EXTERIOR POWERS

KEITH CONRAD

1. Introduction

Let R be a commutative ring. Unless indicated otherwise, all modules are R-modules and all tensor products are taken over R, so we abbreviate $\otimes_R$ to $\otimes$. A bilinear function out of $M_1 \times M_2$ turns into a linear function out of the tensor product $M_1 \otimes M_2$. In a similar way, a multilinear function out of $M_1 \times \cdots \times M_k$ turns into a linear function out of the k-fold tensor product $M_1 \otimes \cdots \otimes M_k$. We will concern ourselves with the case when the component modules are all the same: $M_1 = \cdots = M_k = M$. The tensor power $M^{\otimes k}$ universally linearizes all the multilinear functions on M^k.

A function on M^k can have its variables permuted to give new functions on M^k. When looking at permutations of the variables, two important types of functions on M^k occur: symmetric and alternating. These will be defined in Section 2. We will introduce in Section 3 the module that universally linearizes the alternating multilinear functions on M^k: the exterior power $\Lambda^k(M)$. It is a certain quotient module of $M^{\otimes k}$. The special case of exterior powers of finite free modules will be examined in Section 4. Exterior powers will be extended from modules to linear maps in Section 5. Applications of exterior powers to determinants are in Section 6 and to linear independence are in Section 7. Section 8 will introduce a product $\Lambda^k(M) \times \Lambda^\ell(N) \to \Lambda^{k+\ell}(M)$ called the wedge product, which is applied in Section 9 to a question about torsion modules over a PID. Finally, in Section 10 we will use the wedge product to turn the direct sum of all the exterior powers of a module into a noncommutative ring called the exterior algebra of the module.

The exterior power construction is important in geometry, where it provides the language for discussing differential forms on manifolds. (A differential form on a manifold is related to exterior powers of the dual space of the tangent space of a manifold at each of its points.) Exterior powers also arise in representation theory, as one of several ways of creating new representations of a group from a given representation of the group. In linear algebra, exterior powers provide an algebraic mechanism for detecting linear relations among vectors and for studying the "geometry" of the subspaces of a vector space.

2. Symmetric and alternating functions

For any function $f\colon M^k \to N$ and any $\sigma \in S_k$, we get a new function $M^k \to N$ by permuting the variables in f according to σ:

$$(m_1, \ldots, m_k) \mapsto f(m_{\sigma(1)}, \ldots, m_{\sigma(k)}) \in N.$$

(Warning: if we regard this new function on M^k as the effect of σ on f, and write it as $(\sigma \cdot f)(m_1, \ldots, m_k)$, then $\sigma_1 \cdot (\sigma_2 \cdot f)$ equals $(\sigma_2\sigma_1) \cdot f$, *not* $(\sigma_1\sigma_2) \cdot f$, so we don't have a left action of S_k on the functions $M^k \to N$ but a right action. We won't be using group actions, so don't worry about this.)

Definition 2.1. We call $f\colon M^k \to N$ *symmetric* if

$$f(m_{\sigma(1)}, \ldots, m_{\sigma(k)}) = f(m_1, \ldots, m_k)$$

for all $\sigma \in S_k$. We call f *skew-symmetric* if

$$f(m_{\sigma(1)}, \ldots, m_{\sigma(k)}) = (\operatorname{sign} \sigma) f(m_1, \ldots, m_k) \tag{2.1}$$

for all $\sigma \in S_k$. We call f *alternating* if

$$f(m_1, \ldots, m_k) = 0 \text{ whenever } m_i = m_j \text{ for some } i \neq j$$

when $k \geq 2$, and for $k = 1$ that constraint is empty, so all linear maps $f\colon M \to N$ are alternating.

Symmetric functions are unchanged by each permutation of the variables, while skew-symmetric functions are unchanged by even permutations and change sign under odd permutations. For example, the value of a skew-symmetric function changes by a sign if we permute any two of the variables. Alternating functions are not so intuitive. When R is a field of characteristic 0, like the real numbers, we will see that alternating and skew-symmetric *multilinear* functions are the same thing (Theorem 2.10).

Example 2.2. The function $\mathrm{M}_n(R) \times \mathrm{M}_n(R) \to R$ by $(A, B) \mapsto \operatorname{Tr}(AB)$ is symmetric.

Example 2.3. The function $R^2 \times R^2 \to R$ given by $(\binom{a}{c}, \binom{b}{d}) \mapsto ad - bc$ is skew-symmetric and alternating.

Example 2.4. The cross product $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3$ is skew-symmetric and alternating.

Example 2.5. The function $\mathbf{C} \times \mathbf{C} \to \mathbf{R}$ given by $(z, w) \mapsto \operatorname{Im}(z\overline{w})$ is skew-symmetric and alternating.

Example 2.6. Let R contain $\mathbf{Z}/2\mathbf{Z}$, so $-1 = 1$ in R. The multiplication map $R \times R \to R$ is symmetric and skew-symmetric, but *not* alternating.

In Definition 2.1, the variables are indexed in the order from 1 to k. Let's show that a function $M^k \to N$ being symmetric, skew-symmetric, or alternating does not depend on that particular ordering. There's nothing to do when $k = 1$, so let $k \geq 2$.

Theorem 2.7. *For $k \geq 2$, fix a listing of the numbers from 1 to k as $i_1, i_2, \ldots, i_k$. If a function $f\colon M^k \to N$ is symmetric then*

$$f(m_{\sigma(i_1)}, \ldots, m_{\sigma(i_k)}) = f(m_{i_1}, \ldots, m_{i_k})$$

for all $\sigma \in S_k$. If f is skew-symmetric then

$$f(m_{\sigma(i_1)}, \ldots, m_{\sigma(i_k)}) = (\operatorname{sign} \sigma) f(m_{i_1}, \ldots, m_{i_k})$$

for all $\sigma \in S_k$. If f is alternating then

$$f(m_{i_1}, \ldots, m_{i_k}) = 0 \textit{ whenever } m_{i_s} = m_{i_t} \textit{ for some } i_s \neq i_t.$$

Proof. We will discuss the skew-symmetric case, leaving the other two cases to the reader. Let $\widetilde{\sigma} \in S_k$ be the permutation where $\widetilde{\sigma}(1) = i_1, \ldots, \widetilde{\sigma}(k) = i_k$. If f is skew-symmetric, then

$$f(m_{i_1}, \ldots, m_{i_k}) = f(m_{\widetilde{\sigma}(1)}, \ldots, m_{\widetilde{\sigma}(k)}) = (\operatorname{sign} \widetilde{\sigma}) f(m_1, \ldots, m_k),$$

so for any $\sigma \in S_k$

$$\begin{aligned}
f(m_{\sigma(i_1)}, \dots, m_{\sigma(i_k)}) &= f(m_{(\sigma\widetilde{\sigma})(1)}, \dots, m_{(\sigma\widetilde{\sigma})(k)}) \\
&= \operatorname{sign}(\sigma\widetilde{\sigma}) f(m_1, \dots, m_k) \\
&= \operatorname{sign}(\sigma) \operatorname{sign}(\widetilde{\sigma}) f(m_1, \dots, m_k) \\
&= \operatorname{sign}(\sigma) f(m_{i_1}, \dots, m_{i_k}). \qquad \square
\end{aligned}$$

The next two theorems explain the connection between alternating multilinear functions and skew-symmetric multilinear functions, which is suggested by the above examples.

Theorem 2.8. *A multilinear function $f \colon M^k \to N$ that is alternating is skew-symmetric.*

Proof. The case $k = 1$ is trivial: all linear (= 1-multilinear) maps $M \to N$ are alternating and skew-symmetric. For $k \geq 2$, the basic idea is already present in the case $k = 2$, so we'll treat $k = 2$ first. When $f \colon M^2 \to N$ is alternating, $f(m, m) = 0$ for all m in M. So for all m and m' in M,

$$f(m + m', m + m') = 0.$$

Expanding by linearity in each component,

$$f(m, m) + f(m, m') + f(m', m) + f(m', m') = 0.$$

The first and last terms are 0, so $f(m, m') = -f(m', m)$. This means f is skew-symmetric.

For the general case when $k \geq 2$, we want to show (2.1) for all $\sigma \in S_k$. Notice first that if f satisfies (2.1) for the permutations σ_1 and σ_2 in S_k then

$$\begin{aligned}
f(m_{(\sigma_1\sigma_2)(1)}, \dots, m_{(\sigma_1\sigma_2)(k)}) &= f(m_{\sigma_1(\sigma_2(1))}, \dots, m_{\sigma_1(\sigma_2(k))}) \\
&= (\operatorname{sign} \sigma_1) f(m_{\sigma_2(1)}, \dots, m_{\sigma_2(k)}) \\
&= (\operatorname{sign} \sigma_1)(\operatorname{sign} \sigma_2) f(m_1, \dots, m_k) \\
&= \operatorname{sign}(\sigma_1\sigma_2) f(m_1, \dots, m_k).
\end{aligned}$$

Hence to verify (2.1) for all $\sigma \in S_k$ it suffices to verify (2.1) as σ runs over a generating set of S_k. We will check (2.1) when σ runs over the generating set of transpositions

$$\{(i \ i+1) : 1 \leq i \leq k-1\}.$$

That is, for $(m_1, \dots, m_k) \in M^k$, we want to show any alternating multilinear function $f \colon M^k \to N$ satisfies

$$f(\dots, m_{i-1}, m_i, m_{i+1}, m_{i+2}, \dots) = -f(\dots, m_{i-1}, m_{i+1}, m_i, m_{i+2}, \dots), \tag{2.2}$$

where we only interchange the places of m_i and m_{i+1}.

Fix all components of f except those in positions i and $i+1$, reducing us to a function of two variables: choose $m_1, \dots, m_{i-1}, m_{i+2}, \dots, m_k \in M$ (an empty condition if $k = 2$) and let $g(x, y) = f(m_1, \dots, m_{i-1}, x, y, m_{i+2}, \dots)$. Then g is bilinear and alternating, so by the $k = 2$ case g is skew-symmetric: $g(x, y) = -g(y, x)$. This implies (2.2), so we're done. $\square$

Corollary 2.9. *For $k \geq 2$, a function $f \colon M^k \to N$ is skew-symmetric if and only if it satisfies*

$$f(m_1, \dots, m_{i+1}, m_i, \dots, m_k) = -f(m_1, \dots, m_i, m_{i+1}, \dots, m_k)$$

for $1 \leq i \leq k-1$, and f is alternating if and only if

$$f(m_1, \dots, m_i, m_{i+1}, \dots, m_k) = 0$$

whenever $m_i = m_{i+1}$ for $1 \leq i \leq k-1$.

Proof. The second paragraph of the proof of Theorem 2.8 applies to all functions, not just multilinear functions, so the condition equivalent to skew-symmetry follows.

If we now suppose f vanishes at any k-tuple with a pair of adjacent equal coordinates, then what we just proved shows f is skew-symmetric. Therefore the value of f at any k-tuple with a pair of equal coordinates is, up to sign, its value at a k-tuple with a pair of adjacent equal coordinates, and that value is 0 by hypothesis. □

Example 2.6 shows the converse of Theorem 2.8 can fail: a multilinear function can be skew-symmetric and not alternating. But if 2 is a unit in R (*e.g.*, $R = \mathbf{R}$) then the converse of Theorem 2.8 does hold:

Theorem 2.10. *Let $k \geq 1$. If $2 \in R^\times$ then a multilinear function $f \colon M^k \to N$ that is skew-symmetric is alternating.*

Proof. As with Theorem 2.8, this is trivial for $k = 1$ since all linear maps $M \to N$ are skew-symmetric and alternating. For $k \geq 2$, we will show $f(m_1, m_2, \ldots, m_k) = 0$ when $m_1 = m_2$. The argument when $m_i = m_j$ for other distinct pairs i and j is the same. By skew-symmetry,

$$f(m_2, m_1, m_3, \ldots, m_k) = -f(m_1, m_2, m_3, \ldots, m_k).$$

Therefore

$$f(m, m, m_3, \ldots, m_k) = -f(m, m, m_3, \ldots, m_k),$$

so

$$2f(m, m, m_3, \ldots, m_k) = 0.$$

Since 2 is in $R^\times$, $f(m, m, m_3, \ldots, m_k) = 0$. □

The assumption $2 \in R^\times$ in this theorem could be weakened to $2n = 0 \Rightarrow n = 0$ for $n \in N$, since that's the actual property we used about 2. For example, if $R = \mathbf{Z}$ and $N = \mathbf{Z}^d$, then the proof works even though $2 \notin \mathbf{Z}^\times$.

Over $\mathbf{R}$ and $\mathbf{C}$, the terms "alternating" and "skew-symmetric" can be used interchangeably for multilinear functions. For instance in [7], where only real vector spaces are used, multilinear functions satisfying the skew-symmetric property (2.1) are called alternating. At any point in our discussion where there is a noticeable difference between being alternating and being skew-symmetric, it is just a technicality, so don't worry about it.

3. Exterior powers of modules

Let M and N be R-modules and $k \geq 1$. When $f \colon M^k \to N$ is multilinear and $g \colon N \to P$ is linear, the composite $g \circ f$ is multilinear. And if f is symmetric, skew-symmetric, or alternating then $g \circ f$ has the same property too. So we can create new (symmetric, skew-symmetric, or alternating) multilinear maps from old ones by composing with a linear map.

The kth tensor power $M^k \xrightarrow{\otimes} M^{\otimes k}$, sending $(m_1, \ldots, m_k)$ to $m_1 \otimes \cdots \otimes m_k$, is one example of a multilinear map out of M^k, and every other example comes from this one: given any multilinear map $f \colon M^k \to N$ there is a unique linear map $\widetilde{f} \colon M^{\otimes k} \to N$ whose composite with $M^k \xrightarrow{\otimes} M^{\otimes k}$ is f. That is, there is a unique linear map $\widetilde{f} \colon M^{\otimes k} \to N$

making the diagram

$$\begin{tikzcd} M^k \arrow[r, "f"] \arrow[d, "\otimes"'] & N \\ M^{\otimes k} \arrow[ur, dotted, "\widetilde{f}"'] & \end{tikzcd}$$

commute, which means $\widetilde{f}(m_1 \otimes \cdots \otimes m_k) = f(m_1, \dots, m_k)$.

Now focus on multilinear $f\colon M^k \to N$ that are alternating. If $k = 1$ then f is arbitrary. If $k \geq 2$, then f vanishes on all k-tuples with a pair of equal coordinates, so $\widetilde{f}\colon M^{\otimes k} \to N$ vanishes on all tensors $m_1 \otimes \cdots \otimes m_k$ with $m_i = m_j$ for some $i \neq j$. Let J_k be the submodule spanned by these special tensors when $k \geq 2$, so $\mathrm{J}_k \subset \ker \widetilde{f}$. Set $\mathrm{J}_1 = \{0\}$ in $M^{\otimes 1} = M$, so $\mathrm{J}_1 \subset \ker \widetilde{f}$ when $k = 1$. The quotient module $M^{\otimes k}/\mathrm{J}_k$ will be our main object of interest.

