ij} = w_1 \wedge \cdots \wedge w_i \) of \( \Lambda^k(V) \). Likewise, if \( i \in I \) and \( j \in J \) are increasing multi-indices, the \( ij \)-th
entry of \( \Lambda^k(M) \) is \( \Lambda^k(I \wedge J) \). Now use Corollary 4.28.)

The following two problems are taken from A comprehensive introduction to differential
gometry by M. Spivak.

4.22. An alternating tensor \( v \in \Lambda(V) \) is decomposable if \( v = v_1 \wedge v_2 \wedge \cdots \wedge v_k \) for
some \( v_1, v_2, \ldots, v_k \in V \), and it is indecomposable if it is not decomposable.
(i) If \( \dim V \leq 3 \), then every element of \( \Lambda^2(V) \) is decomposable.
(ii) If \( v \) is decomposable, then \( v^2 = 0 \). (Here \( v^2 \) denotes the square of \( v \) in the
alternating algebra \( \Lambda(V) \), i.e. \( v \wedge v \).)
(iii) Let \( v_1, v_2, \ldots, v_n \in V \) and put \( v = v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge \cdots \wedge v_{2n-1} \wedge v_{2n} \).
Compute \( v^3 = v \wedge v \wedge v \wedge v \) (\( n \) times).
(iv) If $v_1, v_2, \ldots, v_{2n}$ are independent, then the alternating tensor $v$ of degree $2n$
defined in (iii) is indecomposable.

4.23. The annihilator of an alternating tensor $v \in \Lambda^k(V)$ is $v^\circ = \{ w \in V \mid w \wedge v = 0 \}$.

(i) $v^\circ$ is a subspace of $V$ of dimension $\leq k$.

(ii) $\dim v^\circ = k$ if and only if $v$ is decomposable.

(iii) Every finite-dimensional subspace $W$ of $V$ is of the form $W = v^\circ$ for some
decomposable $v \in \Lambda^k(V)$, which is unique up to a multiplicative constant

(iv) If $v_1$ and $v_2$ are decomposable, then $v_1^\circ \subseteq v_2^\circ$ if and only if $v_2 = v_1 \wedge v_3$ for
some $v_3 \in \Lambda(V)$.

(v) If $v_1$ and $v_2$ are decomposable, then $v_1^\circ \cap v_2^\circ = \{0\}$ if and only if $v_1 \wedge v_2 \neq 0$. If
this is the case, $v_1^\circ + v_2^\circ = (v_1 \wedge v_2)^\circ$.

(vi) If $\dim V = n < \infty$, then every $v \in \Lambda^{n-1}(V)$ is decomposable.
APPENDIX A

Bases and dimension

A.1. Some set theory

This section is a selection of facts from set theory presented without proofs. See e.g. Appendix 2 of Algebra by S. Lang for proofs.

The cardinality of a set $X$ is the collection $|X|$ consisting of all sets $Y$ such that there is a bijection $f: X \to Y$. This notion generalizes the notion of "number of elements" to arbitrary, not necessarily finite, sets. Thus two sets $X$ and $Y$ have the same cardinality, $|X| = |Y|$, if and only if there is a bijection $f: X \to Y$. We say that the cardinality of $X$ is less than or equal to the cardinality of $Y$, notation $|X| \leq |Y|$, if there is an injection $f: X \to Y$. The cardinality of $X$ is less than the cardinality of $Y$, notation $|X| < |Y|$, if $|X| \leq |Y|$ and $|X| \neq |Y|$.

A.1. Theorem (Cantor).

(i) $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| < |\mathbb{R}| = |\mathbb{C}|$.

(ii) $|X| < |(0, 1)^X|$ for all sets $X$.

If $X$ and $Y$ are finite, $|X| \leq |Y|$ just means that the number of elements of $X$ is less than or equal to that of $Y$, so for finite sets the following result is not surprising.

A.2. Theorem (Schröder-Bernstein). If $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$.

A partial order on a set $X$ is a binary relation $\leq$ which is reflexive: $x \leq x$ for all $x \in X$; transitive: $x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in X$; and antisymmetric: $x \leq y$ and $y \leq x$ implies $x = y$ for all $x$ and $y \in X$. A set $X$ equipped with a partial order $\leq$ is a partially ordered set or a poset. The partial order is a total order if for all $x$ and $y$ in $Y$ we have either $x \leq y$ or $y \leq x$.