Definition 3.1. For an R-module M and an integer $k \geq 1$, the kth *exterior power* of M, denoted $\Lambda^k(M)$, is the R-module $M^{\otimes k}/\mathrm{J}_k$ where $\mathrm{J}_k = \{0\}$ for $k = 1$ and J_k for $k \geq 2$ is the submodule of $M^{\otimes k}$ spanned by all $m_1 \otimes \cdots \otimes m_k$ with $m_i = m_j$ for some $i \neq j$. For $m_1, \dots, m_k \in M$, the coset of $m_1 \otimes \cdots \otimes m_k$ in $\Lambda^k(M)$ is denoted $m_1 \wedge \cdots \wedge m_k$. For completeness, set $\Lambda^0(M) = R$ (so $\mathrm{J}_0 = \{0\}$ in $M^{\otimes 0} = R$).

We could write $\Lambda^k_R(M)$ to place the ring R in our notation. But since we will never be changing the ring, we suppress this extra decoration. The general element of $\Lambda^k(M)$ will be denoted ω or η (as we write t for a general tensor). Since $M^{\otimes k}$ is spanned tensors of the form $m_1 \otimes \cdots \otimes m_k$, the quotient module $M^{\otimes k}/\mathrm{J}_k = \Lambda^k(M)$ is spanned by their images $m_1 \wedge \cdots \wedge m_k$. That is, any $\omega \in \Lambda^k(M)$ is a finite R-linear combination

$$\omega = \sum r_{i_1,\dots,i_k} m_{i_1} \wedge \cdots \wedge m_{i_k},$$

where the coefficients $r_{i_1,\dots,i_k}$ are in R and the m_i's are in M.

We call $m_1 \wedge m_2 \wedge \cdots \wedge m_k$ an *elementary wedge product* and read it as "m_1 wedge m_2 ... wedge m_k." Another name for elementary wedge products is decomposable elements. (More synonyms: simple, pure, monomial). Since $r(m_1 \wedge m_2 \wedge \cdots \wedge m_k) = (rm_1) \wedge m_2 \wedge \cdots \wedge m_k$, every element of $\Lambda^k(M)$ is a *sum* (not just a linear combination) of elementary wedge products.[1] A linear – or even additive – map out of $\Lambda^k(M)$ is completely determined by its values on elementary wedge products because they additively span $\Lambda^k(M)$. More general wedge products will be met in Section 8.

The modules $\Lambda^k(M)$ were introduced by Grassmann (for $M = \mathbf{R}^n$), who called expressions like $m_1 \wedge m_2$ outer products. Now we use the label "exterior" instead. Perhaps it would be better to call $\Lambda^k(M)$ an alternating power instead of an exterior power, but it's too late to change the terminology.

Example 3.2. Say M is spanned by two elements x and y: $M = Rx + Ry$. (This doesn't mean x and y are a basis, *e.g.*, $R = \mathbf{Z}[\sqrt{-5}]$ and $M = (2, 1 + \sqrt{-5})$.) We will show $\Lambda^2(M)$ is spanned by the single element $x \wedge y$. The tensor square $M^{\otimes 2}$ is spanned by all terms $m \otimes m'$ where m and m' are in M. Write $m = ax + by$ and $m' = cx + dy$. Then in $M^{\otimes 2}$

$$\begin{aligned} m \otimes m' &= (ax + by) \otimes (cx + dy) \\ &= ac(x \otimes x) + ad(x \otimes y) + bc(y \otimes x) + bd(y \otimes y). \end{aligned}$$

[1] Unlike "tensor" as a name for an element of a tensor product module, there is no standard name for an element ω of $\Lambda^k(M)$. It's not called a "wedge" but maybe it should be. Some may call ω a "wedge product" but (i) ω may not be a product and (ii) "wedge product" has another meaning: see Section 8.

The tensors $x \otimes x$ and $y \otimes y$ are in J_2, so in $\Lambda^2(M)$ both $x \wedge x$ and $y \wedge y$ vanish. Therefore

$$m \wedge m' = ad(x \wedge y) + bc(y \wedge x).$$

Moreover, the tensor $(x+y) \otimes (x+y)$ is in J_2, so

$$x \otimes y + y \otimes x = (x+y) \otimes (y+x) - x \otimes x - y \otimes y \in \mathrm{J}_2\,.$$

Therefore in $\Lambda^2(M)$ we have $x \wedge y + y \wedge x = 0$, so

$$m \wedge m' = ad(x \wedge y) + bc(-x \wedge y) = (ad - bc)(x \wedge y),$$

which means $\Lambda^2(M)$ is spanned by the single element $x \wedge y$. It could happen that $x \wedge y = 0$ (so $\Lambda^2(M)$ could be zero), or even if $x \wedge y \neq 0$ it could happen that $r(x \wedge y) = 0$ for some nonzero $r \in R$. It all depends on the nature of the R-linear relations between x and y in M.

By comparison to $\Lambda^2(M)$, $M^{\otimes 2}$ is spanned by $x \otimes x$, $x \otimes y$, $y \otimes x$, and $y \otimes y$, and without further information we have no reason to collapse this spanning set (usually $x \otimes y \neq y \otimes x$, for instance). So when M has a 2-element spanning set, a spanning set for $M^{\otimes 2}$ is typically larger than 2 while a spanning set for $\Lambda^2(M)$ is definitely smaller.

For $k \geq 1$, the standard map $M^k \xrightarrow{\otimes} M^{\otimes k}$ is multilinear and the reduction map $M^{\otimes k} \to M^{\otimes k}/\,\mathrm{J}_k = \Lambda^k(M)$ is linear, so the composite map $\wedge\colon M^k \xrightarrow{\otimes} M^{\otimes k} \to \Lambda^k(M)$ is multilinear. That is, the function

$$(m_1, \ldots, m_k) \mapsto m_1 \wedge \cdots \wedge m_k \tag{3.1}$$

from M^k to $\Lambda^k(M)$ is multilinear in the m_i's. For example,

$$m_1 \wedge \cdots \wedge cm_i \wedge \cdots \wedge m_k = c(m_1 \wedge \cdots \wedge m_i \wedge \cdots \wedge m_k) \text{ and } m_1 \wedge \cdots \wedge 0 \wedge \cdots \wedge m_k = 0.$$

For $k \geq 2$, $m_1 \wedge \cdots \wedge m_k = 0$ in $\Lambda^k(M)$ if $m_i = m_j$ for some $i \neq j$. (Think about what working modulo J_k means for the tensors $m_1 \otimes \cdots \otimes m_k$ when $m_i = m_j$ for some $i \neq j$.) Therefore (3.1) is an *example* of an alternating multilinear map out of M^k. Now we show it is a universal example: all others pass through it using linear maps out of $\Lambda^k(M)$.

Theorem 3.3. *For R-modules M and N and an alternating multilinear map $f\colon M^k \to N$ where $k \geq 1$, there is a unique linear map $\widetilde{f}\colon \Lambda^k(M) \to N$ such that the diagram*

$$\begin{tikzcd} M^k \arrow[r, "f"] \arrow[d, "\wedge"'] & N \\ \Lambda^k(M) \arrow[ur, dotted, "\widetilde{f}"'] & \end{tikzcd}$$

commutes, i.e., $\widetilde{f}(m_1 \wedge \cdots \wedge m_k) = f(m_1, \ldots, m_k)$.

This theorem makes no sense when $k = 0$, unless you want to define $M^0 = R$ and call all linear maps $R \to N$ alternating.

Proof. This is clear when $k = 1$ since $\Lambda^1(M) = M$ and $\wedge\colon M \to \Lambda^1(M)$ is the identity. For $k \geq 2$, f induces a linear map $M^{\otimes k} \to N$ whose effect on elementary tensors is

$$m_1 \otimes \cdots \otimes m_k \mapsto f(m_1, \ldots, m_k).$$

$$\begin{tikzcd} M^k \arrow[r, "f"] \arrow[d, "\otimes"'] & N \\ M^{\otimes k} \arrow[ur, dotted] & \end{tikzcd}$$

(We are going to avoid giving this map $M^{\otimes k} \to N$ a specific notation, since it is just an intermediate device in this proof.) Because f is alternating, it vanishes at any k-tuple $(m_1, \ldots, m_k)$ where $m_i = m_j$ for some $i \neq j$. Thus the linear map that f induces from $M^{\otimes k}$ to N vanishes at any elementary tensor $m_1 \otimes \cdots \otimes m_k$ where $m_i = m_j$ for some $i \neq j$. Hence this linear map out of $M^{\otimes k}$ vanishes on the submodule J_k of $M^{\otimes k}$, so we get an induced linear map $\widetilde{f}$ out of $M^{\otimes k}/J_k = \Lambda^k(M)$. Specifically, $\widetilde{f} \colon \Lambda^k(M) \to N$ is given by

$$m_1 \wedge \cdots \wedge m_k \mapsto f(m_1, \ldots, m_k).$$

$$\begin{tikzcd} M^k \arrow[r, "f"] \arrow[d] & N \\ M^{\otimes k} \arrow[ur, dotted] \arrow[d] & \\ \Lambda^k(M) \arrow[uur, dotted, "\widetilde{f}"'] & \end{tikzcd}$$

Since the elements $m_1 \wedge \cdots \wedge m_k$ span $\Lambda^k(M)$, a linear map out of $\Lambda^k(M)$ is uniquely determined by its effect on these elements. Thus, having constructed a linear map out of $\Lambda^k(M)$ whose effect on any $m_1 \wedge \cdots \wedge m_k$ is the same as the effect of f on $(m_1, \ldots, m_k)$, it is the unique such linear map. $\square$

For $k \geq 1$, call the specific alternating multilinear map $M^k \xrightarrow{\wedge} \Lambda^k(M)$ given by

$$(m_1, \ldots, m_k) \mapsto m_1 \wedge \cdots \wedge m_k$$

the *canonical map.*

Remark 3.4. We could have constructed $\Lambda^k(M)$ as the quotient of a huge free module on the set M^k, bypassing the use of $M^{\otimes k}/\,\mathrm{J}_k$. Then, since the canonical map $M^k \xrightarrow{\wedge} \Lambda^k(M)$ is multilinear, we would get a linear map $M^{\otimes k} \longrightarrow \Lambda^k(M)$ and could recover $\Lambda^k(M)$ as a quotient of $M^{\otimes k}$ anyway.

Corollary 3.5. *Suppose Λ is an R-module and there is an alternating multilinear map $\alpha \colon M^k \to \Lambda$ with the same universal mapping property as the canonical map $M^k \xrightarrow{\wedge} \Lambda^k(M)$: for every R-module N and alternating multilinear map $f \colon M^k \to N$ there is a unique linear map $\Lambda \to N$ making the diagram*

$$\begin{tikzcd} M^k \arrow[r, "f"] \arrow[d, "\alpha"'] & N \\ \Lambda \arrow[ur, dotted] & \end{tikzcd}$$

commute. Then there is a unique R-linear map $\beta\colon \Lambda \to \Lambda^k(M)$ such that the diagram

$$\begin{tikzcd} & M^k \arrow[dl, "\alpha"'] \arrow[dr, "\wedge"] & \\ \Lambda \arrow[rr, "\beta"'] & & \Lambda^k(M) \end{tikzcd}$$

commutes, and β is an isomorphism.

Proof. This is the usual argument that an object equipped with a map satisfying a universal mapping property is determined up to a unique isomorphism: set two such objects and maps against each other to get maps between the objects in both directions whose composites in both orders have to be the identity maps on the two objects by the usual argument. $\square$

Since the canonical map $M^k \xrightarrow{\wedge} \Lambda^k(M)$ is alternating multilinear, by Theorem 2.8 it is skew-symmetric:

$$m_{\sigma(1)} \wedge \cdots \wedge m_{\sigma(k)} = (\operatorname{sign}\sigma) m_1 \wedge \cdots \wedge m_k \tag{3.2}$$

for every $\sigma \in S_k$. In particular, an elementary wedge product $m_1 \wedge \cdots \wedge m_k$ in $\Lambda^k(M)$ is determined up to an overall sign by the terms m_i appearing in it (*e.g.*, $m \wedge m' \wedge m'' = -m' \wedge m \wedge m''$).

Example 3.6. Returning to Example 3.2 and working directly in $\Lambda^2(M)$ from the start, we have

$$(ax+by) \wedge (cx+dy) = ac(x \wedge x) + ad(x \wedge y) + bc(y \wedge x) + bd(y \wedge y) \tag{3.3}$$

by multilinearity. Since (3.1) is alternating, $x \wedge x$ and $y \wedge y$ vanish. By (3.2), $y \wedge x = -x \wedge y$. Feeding this into (3.3) gives

$$(ax+by) \wedge (cx+dy) = ad(x \wedge y) - bc(x \wedge y) = (ad-bc)(x \wedge y),$$

so $\Lambda^2(M)$ is spanned by $x \wedge y$ when M is spanned by x and y. That was faster than Example 3.2!

Example 3.7. Suppose M is spanned by three elements e_1, e_2, and e_3. We will show $e_1 \wedge e_2$, $e_1 \wedge e_3$, and $e_2 \wedge e_3$ span $\Lambda^2(M)$. We know by the definition of $\Lambda^2(M)$ that $\Lambda^2(M)$ is spanned by all $m \wedge m'$, so it suffices to show every $m \wedge m'$ is a linear combination of $e_1 \wedge e_2$, $e_1 \wedge e_3$, and $e_2 \wedge e_3$. Writing

$$m = ae_1 + be_2 + ce_3, \quad m' = a'e_1 + b'e_2 + c'e_3,$$

the multilinearity and the alternating property ($e_i \wedge e_i = 0$ and $e_i \wedge e_j = -e_j \wedge e_i$ for $i \neq j$) imply

$$\begin{aligned} m \wedge m' &= (ae_1 + be_2 + ce_3) \wedge (a'e_1 + b'e_2 + c'e_3) \\ &= ae_1 \wedge (b'e_2 + c'e_3) + be_2 \wedge (a'e_1 + c'e_3) + ce_3 \wedge (a'e_1 + b'e_2) \\ &= ab'(e_1 \wedge e_2) + ac'(e_1 \wedge e_3) + ba'(e_2 \wedge e_1) + bc'(e_2 \wedge e_3) + \\ &\quad\ ca'(e_3 \wedge e_1) + cb'(e_3 \wedge e_2) \\ &= (ab' - ba')(e_1 \wedge e_2) + (ac' - ca')(e_1 \wedge e_3) + (bc' - cb')(e_2 \wedge e_3). \end{aligned}$$

If we write this formula with the first and third terms exchanged as

$$m \wedge m' = (bc' - cb')(e_2 \wedge e_3) + (ac' - ca')(e_1 \wedge e_3) + (ab' - ba')(e_1 \wedge e_2), \tag{3.4}$$

it looks quite close to the cross product on $\mathbf{R}^3$:

$$(a, b, c) \times (a', b', c') = (bc' - cb', -(ac' - ca'), ab' - ba').$$

(There is a way of making $e_3 \wedge e_1$ the more natural wedge product than $e_1 \wedge e_3$ in $\Lambda^2(\mathbf{R}^3)$, so (3.4) would then match the coordinates of the cross-product everywhere. This uses the Hodge-star operator. We don't discuss that here.)

In a tensor power $M^{\otimes k}$ for $k \geq 2$, each tensor is a sum of elementary tensors but usually most tensors are not elementary. The same is true in $\Lambda^k(M)$ for $k \geq 2$: a general element is a sum of elementary wedge products, but is usually not an elementary wedge product.

Example 3.8. Let M be spanned by e_1, e_2, e_3, and e_4. The exterior square $\Lambda^2(M)$ is spanned by the pairs

$$e_1 \wedge e_2,\ e_1 \wedge e_3,\ e_1 \wedge e_4,\ e_2 \wedge e_3,\ e_2 \wedge e_4,\ e_3 \wedge e_4.$$

When M is free and $\{e_1, e_2, e_3, e_4\}$ is a basis of M, the sum $e_1 \wedge e_2 + e_3 \wedge e_4$ in $\Lambda^2(M)$ is *not* an elementary wedge product: it can't be expressed in the form $m \wedge m'$. We'll see why (in most cases) in Example 8.9. On the other hand, the sum

$$e_1 \wedge e_2 + 3(e_1 \wedge e_3) + 3(e_1 \wedge e_4) + 2(e_2 \wedge e_3) + 2(e_2 \wedge e_4)$$

in $\Lambda^2(M)$ doesn't look like an elementary wedge product but it is! It equals

$$(e_1 + e_2 + e_3 + e_4) \wedge (e_2 + 3e_3 + 3e_4).$$

Check equality by expanding this out using multilinearity and the relations $e_i \wedge e_i = 0$ and $e_i \wedge e_j = -e_j \wedge e_i$ for $i \neq j$.

A linear map out of $\Lambda^k(M)$ is completely determined by its values on the elementary wedge products $m_1 \wedge \cdots \wedge m_k$, since they span the module. But elementary wedge products, like elementary tensors, are not linearly independent, so verifying there is a linear map out of $\Lambda^k(M)$ with some prescribed behavior on all the elementary wedge products has to be done carefully. Proceed by first introducing a function on M^k that is multilinear and alternating whose value at $(m_1, \ldots, m_k)$ is what you want the value to be at $m_1 \wedge \cdots \wedge m_k$, and then it automatically factors through $\Lambda^k(M)$ as a linear map with the desired value at $m_1 \wedge \cdots \wedge m_k$. This is like creating homomorphisms out of a quotient group G/N by first making a homomorphism out of G with the desired values and then checking N is in the kernel.

Example 3.9. There is a unique linear map $\Lambda^2(M) \to M^{\otimes 2}$ such that

$$m_1 \wedge m_2 \mapsto m_1 \otimes m_2 - m_2 \otimes m_1.$$

To construct such a map, start by letting $f\colon M^2 \to M^{\otimes 2}$ by $f(m_1, m_2) = m_1 \otimes m_2 - m_2 \otimes m_1$. This is bilinear and $f(m, m) = 0$, so f is alternating and thus there is a linear map $\Lambda^2(M) \to M^{\otimes 2}$ sending any elementary wedge product $m_1 \wedge m_2$ to $f(m_1, m_2) = m_1 \otimes m_2 - m_2 \otimes m_1$. Since a linear map out of $\Lambda^2(M)$ is determined by its values on elementary wedge products, f is the only linear map with the given values on all $m_1 \wedge m_2$.

Here are some basic questions about $\Lambda^k(M)$ for $k \geq 1$.

Questions

(1) What does it mean to say $m_1 \wedge \cdots \wedge m_k = 0$ in $\Lambda^k(M)$?
(2) What does it mean to say $\Lambda^k(M) = 0$?
(3) What does it mean to say $m_1 \wedge \cdots \wedge m_k = m_1' \wedge \cdots \wedge m_k'$?

Answers

(1) Saying $m_1\wedge\cdots\wedge m_k = 0$ means every alternating multilinear map out of M^k vanishes at $(m_1,\dots,m_k)$. Indeed, since every alternating multilinear map out of M^k induces a linear map out of $\Lambda^k(M)$ that sends $m_1\wedge\cdots\wedge m_k$ to the same place as $(m_1,\dots,m_k)$, if $m_1\wedge\cdots\wedge m_k = 0$ then the linear map out of $\Lambda^k(M)$ must send $m_1\wedge\cdots\wedge m_k$ to 0 (linear maps send 0 to 0) so the original alternating multilinear map we started with out of M^k has to equal 0 at $(m_1,\dots,m_k)$. Conversely, if every alternating multilinear map out of M^k sends $(m_1,\dots,m_k)$ to 0 then $m_1\wedge\cdots\wedge m_k = 0$ because the canonical map $M^k \xrightarrow{\wedge} \Lambda^k(M)$ is a particular example of an alternating multilinear map out of M^k and it sends $(m_1,\dots,m_k)$ to $m_1\wedge\cdots\wedge m_k$. (This reasoning is valid at $k = 1$: $\Lambda^1(M) = M$ and all linear maps out of M vanish at m only if $m = 0$ since the identity $M \to M$ is an example of an alternating 1-multilinear map = linear map.)

A lesson: to prove a specific elementary wedge product $m_1\wedge\cdots\wedge m_k$ is *not* 0, find an alternating multilinear map on M^k that is not 0 at $(m_1,\dots,m_k)$.

(2) To say $\Lambda^k(M) = 0$ means every alternating multilinear map on M^k is identically 0.

A lesson: to show $\Lambda^k(M) \neq 0$, find an example of an alternating multilinear map on M^k that is not identically 0.

(3) The condition $m_1\wedge\cdots\wedge m_k = m'_1\wedge\cdots\wedge m'_k$ means every alternating multilinear map on M^k takes the same values at $(m_1,\dots,m_k)$ and at $(m'_1,\dots,m'_k)$.

Remark 3.10. Unlike $M\otimes_R N$, which can be defined for different R-modules M and N, there is no $M\wedge N$ for two unrelated R-modules M and N. That is because the concept of exterior power is bound up with the idea of alternating multilinear functions, and permuting variables in a multivariable function only makes sense when the function has its variables coming from the *same* module.

4. Spanning sets for exterior powers

Let's look more closely at spanning sets of an exterior power module. If M is finitely generated (not necessarily free!), with spanning set $x_1,\dots,x_d$, then any tensor power $M^{\otimes k}$ is finitely generated as an R-module by the d^k tensors $x_{i_1}\otimes\cdots\otimes x_{i_k}$ where $1\le i_1,\dots,i_k\le d$: at first we know $M^{\otimes k}$ is spanned by all the elementary tensors $m_1\otimes\cdots\otimes m_k$, but write each m_i as an R-linear combination of $x_1,\dots,x_d$ and then expand out using the multilinearity of $\otimes$ to express every elementary tensor in $M^{\otimes k}$ as an R-linear combination of the tensors $x_{i_1}\otimes\cdots\otimes x_{i_k}$. (As a general rule this spanning set for $M^{\otimes k}$ can't be reduced further: when M is free and $x_1,\dots,x_d$ is a basis then the d^k elementary tensors $x_{i_1}\otimes\cdots\otimes x_{i_k}$ are a basis of $M^{\otimes k}$.) Since $\Lambda^k(M)$ is a quotient module of $M^{\otimes k}$, it is spanned as an R-module by the d^k elementary wedge products $x_{i_1}\wedge\cdots\wedge x_{i_k}$ where $1\le i_1,\dots,i_k\le d$. Thus exterior powers of a finitely generated R-module are finitely generated.

Theorem 4.1. *If M has a d-element spanning set for $d\ge 1$, then $\Lambda^k(M) = \{0\}$ for $k > d$.*

For example, $\Lambda^2(R) = 0$, and more generally $\Lambda^k(R^d) = 0$ for $k > d$.

Proof. Let $x_1,\dots,x_d$ span M. When $k > d$, so $k\ge 2$, each $x_{i_1}\wedge\cdots\wedge x_{i_k}$ contains two equal terms, so it is zero in in $\Lambda^k(M)$. Thus $\Lambda^k(M)$ is spanned by 0, so it is 0. □

For $d > 1$ and $2\le k\le d$, there is a lot of redundancy in the d^k elementary wedge products $x_{i_1}\wedge\cdots\wedge x_{i_k}$ coming from a spanning set $\{x_1,\dots,x_d\}$ of M. For instance, by the alternating property such an elementary wedge product vanishes if two terms in it are equal. Therefore we can discard from our spanning set for $\Lambda^k(M)$ those $x_{i_1} \wedge \cdots \wedge x_{i_k}$ where an x_i appears twice and we are still left with a spanning set. Moreover, by (3.2) two elementary wedge products containing the same factors in different order are equal up to sign, so our spanning set for $\Lambda^k(M)$ as an R-module can be reduced further to the elements $x_{i_1} \wedge \cdots \wedge x_{i_k}$ where the indices are strictly increasing: $1 \leq i_1 < \cdots < i_k \leq d$. The number of such k-tuples of indices is $\binom{d}{k}$, including when $k = 1$. So $\Lambda^k(M)$ has a spanning set of size $\binom{d}{k}$, which may or may not be reducible further. We will now prove that there can be no further reduction when M is free of rank d and $x_1, \ldots, x_d$ is a basis of M.

Theorem 4.2. *If M is a free R-module and $k \geq 0$, then $\Lambda^k(M)$ is a free R-module.*

Explicitly, if M is finite free of rank d and $k \leq d$ then $\Lambda^k(M)$ is free of rank $\binom{d}{k}$: for a basis $e_1, \ldots, e_d$ of M, the $\binom{d}{k}$ elementary wedge products

$$e_{i_1} \wedge \cdots \wedge e_{i_k} \text{ where } 1 \leq i_1 < \cdots < i_k \leq d$$

are a basis of $\Lambda^k(M)$. If $k > d$ then $\Lambda^k(M) = 0$.

If M has an infinite basis $\{e_i\}_{i \in I}$ and we put a well-ordering on the index set I, then for each $k \geq 1$ a basis of $\Lambda^k(M)$ is $\{e_{i_1} \wedge \cdots \wedge e_{i_k}\}_{i_1 < i_2 < \cdots < i_k}$.

Theorem 4.2 is the first nontrivial result about exterior powers, as it tells us a situation where exterior powers are guaranteed to be nonzero, and in fact be "as big as possible." Read the proof closely, as otherwise you may feel somewhat uneasy about exactly why exterior powers of free modules must be free.

Proof. Let $k \geq 1$, since $k = 0$ is trivial. . (We may also let $M \neq 0$: the zero module is free of rank $d = 0$ and $\binom{0}{k}$ is 1 for $k = 0$ and is 0 for $k \geq 1$.) The idea of the proof is to embed $\Lambda^k(M)$ into $M^{\otimes k}$ and exploit what we already know about $M^{\otimes k}$ for free M. The embedding we use (see (4.1)) may look like it comes out of nowhere, but it is common in differential geometry and we make some remarks about this after the proof.

For $k \geq 1$, the function $M^k \to M^{\otimes k}$ that is given by

$$(m_1, \ldots, m_k) \mapsto \sum_{\sigma \in S_k} (\operatorname{sign} \sigma) m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(k)}$$

is multilinear since each summand contains each m_i once. This function is also alternating. That's obvious for $k = 1$. For $k \geq 2$, by Corollary 2.9 it suffices to check the function vanishes at k-tuples with adjacent equal coordinates. If $m_i = m_{i+1}$ then for each $\sigma \in S_k$ the terms in the sum at σ and $\sigma(i\ i+1)$ are negatives of each other. Now the universal mapping property of exterior powers says there is an R-linear map $\alpha_{k,M} \colon \Lambda^k(M) \to M^{\otimes k}$ such that

$$\alpha_{k,M}(m_1 \wedge \cdots \wedge m_k) = \sum_{\sigma \in S_k} (\operatorname{sign} \sigma) m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(k)}. \tag{4.1}$$

The case $k = 2$ is Example 3.9 and $\alpha_{1,M} \colon M \to M$ is the identity.

Although $\alpha_{k,M}$ exists for all M, injectivity of $\alpha_{k,M}$ is not a general property. Our proof of injectivity for free M will use a basis. We'll write the proof with a finite basis, and the reader can make changes to see the same argument works if M has an infinite basis.

Assume M has a finite basis $e_1, \ldots, e_d$ and $k \leq d$ (since $\Lambda^k(M) = 0$ for $k > d$), so $d \geq 1$. Since the e_i's span M as an R-module, the $\binom{d}{k}$ elementary wedge products

$$e_{i_1} \wedge \cdots \wedge e_{i_k} \text{ where } 1 \leq i_1 < \cdots < i_k \leq d \tag{4.2}$$

span $\Lambda^k(M)$. (Here we need that the indices on the basis of M are totally ordered.) We know already that $M^{\otimes k}$ has a *basis*

$$e_{i_1} \otimes \cdots \otimes e_{i_k} \text{ where } 1 \leq i_1, \ldots, i_k \leq d,$$

where no inequalities are imposed on the indices.

Suppose $\omega \in \Lambda^k(M)$ satisfies $\alpha_{k,M}(\omega) = 0$. Write

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq d} c_{i_1,\ldots,i_k} e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

with $c_{i_1,\ldots,i_k} \in R$. Then the condition $\alpha_{k,M}(\omega) = 0$ becomes

$$\sum_{1 \leq i_1 < \cdots < i_k \leq d} c_{i_1,\ldots,i_k} \sum_{\sigma \in S_k} (\operatorname{sign} \sigma) e_{i_{\sigma(1)}} \otimes \cdots \otimes e_{i_{\sigma(k)}} = 0,$$

which is the same as

$$\sum_{\sigma \in S_k} \sum_I (\operatorname{sign} \sigma) c_I e_{\sigma(I)} = 0,$$

where I runs over all strictly increasing k-tuples $(i_1, \ldots, i_k)$ from 1 to d, with c_I and $e_{\sigma(I)}$ having an obvious meaning in this context. The vectors $\{e_{\sigma(I)}\}_{\sigma,I}$ are a *basis* of $M^{\otimes k}$, so all c_I are 0. This proves $\alpha_{k,M}$ is injective and it also shows (4.2) is a linearly independent subset of $\Lambda^k(M)$, so it is a basis (spans and is linearly independent). □

Exterior powers are closely connected to determinants, and most proofs of Theorem 4.2 for finite free M use the determinant. What we used in lieu of theorems about determinants is our knowledge of bases of tensor powers of a free module. For aesthetic reasons, we want to come back later and prove properties of the determinant using exterior powers, so we did not use the determinant directly in the proof of Theorem 4.2. However, the linear map $\alpha_{k,M}$ looks a lot like a determinant.