A.3. Example. Let $S$ be any set and let $X$ be the power set of $S$, i.e. the set of all subsets of $S$. For $A, B \in X$ define $A \leq B$ if $A \subseteq B$. This is a partial order on $X$, but not a total order (unless $|X| \leq 1$).

An element $x$ of a poset $X$ is maximal if $x \leq y$ implies $x = y$ for all $y \in X$. An upper bound of a subset $Y$ of a poset $X$ is an element $x$ such that $y \leq x$ for all $y \in Y$. A chain in a poset $X$ is a subset that is totally ordered (with respect to the partial order given on $X$). The following statement is a "transfinite" generalization of the principle of induction. It is equivalent to a famous axiom of set theory known as the axiom of choice.

A.4. Lemma (Zorn). Let $X$ be a nonempty poset. Suppose every chain in $X$ has an upper bound. Then $X$ has a maximal element.
A.2. The basis theorem

Let \( V \) be a vector space over a field \( F \). A spanning subset of \( V \) is a subset \( S \) satisfying \( \text{span}(S) = V \). A subset \( J \) of \( V \) is independent if, for all \( v \in J \), \( v \) is not in the span of \( J \setminus \{v\} \). A basis of \( V \) is an independent spanning subset of \( V \).

A.5. Theorem (basis theorem). (i) Let \( J \) be an independent subset and \( S \) a spanning subset of \( V \). Then there exists a basis \( B \) of \( V \) such that \( J \subseteq B \subseteq J \cup S \).
(ii) \( V \) has a basis.

Proof. The set \( J \cup S \) spans \( V \), so after replacing \( S \) with \( J \cup S \) we may assume that \( J \subseteq S \). So now we must find a basis \( B \) of \( V \) satisfying \( J \subseteq B \subseteq S \). Let \( X \) be the collection of all independent subsets \( \mathcal{C} \) of \( V \) satisfying \( J \subseteq \mathcal{C} \subseteq S \). Then \( X \) is nonempty because \( J \in X \). Define a partial order on \( X \) by \( \mathcal{C} \leq \mathcal{D} \) if \( \mathcal{C} \subseteq \mathcal{D} \). Let \( Y \) be a chain in \( X \). Let \( \mathcal{E} \) be the subset of \( V \) defined by

\[
\mathcal{E} = \bigcup_{\mathcal{C} \in Y} \mathcal{C} = \{v \in V \mid \text{there exists } \mathcal{C} \in Y \text{ such that } v \in \mathcal{C}\}.
\]

Since \( Y \) is contained in \( X \), we have \( J \subseteq \mathcal{C} \subseteq S \) for all \( \mathcal{C} \in Y \), and therefore \( J \subseteq \mathcal{E} \subseteq S \).

If \( v_1, v_2, \ldots, v_n \) are vectors in \( \mathcal{E} \), then \( v_i \in \mathcal{C}_i \) for some \( \mathcal{C}_i \in Y \). Since \( Y \) is totally ordered, for all \( i \), \( 1 \leq i \leq n \), we have either \( \mathcal{C}_i \subseteq \mathcal{C}_j \) or \( \mathcal{C}_j \subseteq \mathcal{C}_i \). After renumbering the \( \mathcal{C}_i \) if necessary we may assume \( \mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \cdots \subseteq \mathcal{C}_n \). Hence all vectors \( v_1, v_2, \ldots, v_n \) are in \( \mathcal{C}_n \) and so they are independent. We conclude that every finite subset of \( \mathcal{E} \) is independent, and therefore \( \mathcal{E} \) is independent. In other words, \( \mathcal{E} \) is a basis of \( V \). We assert that \( \mathcal{B} \) spans \( V \). Suppose it did not. Then, since \( S \) spans \( V \), there would exist \( v \in S \) such that \( v \) was not in the span of \( \mathcal{B} \). Let \( \mathcal{C} = \mathcal{B} \cup \{v\} \). Then \( \mathcal{C} \) is independent and \( J \subseteq \mathcal{C} \subseteq S \), so \( \mathcal{C} \in \mathcal{Y} \). But \( \mathcal{B} \subseteq \mathcal{C} \) and \( \mathcal{B} \neq \mathcal{C} \), which contradicts the maximality of \( \mathcal{B} \). Thus \( \mathcal{B} \) is a basis of \( V \) satisfying \( J \subseteq \mathcal{B} \subseteq S \).