When V is a finite-dimensional vector space, it is free so $\alpha_{k,V} \colon \Lambda^k(V) \hookrightarrow V^{\otimes k}$ given by (4.1) is an embedding. It means we can think of $\Lambda^k(V)$ as a *subspace* of the tensor power $V^{\otimes k}$ instead of as a quotient space. This viewpoint is widely used in differential geometry, where vector spaces are defined over $\mathbf{R}$ or $\mathbf{C}$ and the image of $\Lambda^k(V)$ in $V^{\otimes k}$ is called the subspace of skew-symmetric tensors. The embedding $\alpha_{k,V}$ has an unfortunate scaling problem: when we embed $\Lambda^k(V)$ into $V^{\otimes k}$ with $\alpha_{k,V}$ and then reduce $V^{\otimes k}$ back to $\Lambda^k(V)$ with the canonical map $\wedge$, you might think the composite map $\Lambda^k(V) \xrightarrow{\alpha_{k,V}} V^{\otimes k} \xrightarrow{\wedge} \Lambda^k(V)$ is the identity on $\Lambda^k(V)$, but it is multiplication by $k!$. That is obvious when $k = 1$. To verify this for $k \geq 2$, it suffices to check it on elementary wedge products:

$$\begin{aligned}
v_1 \wedge \cdots \wedge v_k &\mapsto \sum_{\sigma \in S_k} (\operatorname{sign} \sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)} \\
&\mapsto \sum_{\sigma \in S_k} (\operatorname{sign} \sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(k)} \\
&= \sum_{\sigma \in S_k} (\operatorname{sign} \sigma)(\operatorname{sign} \sigma) v_1 \wedge \cdots \wedge v_k \\
&= k! v_1 \wedge \cdots \wedge v_k.
\end{aligned}$$

This suggests a better embedding of $\Lambda^k(V)$ into $V^{\otimes k}$ is $\frac{1}{k!}\alpha_{k,V}$, which is given by the formula

$$v_1 \wedge \cdots \wedge v_k \mapsto \frac{1}{k!}\sum_{\sigma \in S_k} (\operatorname{sign}\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}.$$

The composite map $\Lambda^k(V) \xrightarrow{(1/k!)\alpha_{k,V}} V^{\otimes k} \xrightarrow{\wedge} \Lambda^k(V)$ is the identity, but this rescaled embedding only makes sense if $k! \neq 0$ in the scalar field. That is fine for real and complex vector spaces (as in differential geometry), but it is not a universal method. So either you can take your embedding $\Lambda^k(V) \hookrightarrow V^{\otimes k}$ using $\alpha_{k,V}$ for all vector spaces and make $\Lambda^k(V) \longrightarrow V^{\otimes k} \longrightarrow \Lambda^k(V)$ be multiplication by $k!$, or you can have an embedding $\Lambda^k(V) \hookrightarrow V^{\otimes k}$ that only makes sense when $k! \neq 0$. Either way, this mismatch between $\Lambda^k(V)$ as a quotient space of $V^{\otimes k}$ (correct definition) and as a subspace of $V^{\otimes k}$ (incorrect but widely used definition) leads to a lot of excess factorials in formulas when exterior powers are regarded as subspaces of tensor powers instead of as quotient spaces of tensor powers.

There are other approaches to the proof of Theorem 4.2 when M is finite free. In [1, pp. 90–91] an explicit finite free R-module is constructed that has the same universal mapping property as $\Lambda^k(M)$, so $\Lambda^k(M)$ has to be finite free by Corollary 3.5. The other aspects of Theorem 4.2 (the rank of $\Lambda^k(M)$ and an explicit basis) can be read off from the proof in [1]. In [6, pp. 747–751], Theorem 4.2 for finite free M is proved using what is called there the Grassmann algebra of M (which we will meet later under the label exterior algebra of M).

When M is free of rank d and $k \leq d$, the basis $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \leq i_1 < \cdots < i_k \leq d\}$ of $\Lambda^k(M)$ will be called the *corresponding basis* from the choice of basis $e_1, \ldots, e_d$ of M.

Example 4.3. If M is free of rank d with basis $e_1, \ldots, e_d$ then $\Lambda^2(M)$ is free of rank $\binom{d}{2}$ with corresponding basis $\{e_i \wedge e_j : 1 \leq i < j \leq d\}$.

Remark 4.4. When M is a free R-module, its tensor powers $M^{\otimes k}$ are free R-modules. While $\Lambda^k(M)$ is a quotient module of $M^{\otimes k}$, it does not follow from this alone that $\Lambda^k(M)$ must be free: the quotient of a free module is not generally free (consider R/I where I is a proper nonzero ideal). Work was really needed to show exterior powers of free modules are free modules.

When M is free of rank d, $\Lambda^k(M) \neq 0$ when $k \leq d$ and $\Lambda^k(M) = 0$ when $k > d$. This is why we call $\Lambda^d(M)$ the *top exterior power*. It is free of rank 1; a basis of $\Lambda^d(M)$ is $e_1 \wedge \cdots \wedge e_d$ if $e_1, \ldots, e_d$ is a basis of M. Although $\Lambda^d(M)$ is isomorphic to R as an R-module, it is not naturally isomorphic: there is no canonical isomorphism between them.

When M has a d-element spanning set with d minimally chosen, and M is not free, it might happen that $\Lambda^d(M) = 0$. For example, consider a non-principal ideal $I := Rx + Ry$ in R with two generators. The module $\Lambda^2(I)$ is spanned as an R-module by $x \wedge y$.[2] Sometimes $\Lambda^2(I)$ is zero and sometimes it is nonzero.

Example 4.5. Let $R = \mathbf{Z}[\sqrt{-5}]$ and $I = (2, 1+\sqrt{-5})$. Set $\omega := 2 \wedge (1+\sqrt{-5}) \in \Lambda^2(I)$. We will show $2\omega = 0$ and $3\omega = 0$, so $\omega = 0$ (just subtract) and thus $\Lambda^2(I) = 0$:

$$2\omega = 2 \wedge 2(1+\sqrt{-5}) = (1+\sqrt{-5})(2 \wedge 2) = 0,$$

$$3\omega = 6 \wedge (1+\sqrt{-5}) = (1-\sqrt{-5})((1+\sqrt{-5}) \wedge (1+\sqrt{-5})) = 0.$$

[2]It is important to realize $x \wedge y$ here means an elementary wedge product in $\Lambda^2(I)$, *not* in $\Lambda^2(R)$; the latter exterior square is 0 all the time.

Example 4.6. Let $R = A[X, Y]$ be the polynomial ring in two variables over a nonzero commutative ring A. Let $I = (X, Y)$ in R. We will show $X \wedge Y$ in $\Lambda^2(I)$ is nonzero by writing down a linear map out of $\Lambda^2(I)$ whose value on $X \wedge Y$ is nonzero.

Define $B\colon I^2 \to A$ to be the determinant on degree one coefficients:

$$B(aX + bY + \cdots, cX + dY + \cdots) = ad - bc.$$

Regard the target module A as an R-module through scaling by the constant term of a polynomial: $f(X, Y)a = f(0, 0)a$. (That is, we basically treat A as $R/I = A[X, Y]/(X, Y)$.) Then B is R-bilinear, and it is alternating too. Since $B(X, Y) = 1$, B induces a linear map $L\colon \Lambda^2(I) \to A$ where $L(X \wedge Y) = B(X, Y) = 1$, so $\Lambda^2(I) \neq 0$.

Remark 4.7. An analogue of Example 4.5 in the real quadratic ring $\mathbf{Z}[\sqrt{5}]$ has a different result. For $J = (2, 1 + \sqrt{5})$, $2 \wedge (1 + \sqrt{5})$ in $\Lambda^2(J)$ is *not* 0, so $\Lambda^2(J) \neq 0$. Constructing an alternating bilinear map out of $J \times J$ that is not identically 0 can be done using the ideas in Example 4.6, and details are left to the reader (Hint: Find a basis for J as a $\mathbf{Z}$-module.)

The moral from these two examples is that for nonfree M, the highest k for which $\Lambda^k(M) \neq 0$ need not equal the size of a minimal spanning set for the module. When M is finitely generated, it only gives an upper bound.[3]

An important distinction to remember between tensor and exterior powers is that exterior powers are not recursively defined. Whereas $M^{\otimes(k+1)} \cong M \otimes_R M^{\otimes k}$, we can't say that $\Lambda^{k+1}(M)$ is a product of M and $\Lambda^k(M)$; later on (Section 10) we will introduce the exterior algebra, in which something like this does make sense. To appreciate the lack of a recursive definition of exterior powers, consider the following problem. When M has a d-element spanning set, $\Lambda^k(M) = 0$ for $k \geq d + 1$ by Theorem 4.1. Treating M as $\Lambda^1(M)$, we pose a generalization: if $\Lambda^i(M)$ for some $i > 1$ has a d-element spanning set, is $\Lambda^k(M) = 0$ for $k \geq d + i$? The next theorem settles the $d = 0$ case in the affirmative.

Theorem 4.8. *If $\Lambda^i(M) = 0$ for some $i \geq 1$, then $\Lambda^j(M) = 0$ for all $j \geq i$.*

Proof. It suffices to show $\Lambda^i(M) = 0 \Rightarrow \Lambda^{i+1}(M) = 0$, as then we are done by induction. To prove $\Lambda^{i+1}(M) = 0$ we show all $(i + 1)$-fold elementary wedge products

$$m_1 \wedge \cdots \wedge m_i \wedge m_{i+1} \tag{4.3}$$

are 0.

First we give a fake proof. Since $\Lambda^i(M) = 0$, $m_1 \wedge \cdots \wedge m_i = 0$, so (4.3) equals $0 \wedge m_{i+1} = 0$. What makes this absurd, at our present level of understanding, is that there is no sense (yet) in which the notation $\wedge$ is "associative," as the notation $\wedge$ is really just a placeholder to tell us where things go. We can't treat the piece $m_1 \wedge \cdots \wedge m_i$ in (4.3) as its own elementary wedge product having any kind of relation to (4.3). This is like calculus, where students are warned that the separate parts of $\mathrm{d}y/\mathrm{d}x$ do not have an independent meaning, although later they may learn otherwise, as we too will learn otherwise about $\wedge$ in Section 10.

Now we give a real proof, which in fact contains the germ of the idea in Section 10 to make $\wedge$ into a genuine operation and not just a placeholder. For each elementary wedge product (4.3), which belongs to $\Lambda^{i+1}(M)$, we will create a linear map $\Lambda^i(M) \to \Lambda^{i+1}(M)$ with (4.3) in its image. Then since $\Lambda^i(M) = 0$, so a linear map out of $\Lambda^i(M)$ has image 0, (4.3) is 0.

[3]Since $(\mathbf{Q}/\mathbf{Z}) \otimes_{\mathbf{Z}} (\mathbf{Q}/\mathbf{Z}) = 0$, $\Lambda^2(\mathbf{Q}/\mathbf{Z}) = 0$ where we regard $\mathbf{Q}/\mathbf{Z}$ as a $\mathbf{Z}$-module, so the vanishing of $\Lambda^k(M)$ for some k does not force M to be finitely generated.

Consider the function $M^i \to \Lambda^{i+1}(M)$ given by

$$(x_1, \ldots, x_i) \mapsto x_1 \wedge \cdots \wedge x_i \wedge m_{i+1}.$$

This is multilinear and alternating, so by the universal mapping property of exterior powers there is a linear map $\Lambda^i(M) \to \Lambda^{i+1}(M)$ where

$$x_1 \wedge \cdots \wedge x_i \mapsto x_1 \wedge \cdots \wedge x_i \wedge m_{i+1}.$$

The left side is 0 for all choices of $x_1, \ldots, x_i$ in M, so the right side is 0 for all such choices too. In particular, (4.3) is 0.

Here is a different proof. To say $\Lambda^i(M) = 0$ means any alternating multilinear function out of M^i is identically 0. If $\varphi \colon M^{i+1} \to N$ is an alternating multilinear function, and $(m_1, \ldots, m_i, m_{i+1}) \in M^{i+1}$, consider the function $\varphi(x_1, \ldots, x_i, m_{i+1})$ in $x_1, \ldots, x_i$. It is alternating multilinear in i variables from M. Therefore $\varphi(x_1, \ldots, x_i, m_{i+1}) = 0$ for all $x_1, \ldots, x_i$ in M, so $\varphi(m_1, \ldots, m_i, m_{i+1}) = 0$. $\square$

Returning to the general question, where $\Lambda^i(M)$ has a d-element spanning set, asking if $\Lambda^k(M) = 0$ for $k \geq d + i$ is the same as asking if $\Lambda^{d+i}(M) = 0$ by Theorem 4.8. The answer is "yes" although more technique is needed for that than we will develop here.

5. Exterior powers of linear maps

Having constructed exterior powers of modules, we extend the construction to linear maps between modules. First recall any linear map $\varphi \colon M \to N$ between two R-modules induces a linear map $\varphi^{\otimes k} \colon M^{\otimes k} \to N^{\otimes k}$ on the kth tensor powers, for any positive integer k, which has the effect

$$\varphi^{\otimes k}(m_1 \otimes \cdots \otimes m_k) = \varphi(m_1) \otimes \cdots \otimes \varphi(m_k)$$

on elementary tensors.

Theorem 5.1. *Let $\varphi \colon M \to N$ be a linear map of R-modules. Then for each $k \geq 1$ there is a unique linear map $\wedge^k(\varphi) \colon \Lambda^k(M) \to \Lambda^k(N)$ with the effect*

$$m_1 \wedge \cdots \wedge m_k \mapsto \varphi(m_1) \wedge \cdots \wedge \varphi(m_k)$$

on all elementary wedge products. For a second linear map $\psi \colon N \to P$, $\wedge^k(\psi \circ \varphi) = \wedge^k(\psi) \circ \wedge^k(\varphi)$. Moreover, $\wedge^k(\mathrm{id}_M) = \mathrm{id}_{\Lambda^k(M)}$.