Every vector space \( V \) has an independent subset, namely \( J = \emptyset \), and a spanning subset, namely \( S = V \). Hence, by (i), \( V \) has a basis. QED

A.3. The dimension theorem

It is a fact that the cardinality of any independent set \( J \) is less than or equal to that of any spanning set \( S \). Even better, from \( S \) we can obtain a new spanning set \( \{S \setminus \{v\} \cup J \} \) by omitting a suitable subset \( J \) of \( S \) which is in bijective correspondence with \( J \) and then adding on \( J \). This is proved by means of Zorn’s lemma and the following three elementary results.

A.6. Lemma (exchange lemma). Let \( A \) be a subset of \( V \). Let \( v \in A \) and \( w \in V \). Suppose \( w \in \text{span}(A) \) but \( w \notin \text{span}(A \setminus \{v\}) \).

(i) \( \text{span}(A) = \text{span}((A \setminus \{v\}) \cup \{w\}) \).

(ii) If \( A \) is independent, then so is \( (A \setminus \{v\}) \cup \{w\} \).

Proof. See Exercise A.2. QED
A vector space is finite-dimensional if it has no infinite independent subsets. By the basis theorem, this is equivalent to every basis of the space consisting of finitely many elements.

A.7. Lemma. A vector space that is spanned by a finite subset is finite-dimensional.

Proof. See Exercise A.3.

QED

A.8. Lemma. Let \( U \) be a vector space and let \( X \) be a collection of linear subspaces of \( U \). Suppose that \( U \) is finite-dimensional and that \( X \) is totally ordered (i.e. for all \( E \) and \( F \) in \( X \) we have either \( E \subseteq F \) or \( F \subseteq E \)). Then \( X \) is finite.

Proof. See Exercise A.4.

QED


(i) Let \( J \) be an independent subset and \( \delta \) a spanning subset of \( V \). Then there exists an injective map \( f: J \to \delta \) such that \( (\delta \setminus \{f(J)\}) \cup J \) spans \( V \). In particular \(|\delta| \leq |\delta|\).

(ii) Any two bases of \( V \) have the same cardinality.

Proof. Let \( X \) be the collection of all pairs \((\mathcal{K}, g)\) consisting of a subset \( \mathcal{K} \) of \( J \) and an injective map \( g: \mathcal{K} \to \delta \) such that \((\delta \setminus g(\mathcal{K})) \cup \mathcal{K} \) spans \( V \). Note that \((\emptyset, \emptyset)\) is in \( X \), so \( X \) is nonempty. Define a partial order on \( X \) by \((\mathcal{K}, g) \leq (\mathcal{L}, h)\) if \( \mathcal{K} \subseteq \mathcal{L} \) and \( g = h|\mathcal{K} \). Let \( Y \) be a chain in \( X \). Let \( \mathcal{M} \) be the subset of \( V \) defined by

\[
\mathcal{M} = \bigcup_{(\mathcal{K}, g) \in Y} \mathcal{K} = \{ v \in V \mid v \in \mathcal{K} \text{ for some } (\mathcal{K}, g) \in Y \}.
\]

Then \( \mathcal{M} \) is a subset of \( J \) and so is independent. Let \( v \in \mathcal{M} \); choose \((\mathcal{K}, g)\) \in \( Y \) such that \( v \in \mathcal{K} \), and put \( k(v) = g(v) \). If \((\mathcal{L}, h) \in Y \) also satisfies \( v \in \mathcal{L} \), then either \((\mathcal{K}, g) \leq (\mathcal{L}, h)\) or \((\mathcal{L}, h) \leq (\mathcal{K}, g)\), since \( Y \) is totally ordered. In either case we have \( g(v) = h(v) \), so \( k \) is a well-defined map \( \mathcal{M} \to \delta \). We assert that \((\mathcal{M}, k)\) is an element of \( X \). We have yet to show that \( k \) is injective and that \((\delta \setminus k(\mathcal{M})) \cup \mathcal{M} \) spans \( V \).

If \( v, w \in \mathcal{M} \) satisfy \( k(v) = k(w) \), choose \((\mathcal{K}, g)\), \((\mathcal{L}, h) \in Y \) such that \( v \in \mathcal{K} \) and \( w \in \mathcal{L} \). Again because \( Y \) is a chain, we have either \((\mathcal{K}, g) \leq (\mathcal{L}, h)\) or \((\mathcal{L}, h) \leq (\mathcal{K}, g)\), so we may just as well assume \((\mathcal{K}, g) \leq (\mathcal{L}, h)\). Then \( g = h|\mathcal{K} \), so \( k(v) = g(v) = h(v) \), and \( k(w) = h(w) \). Hence \( h(v) = h(w) \), so \( v = w \), because \( h \) is injective. We conclude that \( k \) is injective.

Let \( u \in V \). Since \( \delta \) spans \( V \), we can write \( u \) as a linear combination of finitely many elements of \( \delta \). Let \( U \) be the span of these elements. By Lemma A.7, \( U \) is a finite-dimensional subspace of \( V \). For each \( x = (\mathcal{K}, g) \in X \) let us write

\[
V_x = U \cap \text{span}(\delta \setminus g(\mathcal{K})) \quad \text{and} \quad W_x = U \cap \text{span} \mathcal{K}.
\]

By definition of the poset \( X \) we have \( U = V_x + W_x \) for all \( x \in X \), and \( V_x \supseteq V_y \) and \( W_x \subseteq W_y \) if \( x \leq y \). Since \( Y \) is totally ordered, the two collections \( \{ V_y \mid y \in Y \} \) and \( \{ W_y \mid y \in Y \} \) of subspaces of \( U \) are totally ordered. Therefore they are finite by Lemma A.8. Hence there exists \( z \in Y \) such that \( V_y \supseteq V_z \) and \( W_y \subseteq W_z \) for all
$y \in Y$. From the definition of $M$ it follows that
\[
\text{span}(S \setminus k(M)) = \bigcap_{(L,h) \in Y} \text{span}(S \setminus h(L)) \quad \text{and} \quad \text{span} M = \bigcup_{(L,h) \in Y} \text{span} L,
\]
and therefore $V_z = U \cap \text{span}(S \setminus k(M))$ and $W_z = U \cap \text{span} M$. Thus
\[
u \in V_z + W_z = \text{span}((S \setminus k(M)) \cup M),
\]
which proves that $(M,k) \in X$. By construction, $(M,k)$ is an upper bound of $Y$.

We can now apply Zorn’s lemma to obtain a maximal element $(K,g)$ of $X$. To
finish the proof it suffices to show that $K = J$. Suppose there existed $u \in J$ which
was not in $K$. Put $L = K \cup \{u\}$. Since $(K,g)$ is in $X$, $V$ is spanned by $(S \setminus g(K)) \cup K$, so we can write
\[
u = \sum_{i=1}^{p} \lambda_i v_i + w \quad \text{(A.1)}
\]
with $v_i \in S \setminus g(K)$ and $w \in \text{span} K$. Then $p \geq 1$, or else $u$ would be in the
span of $K$, which is impossible since $J$ is independent. Among all expressions for
$u$ of the form (A.1) let us choose one which has $p$ as small as possible. Then
$u \in \text{span } K \cup \{v_1,v_2,\ldots,v_p\}$, but $u \notin \text{span } K \cup \{v_1,v_2,\ldots,v_{p-1}\}$. Hence
\[
\text{span } K \cup \{v_1,v_2,\ldots,v_p\} = \text{span } K \cup \{v_1,v_2,\ldots,v_{p-1},u\} \quad \text{(A.2)}
\]
by the exchange lemma. Now define $h: L \to S$ by $h[v] = g[v]$ if $v \in K$ and
$h(u) = v_p$. Then $h$ is injective since $g$ is injective and $v_p$ is not in $g(K)$. By (A.2),
$V$ is the span of
\[
(S \setminus (g(K) \cup \{v_p\})) \cup K \cup \{u\} = (S \setminus h(L)) \cup L.
\]
Thus $(L,h)$ is an element of $X$ greater than $(K,g)$, which contradicts the maximality
of $(K,g)$. The conclusion is that $K = J$.

Since $g$ is injective, we have $|J| \leq |(S \setminus g(K)) \cup \{u\}| = |S|$. Hence $|J| \leq |S|.$

Let $B_1$ and $B_2$ be two bases of $V$. Applying (i) to $J = B_1$ and $S = B_2$ we find $|B_1| \leq |B_2|$. By interchanging the roles of $B_1$ and $B_2$ we get $|B_2| \leq |B_1|$. Hence $|B_1| = |B_2|$ by the Schröder-Bernstein theorem.

QED

Theorems A.5 and A.9 justify the next definition.

A.10. DEFINITION. The dimension of $V$, denoted $\dim V$ or $\dim_{F} V$, is the
cardinality of any basis of $V$.

Exercises

In the following problems $V$ and $W$ denote not necessarily finite-dimensional vector
spaces over a fixed field $F$.