Proof. This is trivial when $k = 1$, so let $k \geq 2$. There is at most one such linear map $\Lambda^k(M) \to \Lambda^k(N)$ since the elementary wedge products span $\Lambda^k(M)$. To show there is such a linear map, start by backing up and defining a function $f \colon M^k \to \Lambda^k(N)$ by

$$f(m_1, \ldots, m_k) = \varphi(m_1) \wedge \cdots \wedge \varphi(m_k).$$

This is a multilinear map that is alternating, so by the universal mapping property of the kth exterior power there is a linear map $\Lambda^k(M) \to \Lambda^k(N)$ with the effect

$$m_1 \wedge \cdots \wedge m_k \mapsto f(m_1, \ldots, m_k) = \varphi(m_1) \wedge \cdots \wedge \varphi(m_k).$$

$$\begin{tikzcd} M^k \arrow[r, "f"] \arrow[d, "\wedge"'] & \Lambda^k(N) \\ \Lambda^k(M) \arrow[ur, dotted] & \end{tikzcd}$$

This proves the existence of the linear map $\wedge^k(\varphi)$ we are seeking.

To show $\wedge^k(\psi \circ \varphi) = \wedge^k(\psi) \circ \wedge^k(\varphi)$, it suffices since both sides are linear to check that both sides have the same value on each elementary wedge product in $\Lambda^k(M)$. At any elementary wedge product $m_1 \wedge \cdots \wedge m_k$, the left side and right side have the common value $\psi(\varphi(m_1)) \wedge \cdots \wedge \psi(\varphi(m_k))$. That $\wedge^k(\mathrm{id}_M) = \mathrm{id}_{\Lambda^k(M)}$ is easy: $\wedge^k(\mathrm{id}_M)$ is linear and fixes every $m_1 \wedge \cdots \wedge m_k$ and these span $\Lambda^k(M)$, so $\wedge^k(\mathrm{id}_M)$ fixes everything. $\square$

Theorem 5.1 is also true for $k = 0$ by setting $\wedge^0(\varphi) = \mathrm{id}_R$. Recall $\Lambda^0(M) = R$

That the passage from φ to $\wedge^k(\varphi)$ respects composition and sends the identity map on a module to the identity map on its kth exterior power is called *functoriality* of the kth exterior power.

Armed with bases for exterior powers of finite free modules, we can write down matrices for exterior powers of linear maps between them. When M and N are finite free of respective ranks m and n, a choice of bases of M and N turns any linear map $\varphi\colon M \to N$ into an $n \times m$ matrix. Using the corresponding bases on $\Lambda^k(M)$ and $\Lambda^k(N)$, we can write $\wedge^k(\varphi)$ as an $\binom{n}{k} \times \binom{m}{k}$ matrix. If we want to look at an example, we need to keep m and n small or we will face very large matrices.

Example 5.2. If M and N are free of rank 5 and $\varphi\colon M \to N$ is linear, then $\wedge^2(\varphi)$ is represented by a 10×10 matrix since $\binom{5}{2} = 10$.

Example 5.3. Let $L\colon \mathbf{R}^3 \to \mathbf{R}^3$ be the linear map given by the matrix

$$\begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 3 & 2 \end{pmatrix}. \tag{5.1}$$

We will compute the matrix for $\wedge^2(L)\colon \Lambda^2(\mathbf{R}^3) \to \Lambda^2(\mathbf{R}^3)$ with respect to the basis $e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3$, where the e_i's are the standard basis of $\mathbf{R}^3$. Going in order,

$$\begin{aligned} \wedge^2(L)(e_1 \wedge e_2) &= L(e_1) \wedge L(e_2) \\ &= e_2 \wedge (2e_1 + e_2 + 3e_3) \\ &= -2(e_1 \wedge e_2) + 3(e_2 \wedge e_3), \end{aligned}$$

$$\begin{aligned} \wedge^2(L)(e_1 \wedge e_3) &= L(e_1) \wedge L(e_3) \\ &= e_2 \wedge (e_2 + 2e_3) \\ &= 2(e_2 \wedge e_3), \end{aligned}$$

and

$$\begin{aligned} \wedge^2(L)(e_2 \wedge e_3) &= L(e_2) \wedge L(e_3) \\ &= (2e_1 + e_2 + 3e_3) \wedge (e_2 + 2e_3) \\ &= 2(e_1 \wedge e_2) + 4(e_1 \wedge e_3) - e_2 \wedge e_3. \end{aligned}$$

Therefore the matrix for $\wedge^2(L)$ relative to this ordered basis is

$$\begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 4 \\ 3 & 2 & -1 \end{pmatrix}.$$

Theorem 5.4. *Let $\varphi\colon M \to N$ be linear. If φ is an R-module isomorphism then $\wedge^k(\varphi)$ is an R-module isomorphism for every k. If φ is surjective then every $\wedge^k(\varphi)$ is surjective.*

Proof. It is clear for $k = 0$ and $k = 1$. Let $k \geq 2$. Suppose φ is an isomorphism of R-modules, with inverse $\psi\colon N \to M$. Then $\varphi \circ \psi = \mathrm{id}_N$ and $\psi \circ \varphi = \mathrm{id}_M$, so by Theorem 5.1 we have $\wedge^k(\varphi) \circ \wedge^k(\psi) = \wedge^k(\mathrm{id}_N) = \mathrm{id}_{\Lambda^k(N)}$ and similarly $\wedge^k(\psi) \circ \wedge^k(\varphi) = \mathrm{id}_{\Lambda^k(M)}$.

If φ is surjective, then $\wedge^k(\varphi)$ is surjective because $\Lambda^k(N)$ is spanned by the elementary wedge products $n_1 \wedge \cdots \wedge n_k$ for all $n_i \in N$ and these are in the image of $\wedge^k(\varphi)$ explicitly: writing $n_i = \varphi(m_i)$, the elementary wedge product of the n_i's in $\Lambda^k(N)$ is $\varphi(m_1 \wedge \cdots \wedge m_k)$. Since the image of the linear map $\wedge^k(\varphi)$ is a submodule of $\Lambda^k(N)$ that contains a spanning set for $\Lambda^k(N)$, the image is all of $\Lambda^k(N)$. $\square$

As with tensor products of linear maps, it is *false* that $\wedge^k(\varphi)$ has to be injective if φ is injective.

Example 5.5. Let $R = A[X, Y]$ with A a nonzero commutative ring and let $I = (X, Y)$. In our discussion of tensor products, it was seen that the inclusion map $i\colon I \to R$ is injective while its induced R-linear map $i^{\otimes 2}\colon I^{\otimes 2} \to R^{\otimes 2} \cong R$ is not injective. Therefore it should come as no surprise that the map $\wedge^2(i)\colon \Lambda^2(I) \to \Lambda^2(R)$ also is not injective. Indeed, $\Lambda^2(R) = 0$ and we saw in Example 4.6 that $\Lambda^2(I) \neq 0$ because there is an R-linear map $L\colon \Lambda^2(I) \to A$ where $L(X \wedge Y) = 1$.

(In $\Lambda^2(I)$ we have $X \wedge Y \neq 0$ while in $\Lambda^2(R)$ we have $X \wedge Y = XY(1 \wedge 1) = 0$. There is nothing inconsistent about this, even though $I \subset R$, because the natural map $\wedge^2(i)\colon \Lambda^2(I) \to \Lambda^2(R)$ is not injective. The $X \wedge Y$'s in $\Lambda^2(I)$ and $\Lambda^2(R)$ lie in different modules, and although $\wedge^2(i)$ sends $X \wedge Y$ in the first module to $X \wedge Y$ in the second, linear maps can send a nonzero element to 0 and that is what is happening.)

We now show the linear map $L\colon \Lambda^2(I) \to A$ is actually an isomorphism of R-modules. In $\Lambda^2(I) = \{r(X \wedge Y) : r \in R\}$, $X \wedge Y$ is killed by multiplication by X and Y since $X(X \wedge Y) = X \wedge XY = Y(X \wedge X) = 0$ and likewise for $Y(X \wedge Y)$. So $f(X, Y)(X \wedge Y) = f(0, 0)(X \wedge Y)$ in $\Lambda^2(I)$, which means every element of $\Lambda^2(I)$ has the form $a(X \wedge Y)$ for some $a \in A$. Thus the function $L'\colon A \to \Lambda^2(I)$ given by $L'(a) = a(X \wedge Y)$ is R-linear and is an inverse to L.

The isomorphism $\Lambda^2(I) \cong A$ generalizes to the polynomial ring $R = A[X_1, \ldots, X_n]$ for any $n \geq 2$: the ideal $I = (X_1, \ldots, X_n)$ in R has $\Lambda^n(I) \cong A \cong R/I$, so the inclusion $i\colon I \to R$ is injective but $\wedge^n(i)\colon \Lambda^n(I) \to \Lambda^n(R) = 0$ is not injective.

Although exterior powers don't preserve injectivity of linear maps in general, there are some cases when they do. This is the topic of the rest of this section. A number of ideas here and later are taken from [2].

Theorem 5.6. *Suppose $\varphi\colon M \to N$ is injective and the image $\varphi(M) \subset N$ is a direct summand: $N = \varphi(M) \oplus P$ for some submodule P of N. Then $\wedge^k(\varphi)$ is injective for all $k \geq 0$ and $\Lambda^k(M)$ is isomorphic to a direct summand of $\Lambda^k(N)$.*

Proof. The result is trivial for $k = 0$ and $k = 1$. Suppose $k \geq 2$.

We will use the splitting criteria for short exact sequences of modules. Since $N = \varphi(M) \oplus P$, we have a linear map $\psi\colon N \twoheadrightarrow M$ that undoes the effect of φ: let $\psi(\varphi(m) + p) = m$. Then $\psi(\varphi(m)) = m$ for all $m \in M$, so

$$\psi \circ \varphi = \mathrm{id}_M\,. \tag{5.2}$$

(The composite in the other direction, $\varphi \circ \psi$, is definitely not id_N unless $P = 0$, but this does not matter.) Applying $\wedge^k$ to (5.2) gives us linear maps $\wedge^k(\varphi)\colon \Lambda^k(M) \to \Lambda^k(N)$ and $\wedge^k(\psi)\colon \Lambda^k(N) \to \Lambda^k(M)$ with

$$\wedge^k(\psi) \circ \wedge^k(\varphi) = \wedge^k(\psi \circ \varphi) = \wedge^k(\mathrm{id}_M) = \mathrm{id}_{\Lambda^k(M)}$$

by functoriality (Theorem 5.1). In particular, if $\wedge^k(\varphi)(\omega) = 0$ for some $\omega \in \Lambda^k(M)$, then applying $\wedge^k(\psi)$ to both sides gives us $\omega = \wedge^k(\psi)(0) = 0$, so $\wedge^k(\varphi)$ has kernel 0 and thus is injective.

For the short exact sequence $0 \longrightarrow \Lambda^k(M) \xrightarrow{\wedge^k(\varphi)} \Lambda^k(N) \longrightarrow \Lambda^k(N)/\Lambda^k(M) \longrightarrow 0$, the fact that $\wedge^k(\psi)$ is a left inverse to $\wedge^k(\varphi)$ implies by the splitting criteria for short exact sequences that $\Lambda^k(N) \cong \Lambda^k(M) \oplus (\Lambda^k(N)/\Lambda^k(M))$, so $\Lambda^k(M)$ is isomorphic to a direct summand of $\Lambda^k(N)$. $\square$

Example 5.7. For any R-modules M and M', the inclusion $i\colon M \to M \oplus M'$ where $i(m) = (m, 0)$ has image $M \oplus 0$, which is a direct summand of $M \oplus M'$, so the induced linear map $\Lambda^k(M) \to \Lambda^k(M \oplus M')$ sending $m_1 \wedge \cdots \wedge m_k$ to $(m_1, 0) \wedge \cdots \wedge (m_k, 0)$ is one-to-one.

Remark 5.8. It is instructive to check that the hypothesis of Theorem 5.6 does not apply to the inclusion $i\colon (X, Y) \to A[X, Y]$ in Example 5.5 (which must be so because $\wedge^2(i)$ is not injective). We can write $A[X, Y] = A \oplus (X, Y)$, so (X, Y) is a direct summand of $A[X, Y]$ as abelian groups, or even as A-modules, but this is not a direct sum of $A[X, Y]$-modules: A is not an ideal in $A[X, Y]$.

Corollary 5.9. *Let K be a field and V and W be K-vector spaces. If the K-linear map $\varphi\colon V \to W$ is injective then $\wedge^k(\varphi)\colon \Lambda^k(V) \to \Lambda^k(W)$ is injective for all $k \geq 0$.*

Proof. The subspace $\varphi(V) \subset W$ is a direct summand of W: pick a basis of $\varphi(V)$ over K, extend it to a basis of the whole space W, and let P be the span of the new part of this full basis: $W = \varphi(V) \oplus P$. Thus the hypothesis of Theorem 5.6 applies to this situation.[4] $\square$

When working with linear maps of vector spaces (not necessarily finite-dimensional), we have shown

$$\begin{aligned}
\varphi\colon V \to W \text{ injective} &\implies \wedge^k(\varphi) \text{ injective for all } k \text{ (Corollary 5.9)},\\
\varphi\colon V \to W \text{ surjective} &\implies \wedge^k(\varphi) \text{ surjective for all } k \text{ (Theorem 5.4)},\\
\varphi\colon V \to W \text{ an isomorphism} &\implies \wedge^k(\varphi) \text{ an isomorphism for all } k.
\end{aligned}$$

Replacing vector spaces with modules, the second and third properties are true but the first one may fail (Example 5.5 with $R = A[X, Y]$ and $k = 2$). The first property does remain true for free modules, however.

Theorem 5.10. *Suppose M and N are free R-modules. If a linear map $\varphi\colon M \to N$ is injective then $\wedge^k(\varphi)\colon \Lambda^k(M) \to \Lambda^k(N)$ is injective for all k.*

Notice the free hypothesis! We can't use Corollary 5.9 here (if R is not a field), as the image of φ need not be a direct summand of N. That is, a submodule of a free module is often not a direct summand (unless R is a field). We are not assuming M and N have finite bases.

[4]Picking a basis of $\varphi(V)$ and extending a basis of it to all of W, for possibly infinite-dimensional W, needs the axiom of choice. The proof of Corollary 5.9 can be modified to avoid using the axiom of choice. See https://mathoverflow.net/questions/332043/exterior-powers-and-choice.