A.1. Let $W_1$ be a subspace of $V$.

(i) There exists a basis $S$ of $V$ such that $S \cap W_1$ is a basis of $W_1$. Hence $\dim W_1 \leq \dim V$.

(ii) There exists a linear subspace $W_2$ of $V$, called a complement to $W_1$, such that
$W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$.

(Use the basis theorem.)

A.3. Prove Lemma A.7. (Don't use the dimension theorem, or your proof will be circular. Let $V$ be a vector space; suppose $S = \{w_1, w_2, \ldots, w_m\}$ is a finite spanning set. By the basis theorem we may assume, after selecting a subset of $S$ if necessary, that $S$ is a basis. Suppose $J$ was an independent subset of $V$ containing $n > m$ elements and try to get a contradiction, for instance by resorting to the theory of linear equations.)

A.4. (i) Prove Lemma A.8. (Argue by contradiction, assuming that $X$ is infinite. Show that $X$ must contain an infinite sequence of subspaces $E_1, E_2, E_3, \ldots$ such that either $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$ or $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$.)

(ii) Give a counterexample to the conclusion of Lemma A.8 if you drop the assumption that $U$ is finite-dimensional.

(iii) Give a counterexample to the conclusion of Lemma A.8 if you drop the assumption that $X$ is totally ordered.

A.5. Let $f: V \to W$ be a linear map.

(i) $f$ is surjective if and only if there exists a linear map $g: W \to V$ (called a right inverse of $f$) such that $f \circ g = \text{id}_W$.

(ii) $f$ is injective if and only if there exists a linear map $g: W \to V$ (called a left inverse of $f$) such that $g \circ f = \text{id}_V$.

(Use the basis theorem.)

A.6 (the pigeonhole principle for linear maps). Suppose $f: V \to W$ is a linear map and that $\dim V = \dim W < \infty$. The following statements are equivalent.

(i) $f$ is injective.

(ii) $f$ is surjective.

(iii) $f$ is bijective (and hence an isomorphism).

(Use Exercise A.5.) Give counterexamples to (i) $\implies$ (ii) and to (ii) $\implies$ (i) when $V$ is infinite-dimensional.
The Greek alphabet

<table>
<thead>
<tr>
<th>upper case</th>
<th>lower case</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>B</td>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>Γ</td>
<td>γ</td>
<td>gamma</td>
</tr>
<tr>
<td>Δ</td>
<td>δ</td>
<td>delta</td>
</tr>
<tr>
<td>E</td>
<td>ε, ϵ</td>
<td>epsilon</td>
</tr>
<tr>
<td>Z</td>
<td>ζ</td>
<td>zeta</td>
</tr>
<tr>
<td>H</td>
<td>η</td>
<td>eta</td>
</tr>
<tr>
<td>Θ</td>
<td>θ, θ</td>
<td>theta</td>
</tr>
<tr>
<td>I</td>
<td>ι</td>
<td>iota</td>
</tr>
<tr>
<td>K</td>
<td>κ</td>
<td>kappa</td>
</tr>
<tr>
<td>Λ</td>
<td>λ</td>
<td>lambda</td>
</tr>
<tr>
<td>M</td>
<td>μ</td>
<td>mu</td>
</tr>
<tr>
<td>N</td>
<td>ν</td>
<td>nu</td>
</tr>
<tr>
<td>Ξ</td>
<td>ξ</td>
<td>xi</td>
</tr>
<tr>
<td>Ο</td>
<td>ο</td>
<td>omicron</td>
</tr>
<tr>
<td>Π</td>
<td>π, ϖ</td>
<td>pi</td>
</tr>
<tr>
<td>P</td>
<td>ρ</td>
<td>rho</td>
</tr>
<tr>
<td>Σ</td>
<td>σ</td>
<td>sigma</td>
</tr>
<tr>
<td>T</td>
<td>τ</td>
<td>tau</td>
</tr>
<tr>
<td>Υ</td>
<td>υ</td>
<td>upsilon</td>
</tr>
<tr>
<td>Φ</td>
<td>φ, ϕ</td>
<td>phi</td>
</tr>
<tr>
<td>Χ</td>
<td>χ</td>
<td>chi</td>
</tr>
<tr>
<td>Ψ</td>
<td>ψ</td>
<td>psi</td>
</tr>
<tr>
<td>Ω</td>
<td>ω</td>
<td>omega</td>
</tr>
</tbody>
</table>