Proof. The diagram

$$\begin{tikzcd} \Lambda^k(M) \arrow[r, "\alpha_{k,M}"] \arrow[d, "\wedge^k(\varphi)"'] & M^{\otimes k} \arrow[d, "\varphi^{\otimes k}"] \\ \Lambda^k(N) \arrow[r, "\alpha_{k,N}"] & N^{\otimes k} \end{tikzcd}$$

commutes, where the top and bottom maps come from Theorem 4.2. The explicit effect in this diagram on an elementary wedge product in $\Lambda^k(M)$ is given by

$$\begin{tikzcd} m_1 \wedge \cdots \wedge m_k \arrow[r, mapsto, "\alpha_{k,M}"] \arrow[d, mapsto, "\wedge^k(\varphi)"'] & \Sigma_{\sigma \in S_k} (\operatorname{sign} \sigma) m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(k)} \arrow[d, mapsto, "\varphi^{\otimes k}"] \\ \varphi(m_1) \wedge \cdots \wedge \varphi(m_k) \arrow[r, mapsto, "\alpha_{k,N}"] & \Sigma_{\sigma \in S_k} (\operatorname{sign} \sigma) \varphi(m_{\sigma(1)}) \otimes \cdots \otimes \varphi(m_{\sigma(k)}) \end{tikzcd}$$

Since $\alpha_{k,M}$ and $\alpha_{k,N}$ are injective by Theorem 4.2, and $\varphi^{\otimes k}$ is injective from our development of the tensor product (any tensor power of an injective linear map of *flat* modules is injective, and free modules are flat), $\wedge^k(\varphi)$ has to be injective from commutativity of the diagram. $\square$

Here is a nice application of Theorem 5.10 and the nonvanishing of $\Lambda^k(R^n)$ for $k \leq n$ (but not $k > n$).

Corollary 5.11. *If $\varphi \colon R^m \to R^n$ is a linear map, where m and n are positive integers, then surjectivity of φ implies $m \geq n$ and injectivity of φ implies $m \leq n$.*

Proof. First suppose φ is onto. Taking nth exterior powers, $\wedge^n(\varphi) : \Lambda^n(R^m) \to \Lambda^n(R^n)$ is onto by Theorem 5.4. Since $\Lambda^n(R^n) \neq 0$ we get $\Lambda^n(R^m) \neq 0$, so $n \leq m$.

Now suppose φ is one-to-one. Taking mth exterior powers, $\wedge^m(\varphi) \colon \Lambda^m(R^m) \to \Lambda^m(R^n)$ is one-to-one by Theorem 5.10, so the nonvanishing of $\Lambda^m(R^m)$ implies $\Lambda^m(R^n) \neq 0$, so $m \leq n$. $\square$

The proof of Corollary 5.11 is short, but if we unravel it we see that the injectivity part of Corollary 5.11 is a deeper result than the surjectivity, because exterior powers of linear maps don't preserve injectivity in general (Example 5.5). There will be another interesting application of Theorem 5.10 in Section 7 (Theorem 7.8).

Corollary 5.12. *If M is a nonzero free module and $\{m_1, \ldots, m_s\}$ is a finite linearly independent subset then for any $k \leq s$ the $\binom{s}{k}$ elementary wedge products*

$$m_{i_1} \wedge \cdots \wedge m_{i_k} \text{ where } 1 \leq i_1 < \cdots < i_k \leq s \tag{5.3}$$

are linearly independent in $\Lambda^k(M)$.

Proof. We have an embedding $R^s \hookrightarrow M$ by $\sum_{i=1}^s r_i e_i \mapsto \sum_{i=1}^s r_i m_i$. Since R^s and M are free, the kth exterior power of this linear map is an embedding $\Lambda^k(R^s) \hookrightarrow \Lambda^k(M)$ that sends the basis

$$e_{i_1} \wedge \cdots \wedge e_{i_k}$$

of $\Lambda^k(R^s)$, where $1 \leq i_1 < \cdots < i_k \leq s$, to the elementary wedge products in (5.3), so they are linearly independent in $\Lambda^k(M)$. $\square$

Remark 5.13. In a vector space, any linearly independent subset extends to a basis. The corresponding result in R^n is generally false. In fact $\mathbf{Z}^2$ already provides counterexamples: the vector $(2, 2)$ is linearly independent by itself but can't belong to a basis because a basis vector in $\mathbf{Z}^2$ must have relatively prime coordinates. Since any linearly independent subset of R^n has at most n terms in it, by Corollary 5.11, it is natural to ask if every maximal linearly independent subset of R^n has n vectors in it (which need not be a basis). This is true in $\mathbf{Z}^2$, *e.g.*, $(2,2)$ is part of the linearly independent subset $\{(2,2),(1,0)\}$.

However, it is *not* true in general that every maximal linearly independent subset of R^n has n vectors in it. There are rings R such that R^2 contains a vector v such that $\{v\}$ is linearly independent (meaning it's torsion-free) but there is no linearly independent subset $\{v,w\}$ in R^2. An example, due to David Speyer, is the following: let R be the ring of functions $\mathbf{C}^2 - \{(0,0)\} \to \mathbf{C}$ that coincide with a polynomial function at all but finitely many points. (The finitely many exceptional points can vary.) Letting z and w be the coordinate functions on $\mathbf{C}^2$, in R^2 the vector (z,w) is linearly independent and is not part of any larger linearly independent subset.

6. Determinants

Now we put exterior powers to work in the development of the determinant, whose properties have up until now played no role except for a 2×2 determinant in Example 5.5.

Let M be a free R-module of rank $d \geq 1$. The top exterior power $\Lambda^d(M)$ is a free R-module of rank 1, so any linear map $\Lambda^d(M) \to \Lambda^d(M)$ is scaling by an element of R. For a linear map $\varphi\colon M \to M$, the induced linear map $\wedge^d(\varphi)\colon \Lambda^d(M) \to \Lambda^d(M)$ is scaling by what element of R?

Theorem 6.1. *If M is a free R-module of rank $d \geq 1$ and $\varphi\colon M \to M$ is a linear map, its top exterior power $\wedge^d(\varphi)\colon \Lambda^d(M) \to \Lambda^d(M)$ is multiplication by $\det\varphi \in R$.*

Proof. We want to show $\wedge^d(\varphi)(\omega) = (\det\varphi)\omega$ for all $\omega \in \Lambda^d(M)$. It suffices to check this when ω is a basis of $\Lambda^d(M)$. Let $e_1,\dots,e_d$ be a basis for M, so $e_1\wedge\cdots\wedge e_d$ is a basis for $\Lambda^d(M)$. We will show

$$\wedge^d(\varphi)(e_1\wedge\cdots\wedge e_d) = (\det\varphi)(e_1\wedge\cdots\wedge e_d).$$

By definition,

$$\wedge^d(\varphi)(e_1\wedge\cdots\wedge e_d) = \varphi(e_1)\wedge\cdots\wedge\varphi(e_d).$$

Let $\varphi(e_j) = \sum_{i=1}^d a_{ij}e_i$. Then (a_{ij}) is the matrix representation for φ in the ordered basis $e_1,\dots,e_d$ and

$$\begin{aligned}\wedge^d(\varphi)(e_1\wedge\cdots\wedge e_d) &= \sum_{i=1}^d a_{i1}e_i\wedge\cdots\wedge\sum_{i=1}^d a_{id}e_i\\ &= \sum_{i_1=1}^d a_{i_11}e_{i_1}\wedge\cdots\wedge\sum_{i_d=1}^d a_{i_dd}e_{i_d},\end{aligned}$$

where we introduce different labels for the summation indices because we are about to combine terms using multilinearity:

$$\wedge^d(\varphi)(e_1\wedge\cdots\wedge e_d) = \sum_{i_1,\dots,i_d=1}^d a_{i_11}\cdots a_{i_dd}e_{i_1}\wedge\cdots\wedge e_{i_d}.$$

In this sum, terms with equal indices can be dropped (the wedge product vanishes) so all we are left with is a sum over d-tuples of distinct indices. Since d distinct integers from 1 to d must be $1, 2, \dots, d$ in some rearrangement, we can write $i_1 = \sigma(1), \dots, i_d = \sigma(d)$ as σ runs over S_d:

$$\wedge^d(\varphi)(e_1 \wedge \cdots \wedge e_d) = \sum_{\sigma \in S_d}^{d} a_{\sigma(1)1} \cdots a_{\sigma(d)d}(e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(d)}).$$

By (3.2), this becomes

$$\wedge^d(\varphi)(e_1 \wedge \cdots \wedge e_d) = \sum_{\sigma \in S_d}^{d} (\operatorname{sign}\sigma) a_{\sigma(1)1} \cdots a_{\sigma(d)d}(e_1 \wedge \cdots \wedge e_d).$$

Thus $\wedge^d(\varphi)$ acts on the 1-element basis $e_1 \wedge \cdots \wedge e_d$ of $\Lambda^d(M)$ as multiplication by the number we recognize as $\det((a_{ij})^\top) = \det(a_{ij}) = \det(\varphi)$, so it acts on every element of $\Lambda^d(M)$ as scaling by $\det(\varphi)$. $\square$

Since we did not use determinants before, we could *define* the determinant of a linear operator φ on a (nonzero) finite free module M to be the scalar by which φ acts on the top exterior power of M. Then the proof of Theorem 6.1 shows $\det \varphi$ can be computed from any matrix representation of φ by the usual formula, and it shows this formula is independent of the choice of matrix representation for φ since our construction of exterior powers was coordinate-free. Since $\wedge^k(\mathrm{id}_M) = \mathrm{id}_{\Lambda^k(M)}$, the determinant of the identity map is 1. Here is a slick proof that the determinant is multiplicative:

Corollary 6.2. *If M is a nonzero finite free R-module and φ and ψ are linear maps $M \to M$, then $\det(\psi \circ \varphi) = \det(\psi)\det(\varphi)$.*

Proof. Let M have rank $d \geq 1$. By Theorem 5.1, $\wedge^d(\psi \circ \varphi) = \wedge^d(\psi) \circ \wedge^d(\varphi)$. Both sides are linear maps $\Lambda^d(M) \to \Lambda^d(M)$ and $\Lambda^d(M)$ is free of rank 1. The left side is multiplication by $\det(\psi \circ \varphi)$. The right side is multiplication by $\det(\varphi)$ followed by multiplication by $\det(\psi)$, which is multiplication by $\det(\psi)\det(\varphi)$. Thus, by checking both sides on a one-element basis of $\Lambda^d(M)$, we obtain $\det(\psi \circ \varphi) = \det(\psi)\det(\varphi)$. $\square$

Continuing a purely logical development (not assuming prior knowledge of determinants, that is), at this point we could introduce the characteristic polynomial and prove the Cayley-Hamilton theorem. One of the corollaries of the Cayley-Hamilton theorem is that $\mathrm{GL}_d(R) = \{A \in \mathrm{M}_d(R) : \det A \in R^\times\}$. That is used in the next result, which characterizes bases in R^d using $\Lambda^d(R)$.

Corollary 6.3. *Let M be a free R-module of rank $d \geq 1$. For $x_1, \dots, x_d \in M$, $\{x_1, \dots, x_d\}$ is a basis of M if and only if $x_1 \wedge \cdots \wedge x_d$ is a basis of $\Lambda^d(M)$.*

Proof. We know by Theorem 4.2 that if $\{x_1, \dots, x_d\}$ is a basis of M then $x_1 \wedge \cdots \wedge x_d$ is a basis of $\Lambda^d(M)$. We now want to go the other way: if $x_1 \wedge \cdots \wedge x_d$ is a basis of $\Lambda^d(M)$ we show $\{x_1, \dots, x_d\}$ is a basis of M.

Since M is free of rank d it has some basis, say $\{e_1, \dots, e_d\}$. Write the x_j's in terms of this basis: $x_j = \sum_{i=1}^n a_{ij} e_j$, where $a_{ij} \in R$. That means the linear map $A \colon M \to M$ given by $A(e_j) = x_j$ for all j has matrix representation (a_{ij}) in the basis $\{e_1, \dots, e_d\}$. Therefore

$$x_1 \wedge \cdots \wedge x_d = Ae_1 \wedge \cdots \wedge Ae_d = \wedge^d(A)(e_1 \wedge \cdots \wedge e_d) = (\det A)(e_1 \wedge \cdots \wedge e_d).$$

Since $x_1 \wedge \cdots \wedge x_d$ and $e_1 \wedge \cdots \wedge e_d$ are both bases of $\Lambda^d(M)$ (the first by hypothesis and the second by Theorem 4.2), the scalar by which they differ must be a unit: $\det A \in R^\times$.

Therefore $A \in \mathrm{GL}_d(R)$, so the x_j's must be a basis of R^d because $A(e_j) = x_j$ and the e_j's are a basis. $\square$

While the top exterior power of a linear operator on a finite free module is multiplication by its determinant, the lower-order exterior powers of a linear map between finite free modules have matrix representations whose entries are determinants. Let $\varphi\colon M \to N$ be linear, with M and N finite free of positive ranks m and n, respectively. Take $k \leq m$ and $k \leq n$ since otherwise $\Lambda^k(M)$ or $\Lambda^k(N)$ is 0. Pick bases $e_1, \dots, e_m$ for M and $f_1, \dots, f_n$ for N. The corresponding basis of $\Lambda^k(M)$ is all $e_{j_1} \wedge \cdots \wedge e_{j_k}$ where

$$1 \leq j_1 < \cdots < j_k \leq m,$$

Similarly, the corresponding basis for $\Lambda^k(N)$ is all $f_{i_1} \wedge \cdots \wedge f_{i_k}$ where $1 \leq i_1 < \cdots < i_k \leq n$. The matrix for $\wedge^k(\varphi)$ relative to these bases of $\Lambda^k(M)$ and $\Lambda^k(N)$ is therefore naturally indexed by pairs of increasing k-tuples.

Theorem 6.4. *With notation as above, let $[\varphi]$ denotes the $n \times m$ matrix for φ relative to the choice of bases for M and N. Relative to the corresponding bases on $\Lambda^k(M)$ and $\Lambda^k(N)$, the matrix entry for $\wedge^k(\varphi)$ in row position $(i_1, \dots, i_k)$ and column position $(j_1, \dots, j_k)$ is the determinant of the $k \times k$ matrix built from rows $i_1, \dots, i_k$ and columns $j_1, \dots, j_k$ of $[\varphi]$.*

Proof. The matrix entry in question is the coefficient of $f_{i_1} \wedge \cdots \wedge f_{i_k}$ in the expansion of $\wedge^k(\varphi)(e_{j_1} \wedge \cdots \wedge e_{j_k}) = \varphi(e_{j_1}) \wedge \cdots \wedge \varphi(e_{j_k})$. Details are left to the reader. $\square$

In short, this says the kth exterior power of a linear map φ has matrix entries that are $k \times k$ determinants of submatrices of the matrix of φ.

Example 6.5. In Example 5.3 we computed the second exterior power of the linear map $\mathbf{R}^3 \to \mathbf{R}^3$ given by the 3×3 matrix L in (5.1). The result is a matrix of size $\binom{3}{2} \times \binom{3}{2} = 3 \times 3$ matrix whose rows and columns are associated to basis pairs $e_i \wedge e_{i'}$ and $e_j \wedge e_{j'}$, where the ordering of the basis was $e_1 \wedge e_2, e_1 \wedge e_3$, and $e_2 \wedge e_3$. For instance, the upper right entry in the matrix for $\wedge^2(L)$ is in its first row and third column, so it is the $e_1 \wedge e_2$-coefficient of $\wedge^2(L)(e_2 \wedge e_3)$. This matrix entry has row position $(1, 2)$ (index for the first basis vector) and column position $(2, 3)$ (index for the third basis vector). The 2×2 submatrix of L using rows 1 and 2 and columns 2 and 3 is $\left(\begin{smallmatrix} 2 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, whose determinant is 2, which matches the upper right entry in the matrix at the end of Example 5.3.

7. Exterior powers and linear independence

This section discusses the connection between elementary wedge products and linear independence in a free module. We will start off with vector spaces, which are easier to handle and provide a geometric perspective on exterior powers that sounds different from all this business about alternating multilinear maps.

If we are given d vectors in $\mathbf{R}^d$ for $d \geq 1$, then there are two ways to determine if they are linearly independent by using an $d \times d$ matrix with those vectors as its columns:

(1) Row reduce the matrix and see if you get the identity matrix.
(2) Compute the determinant of the matrix and see if you get a nonzero value.

These methods can be generalized to decide if k vectors in $\mathbf{R}^d$ are linearly dependent when $1 \leq k < d$ by using a $d \times k$ matrix having the k vectors as its columns:

(1) Row reduce the matrix and see if a $k \times k$ submatrix is the identity.
(2) Compute the determinants of all $k \times k$ submatrices and see if any of them is not 0.

While the first way (row reduction) involves a set of steps that always keeps you going in the right direction, the second way involves one determinant computation after the other, and that will take longer to complete (particularly if all the determinants turn out to be 0!). Using exterior powers, we can carry out all these determinant computations at once. That is the algorithmic content of the next theorem for finite-dimensional vector spaces, as we'll see after its proof.

Theorem 7.1. *Let V be a nonzero real vector space. For $k \geq 1$, the vectors $v_1, \ldots, v_k$ in V are linearly independent if and only if $v_1 \wedge \cdots \wedge v_k \neq 0$ in $\Lambda^k(V)$.*

Proof. We can assume $k \leq \dim_{\mathbf{R}}(V)$, since $k > \dim_{\mathbf{R}}(V)$ then $\{v_1, \ldots, v_k\}$ can't be linearly dependent and $\Lambda^k(V) = 0$.

The case $k = 1$ is trivial, so take $k \geq 2$. First assume $\{v_1, \ldots, v_k\}$ is a linearly independent set. This set extends to a basis of V (we are working over a field!), so $v_1 \wedge \cdots \wedge v_k$ is part of a basis of $\Lambda^k(V)$ by Theorem 4.2. In particular, $v_1 \wedge \cdots \wedge v_k \neq 0$ in $\Lambda^k(V)$.

Now suppose $\{v_1, \ldots, v_k\}$ is linearly dependent, so one v_i is a linear combination of the others. Whether or not $v_1 \wedge \cdots \wedge v_k$ is nonzero in $\Lambda^k(V)$ is independent of the order of the factors, since permuting the terms only changes the elementary wedge product by a sign, so we may suppose v_k is a linear combination of the rest:

$$v_k = c_1 v_1 + \cdots + c_{k-1} v_{k-1}.$$

Then in $\Lambda^k(V)$,

$$v_1 \wedge \cdots \wedge v_{k-1} \wedge v_k = v_1 \wedge \cdots \wedge v_{k-1} \wedge \left(\sum_{i=1}^{k-1} c_i v_i \right).$$

Expanding the right side gives a sum of $k - 1$ wedge products, each containing a repeated vector, so every term vanishes. □

Taking $k = 2$, v and w in V are linearly independent if and only if $v \wedge w \neq 0$ in $\Lambda^2(V)$. In Appendix A, we'll see that the linear maps $A \colon V \to V$ preserving wedge products, by which we mean $Av \wedge Aw = v \wedge w$ for all v and w in V, are extremely limited when $\dim V > 1$. So the linear maps preserving wedge products are not nearly as interesting as the linear maps preserving bilinear forms (which lead to orthogonal or symplectic groups).

There is nothing important in Theorem 7.1 about working over $\mathbf{R}$: its proof works for vector spaces over all fields. And in Theorem 7.1, V can be an infinite-dimensional vector space since we never required finite-dimensionality in the proof. At one point we invoked Theorem 4.2, which was proved for all free modules, not just free modules with a finite basis.

Remark 7.2. Theorem 7.1 provides an intuition about what exterior powers mean: the kth exterior power of a vector space is what we get if we want a vector space built from k-tuples of vectors where the linearly dependent sets of k vectors are set equal to 0. The fact that linear dependence of k vectors in $\mathbf{R}^d$ can be detected by a vanishing property of a product (the elementary wedge product of the k vectors) is a counterpart to the vanishing of the dot product of two vectors in $\mathbf{R}^d$ for detecting their orthogonality.

When the vector space in Theorem 7.1 is finite dimensional, we can interpret the conclusion of the theorem in terms of determinants of $k \times k$ submatrices. Let $d = \dim_{\mathbf{R}}(V)$, so $k \le d$. Identity V with $\mathbf{R}^d$ by picking a basis. Let $A \in \mathrm{M}_{d\times k}(\mathbf{R})$ be the matrix whose columns are $v_1, \ldots, v_k$. Writing $v_j = \sum_{i=1}^{d} a_{ij} e_i$ using the standard basis $e_1, \ldots, e_d$ of $\mathbf{R}^d$,

$$v_1 \wedge \cdots \wedge v_k = \sum_{1 \le i_1 < \cdots < i_k \le d} \begin{vmatrix} a_{1i_1} & \cdots & a_{1i_k} \\ \vdots & \ddots & \vdots \\ a_{ki_1} & \cdots & a_{ki_k} \end{vmatrix} e_{i_1} \wedge \cdots \wedge e_{i_k}. \tag{7.1}$$

Since the k-fold elementary wedge products of the standard basis in this sum form a basis of $\Lambda^k(\mathbf{R}^d)$, $v_1 \wedge \cdots \wedge v_k$ is nonzero precisely when at least one of the coefficients in the sum is nonzero. Therefore $v_1, \ldots, v_k$ is linearly independent in $\mathbf{R}^d$ if and only if at least one of the $k \times k$ submatrices of A has a nonzero determinant; this is the numerical criterion for linear independence of k vectors in $\mathbf{R}^d$ that was mentioned earlier.

Linear independence of the columns of A is the same thing as injectivity of A as a linear map $\mathbf{R}^k \to \mathbf{R}^d$, so we can also say (for $k \le d$) that a matrix in $\mathrm{M}_{d\times k}(\mathbf{R})$ is injective as a linear transformation $\mathbf{R}^k \to \mathbf{R}^d$ if and only if at least one of the $k \times k$ submatrices of A has a nonzero determinant. This all works for vector spaces over a general field, not just $\mathbf{R}$.

Corollary 7.3. *Let V is a nonzero vector space. For $k \ge 1$ and linear independent $v_1, \ldots, v_k$ in V, an element $v \in V$ is a linear combination of $v_1, \ldots, v_k$ if and only if $v_1 \wedge \cdots \wedge v_k \wedge v = 0$ in $\Lambda^{k+1}(V)$.*

Proof. Because the v_i's are linearly independent, v is a linear combination of them if and only if $\{v_1, \ldots, v_k, v\}$ is linearly dependent, and that is equivalent to their wedge product vanishing by Theorem 7.1 (as a theorem for vector spaces over all fields). □

Example 7.4. In $\mathbf{R}^3$, does the vector $v = (4, -1, 1)$ lie in the span of $v_1 = (2, 1, 3)$ and $v_2 = (1, 2, 4)$? Since v_1 and v_2 are linearly independent, v is in their span if and only if $v_1 \wedge v_2 \wedge v$ vanishes in $\Lambda^3(\mathbf{R}^3)$. Let e_1, e_2, e_3 be the standard basis of $\mathbf{R}^3$. Then we compute

$$\begin{aligned} v_1 \wedge v_2 &= (2e_1 + e_2 + 3e_3) \wedge (e_1 + 2e_2 + 4e_3) \\ &= 4(e_1 \wedge e_2) + 8(e_1 \wedge e_3) - e_1 \wedge e_2 + 4(e_2 \wedge e_3) \\ &\quad\ -3(e_1 \wedge e_3) - 6(e_2 \wedge e_3) \\ &= 3(e_1 \wedge e_2) + 5(e_1 \wedge e_3) - 2(e_2 \wedge e_3) \end{aligned}$$

so

$$\begin{aligned} v_1 \wedge v_2 \wedge v &= (3(e_1 \wedge e_2) + 5(e_1 \wedge e_3) - 2(e_2 \wedge e_3)) \wedge (4e_1 - e_2 + e_3) \\ &= 3(e_1 \wedge e_2 \wedge e_3) - 5(e_1 \wedge e_3 \wedge e_2) - 8(e_2 \wedge e_3 \wedge e_1) \\ &= 3(e_1 \wedge e_2 \wedge e_3) + 5(e_1 \wedge e_2 \wedge e_3) - 8(e_1 \wedge e_2 \wedge e_3) \\ &= 0. \end{aligned}$$

Since this vanishes, $(4, -1, 1)$ is in the span of $(2, 1, 3)$ and $(1, 2, 4)$.

This method does not explicitly represent $(4, -1, 1)$ as a linear combination in the span of $(2, 1, 3)$ and $(1, 2, 4)$. (Here is one: $(4, -1, 1) = 3(2, 1, 3) - 2(1, 2, 4)$.) On the other hand, if we are only concerned with an existence question (is it in the span, not how is it in the span) then this procedure works just as the computation of a determinant does to detect invertibility of a matrix without providing a formula for the inverse matrix.

Theorem 7.5. *Let V be a nonzero vector space and $1 \le k \le \dim(V)$. Two k-element linearly independent subsets $w_1, \ldots, w_k$ and $w_1', \ldots, w_k'$ span the same subspace of V if and only if $w_1 \wedge \cdots \wedge w_k$ and $w_1' \wedge \cdots \wedge w_k'$ are nonzero scalar multiples in $\Lambda^k(V)$.*

Proof. The k-fold elementary wedge products $w_1 \wedge \cdots \wedge w_k$ and $w'_1 \wedge \cdots \wedge w'_k$ are not zero in $\Lambda^k(V)$ by Theorem 7.1, so if they are scalar multiples of each other then the scaling factor is nonzero.

First assume the w_i's and w'_i's span the same subspace W. Write the elements of one set in terms of the other, say

$$w_i = \sum_{j=1}^{k} a_{ij} w'_j$$

for scalars a_{ij}. Then

$$w_1 \wedge \cdots \wedge w_k = \left(\sum_{j=1}^{k} a_{1j} w'_j\right) \wedge \cdots \wedge \left(\sum_{j=1}^{k} a_{kj} w'_j\right),$$

and the right side can be expanded by alternating multilinearity into a scalar multiple of $w'_1 \wedge \cdots \wedge w'_k$: an elementary wedge product with a repeated w'_i is 0, and an elementary wedge product of $w'_1, \ldots, w'_k$ in some order $\pm w'_1 \wedge \cdots \wedge w'_k$.

Conversely, if $w_1 \wedge \cdots \wedge w_k = c w'_1 \wedge \cdots \wedge w'_k$ for a scalar c (necessarily nonzero), then we want to show $w_1, \ldots, w_k$ and $w'_1, \ldots, w'_k$ span the same subspace of V. Since $w_1 \wedge w_1 \wedge \cdots \wedge w_k = 0$ in $\Lambda^{k+1}(V)$ due to the repeated w_1, also

$$w_1 \wedge c w'_1 \wedge \cdots \wedge w'_k = 0,$$

so $w_1 \wedge w'_1 \wedge \cdots \wedge w'_k = 0$ since $c \neq 0$. Therefore $\{w_1, w'_1, \ldots, w'_k\}$ is linearly dependent in V by Theorem 7.1. Since $\{w'_1, \ldots, w'_k\}$ is linearly independent, w_1 must be a linear combination of $w'_1, \ldots, w'_k$. The same argument shows each $w_i \in W$ is in the span of $w'_1, \ldots, w'_k$. Therefore the span of the w_i's is contained in the span of the w'_i's, so the spans are equal by comparing dimensions. □

Corollary 7.6. *For $1 \leq k \leq \dim(V)$, we can embed the k-dimensional subspaces W of V into $\Lambda^k(V)$ as linear subspaces by associating to each basis $w_1, \ldots, w_k$ of W the linear subspace of $\Lambda^k(V)$ spanned by $w_1 \wedge \cdots \wedge w_k$.*

Proof. By Theorem 7.1, the elementary wedge product $w_1 \wedge \cdots \wedge w_k$ is not 0 in $\Lambda^k(V)$, and the scalar multiples of this elementary wedge product are independent of the choice of basis for W by the "only if" direction of Theorem 7.5. Thus it is well-defined to map the k-dimensional subspaces of W to the 1-dimensional spaces of $\Lambda^k(V)$ by sending the span of a linearly independent set $\{w_1, \ldots, w_k\}$ to the scalar multiples of $w_1 \wedge \cdots \wedge w_k$. This mapping is injective by the "if" direction ofTheorem 7.5. □

Remark 7.7. The collection of all k-dimensional subspaces of a vector space V is denoted $G_k(V)$ and is called a Grassmannian after Hermann Grassmann, who first developed exterior algebra several decades before it was appreciated by anybody else. The special case $k = 1$ is called a projective space: $\mathbf{P}(V) = G_1(V)$ is the set of one-dimensional subspaces of V. For each $k \leq \dim V$, Corollary 7.6 says there is an embedding $G_k(V) \hookrightarrow \mathbf{P}(\Lambda^k(V)) = G_1(\Lambda^k(V))$, that turns k-dimensional subspaces of V into 1-dimensional subspaces of $\Lambda^k(V)$. This embedding is called the *Plücker embedding.* The image of the Plücker embedding is typically not all of $\mathbf{P}(\Lambda^k(V))$ since elements of $\Lambda^k(V)$ are typically not elementary wedge products. The image of the Plücker embedding is a subset of $\mathbf{P}(\Lambda^k(V))$ cut out by several quadratic relations on the coordinates (the Plücker relations).

There is a generalization of Theorem 7.1 to describe linear independence in a free R-module M rather than in a vector space. Given $m_1, \ldots, m_k$ in M, their linear independence is equivalent to a property of $m_1 \wedge \cdots \wedge m_k$ in $\Lambda^k(M)$, but the property is *not* the nonvanishing of this elementary wedge product:

Theorem 7.8. *Let M be a nonzero free R-module. For $k \geq 1$, elements $m_1, \ldots, m_k$ in M are linearly independent in M if and only if $m_1 \wedge \cdots \wedge m_k$ in $\Lambda^k(M)$ is torsion-free: for $r \in R$, $r(m_1 \wedge \cdots \wedge m_k) = 0$ only when $r = 0$.*

When R is a field, we recover Theorem 7.1. What makes Theorem 7.8 more subtle than Theorem 7.1 is that a linearly independent set in a free module usually does not extend to a basis, so we can't blindly adapt the proof of Theorem 7.1 to the case of free modules.

Proof. The case $k = 1$ is trivial by the definition of a linearly independent set in a module, so we can take $k \geq 2$.

Assume $m_1, \ldots, m_k$ is a linearly independent set in M. By Corollary 5.12, the elementary wedge product $m_1 \wedge \cdots \wedge m_k$ is a linearly independent one-element subset in $\Lambda^k(M)$, so $m_1 \wedge \cdots \wedge m_k$ has no R-torsion.

If the m_i's are linearly dependent, say $r_1 m_1 + \cdots + r_k m_k = 0$ with $r_i \in R$ not all 0, we may re-index and assume $r_1 \neq 0$. Then

$$0 = (r_1 m_1 + \cdots + r_k m_k) \wedge m_2 \wedge \cdots \wedge m_k = r_1(m_1 \wedge m_2 \wedge \cdots \wedge m_k),$$

so $m_1 \wedge m_2 \wedge \cdots \wedge m_k$ has R-torsion. $\square$

Corollary 7.9. *For $d \geq 1$, a system of d equations in d unknowns*

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1d}x_d &= 0 \\ a_{21}x_1 + \cdots + a_{2d}x_d &= 0 \\ &\vdots \\ a_{d1}x_1 + \cdots + a_{dd}x_d &= 0 \end{aligned}$$

in a commutative ring R has a nonzero solution $x_1, \ldots, x_d \in R$ if and only if $\det(a_{ij})$ is a zero divisor in R.

Proof. We rewrite the theorem in terms of matrices: for $A \in \mathrm{M}_d(R)$, the equation $Av = 0$ has a nonzero solution $v \in R^d$ if and only if $\det A$ is a zero divisor in R.

We consider the negated property, that the only solution of $Av = 0$ is $v = 0$. This is equivalent to the columns of A being linearly independent. The columns are $Ae_1, \ldots, Ae_d$, and their elementary wedge product is

$$Ae_1 \wedge \cdots \wedge Ae_d = (\det A)e_1 \wedge \cdots \wedge e_d \in \Lambda^d(R^d).$$

Since $e_1 \wedge \cdots \wedge e_d$ is a basis of $\Lambda^d(R^d)$ as an R-module, $(\det A)e_1 \wedge \cdots \wedge e_d$ is torsion-free if and only if the only solution of $r \det A = 0$ is $r = 0$, which means $\det A$ is *not* a zero divisor. Thus $Av = 0$ has a nonzero solution if and only if $\det A$ is a zero divisor. $\square$

For $d \geq 1$, Theorem 7.8 characterizes linear independence of k vectors $v_1, \ldots, v_k$ in R^d when $1 \leq k \leq d$.[5] The test is that $v_1 \wedge \cdots \wedge v_k$ is torsion-free in $\Lambda^k(R^d)$. If we let

[5] A linearly independent subset of R^d has at most d terms by Corollary 5.11.

$A \in \mathrm{M}_{d\times k}(R)$ be the matrix whose columns are $v_1, \ldots, v_k$ and write $v_j = \sum_{i=1}^d a_{ij}e_i$ in terms of the the standard basis $e_1, \ldots, e_d$ of R^d then (7.1) works in R^d with no changes:

$$v_1 \wedge \cdots \wedge v_k = \sum_{1\le i_1<\cdots<i_k\le d} \begin{vmatrix} a_{1i_1} & \cdots & a_{1i_k} \\ \vdots & \ddots & \vdots \\ a_{ki_1} & \cdots & a_{ki_k} \end{vmatrix} e_{i_1} \wedge \cdots \wedge e_{i_k}. \tag{7.2}$$

Since the k-fold wedge products of $e_1, \ldots, e_d$ here are a basis of $\Lambda^k(R^d)$, $v_1 \wedge \cdots \wedge v_k$ is torsion-free in $\Lambda^k(R^d)$ precisely when the coefficients on the right side of (7.2) are not all killed by a common nonzero element of R. And since linear independence of the columns of A is the same thing as injectivity of A as a linear map $R^k \to R^d$, we can say a matrix $A \in \mathrm{M}_{d\times k}(R)$ is injective as a linear map $R^k \to R^d$ if and only if the determinants of its $k \times k$ submatrices have no common nonzero annihilator in R. (When R is a field, the only element of R with a nonzero annihilator is 0, so we recover the injectivity criterion for $d \times k$ matrices over fields when $k \le d$: some $k \times k$ submatrix has a nonzero determinant.)

Example 7.10. Let

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 5 \\ 1 & 2 \end{pmatrix} \in \mathrm{M}_{3\times 2}(\mathbf{Z}/6\mathbf{Z}).$$

Its 2×2 submatrices have determinants $8, -3, 2$, which equal $2, 3, 2$ in $\mathbf{Z}/6\mathbf{Z}$. Although none of these determinants is a unit in $\mathbf{Z}/6\mathbf{Z}$, which would be an *easy* way to see injectivity, A is still injective because 2 and 3 have no common nonzero annihilator in $\mathbf{Z}/6\mathbf{Z}$ (that is, $2r = 0$ and $3r = 0$ only for $r = 0$). In the terminology of linear independence, we showed the two columns of A are linearly independent in $(\mathbf{Z}/6\mathbf{Z})^3$. (This does *not* mean neither column is a scalar multiple of the other, but it means the stronger assertion that no linear combination of the columns is 0 except for the trivial combination with both coefficients equal to 0.)

We can also check this in $\Lambda^2((\mathbf{Z}/6\mathbf{Z})^3)$: the two columns of A are $2e_1 + e_2 + e_3$ and $2e_1 + 5e_2 + 2e_3$, and

$$(2e_1 + e_2 + e_3) \wedge (2e_1 + 5e_2 + 2e_3) = 2e_1 \wedge e_2 + 2e_1 \wedge e_3 + 3e_2 \wedge e_3,$$

which is torsion-free since $2r = 0$ and $3r = 0$ in $\mathbf{Z}/6\mathbf{Z}$ only for $r = 0$.

That the coefficients in the elementary wedge product match the determinants of the 2×2 submatrices illustrates that the rules about elementary wedge product computations really do encode everything about determinants of submatrices.

I learned the following neat use of Corollary 7.9 from [2, pp. 6–7].

Theorem 7.11. *Let M be an R-module admitting a linear injection $R^d \hookrightarrow M$ and a linear surjection $R^d \twoheadrightarrow M$, where $d \ge 1$. Then $M \cong R^d$.*

Proof. The hypotheses say M has a d-element linearly independent subset and a d-element spanning set. Call the former $e_1, \ldots, e_d$ and the latter $x_1, \ldots, x_d$, so we can write

$$M = Rx_1 + \cdots + Rx_d, \quad e_i = \sum_{j=1}^d a_{ij}x_j.$$

The x_j's span M. We want to show they are linearly independent, so they form a basis and M is free of rank d.

The expression of the e_i's in terms of the x_j's can be written as the vector-matrix equation

$$\begin{pmatrix} e_1 \\ \vdots \\ e_d \end{pmatrix} = (a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \tag{7.3}$$

where we treat $(a_{ij}) \in \mathrm{M}_d(R)$ as a matrix acting on the d-tuples in M^d.

We show by contradiction that $\Delta := \det(a_{ij})$ is not a zero divisor in R. If Δ is a zero divisor then the columns of (a_{ij}) are linearly dependent by Corollary 7.9. A square matrix and its transpose have the same determinant, so the rows of (a_{ij}) are also linearly dependent. That means there are $c_1, \ldots, c_d \in R$ not all 0 such that

$$(c_1, \ldots, c_d)(a_{ij}) = (0, \ldots, 0).$$

Using this, we multiply both sides of (7.3) on the left by $(c_1, \ldots, c_d)$ to get

$$(c_1, \ldots, c_d) \begin{pmatrix} e_1 \\ \vdots \\ e_d \end{pmatrix} = 0.$$

That says $c_1e_1 + \cdots + c_de_d = 0$, which contradicts linear independence of the e_i's. So Δ is not a zero divisor.

Suppose now that $a_1x_1 + \cdots + a_dx_d = 0$. We want to show every a_i is 0. Multiply both sides of (7.3) on the left by the cofactor matrix for (a_{ij}) (that's the matrix you multiply by to get the scalar diagonal matrix with the determinant Δ along the main diagonal):

$$\operatorname{cof}(a_{ij}) \begin{pmatrix} e_1 \\ \vdots \\ e_d \end{pmatrix} = \Delta \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}.$$

Now multiply both sides of this equation on the left by $(a_1, \ldots, a_d)$:

$$(a_1, \ldots, a_d) \operatorname{cof}(a_{ij}) \begin{pmatrix} e_1 \\ \vdots \\ e_d \end{pmatrix} = \Delta(a_1x_1 + \cdots + a_dx_d) = 0.$$

The product $(a_1, \ldots, a_d) \operatorname{cof}(a_{ij})$ is a (row) vector, say $(b_1, \ldots, b_d)$. Then $b_1e_1 + \cdots + b_de_d = 0$, so every b_i is 0 by linear independence of the e_i's. Therefore $(a_1, \ldots, a_d) \operatorname{cof}(a_{ij}) = (0, \ldots, 0)$. Multiply both sides of this on the *right* by (a_{ij}) to get $(a_1, \ldots, a_d)\Delta = (0, \ldots, 0)$, so $\Delta a_i = 0$ for all i. Since Δ is not a zero divisor, every a_i is 0 and we are done. □

8. The wedge product

By concatenating elementary wedge products, we introduce a multiplication operation between different exterior powers of a module.

Lemma 8.1. *If* $f \colon \underbrace{M \times \cdots \times M}_{k \text{ times}} \times N_1 \cdots \times N_\ell$ *is multilinear and is alternating in the M's, for $k \geq 1$, then there is a unique multilinear map* $\widetilde{f} \colon \Lambda^k(M) \times N_1 \times \cdots \times N_\ell \to N$ *such that*

$$\widetilde{f}(m_1 \wedge \cdots \wedge m_k, n_1, \ldots, n_\ell) = f(m_1, \ldots, m_k, n_1, \ldots, n_\ell).$$

Proof. Uniqueness of $\widetilde{f}$ follows from multilinearity. To prove existence of $\widetilde{f}$, fix a choice of $n_1 \in N_1, \dots, n_\ell \in N_\ell$. Define $f_{n_1,\dots,n_\ell} \colon M^k \to N$ by

$$f_{n_1,\dots,n_\ell}(m_1, \dots, m_k) = f(m_1, \dots, m_k, n_1, \dots, n_\ell).$$

Since f is multilinear, $f_{n_1,\dots,n_\ell}$ is multilinear. Since f is alternating in the m_i's, $f_{n_1,\dots,n_\ell}$ is an alternating map. Therefore there is a unique linear map $\widetilde{f}_{n_1,\dots,n_\ell} \colon \Lambda^k(M) \to N$ such that

$$\widetilde{f}_{n_1,\dots,n_\ell}(m_1 \wedge \cdots \wedge m_k) = f_{n_1,\dots,n_\ell}(m_1, \dots, m_k) = f(m_1, \dots, m_k, n_1, \dots, n_\ell).$$

on elementary wedge products.

Define $\widetilde{f} \colon \Lambda^k(M) \times N_1 \times \cdots \times N_\ell \to N$ by $\widetilde{f}(\omega, n_1, \dots, n_\ell) = f_{n_1,\dots,n_\ell}(\omega)$. Since each $f_{n_1,\dots,n_\ell}$ is linear, $\widetilde{f}$ is linear in its first component. To check $\widetilde{f}$ is linear in one of its other coordinates, we carry it out for n_1 (all the rest are similar). We want to verify that

$$\widetilde{f}(\omega, n_1 + n_1', n_2, \dots, n_\ell) = \widetilde{f}(\omega, n_1, n_2, \dots, n_\ell) + \widetilde{f}(\omega, n_1', n_2, \dots, n_\ell)$$

and

$$\widetilde{f}(\omega, rn_1, n_2, \dots, n_\ell) = r\widetilde{f}(\omega, n_1, n_2, \dots, n_\ell).$$

Both sides of both equations are additive in ω, so it suffices to check the identity when $\omega = m_1 \wedge \cdots \wedge m_k$ is an elementary wedge product. In that case the two equations turn into linearity of f in its N_1-component, and that is just a special case of the multilinearity of f. $\square$

Theorem 8.2. *Let M be an R-module. For positive integers k and ℓ, there is a unique R-bilinear map $\Lambda^k(M) \times \Lambda^\ell(M) \to \Lambda^{k+\ell}(M)$ satisfying the rule*

$$(m_1 \wedge \cdots \wedge m_k, m_1' \wedge \cdots \wedge m_\ell') \mapsto m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell'$$

on pairs of elementary wedge products.

Proof. Since the elementary wedge products span each exterior power module, and a bilinear map is determined by its values on pairs coming from spanning sets, there is *at most one* bilinear map with the prescribed behavior. As usual in this game, what needs proof is the *existence* of such a map.

Start by backing up and considering the function $f \colon M^k \times M^\ell \to \Lambda^{k+\ell}(M)$ by

$$f(m_1, \dots, m_k, m_1', \dots, m_\ell') = m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell'.$$

This is multilinear and alternating in the first k-coordinates, so by Lemma 8.1 (with $N_1, \dots, N_\ell$ being M) there is a multilinear map $\widetilde{f} \colon \Lambda^k(M) \times M^\ell \to \Lambda^{k+\ell}(M)$ such that

$$\widetilde{f}(m_1 \wedge \cdots \wedge m_k, m_1', \dots, m_\ell') = m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell'.$$

and it is alternating in its last ℓ coordinates. Therefore, again by Lemma 8.1 (with the N_j's being the single module $\Lambda^k(M)$) there is a bilinear map $B \colon \Lambda^k(M) \times \Lambda^\ell(M) \to \Lambda^{k+\ell}(M)$ such that

$$B(\omega, m_1' \wedge \cdots m_\ell') = \widetilde{f}(\omega, m_1', \dots, m_\ell'),$$

so

$$\begin{aligned} B(m_1 \wedge \cdots \wedge m_k, m_1' \wedge \cdots m_\ell') &= \widetilde{f}(m_1 \wedge \cdots \wedge m_k, m_1', \dots, m_\ell') \\ &= m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell'. \end{aligned}$$

We have produced a bilinear map on the kth and ℓth exterior powers with the desired value on pairs of elementary wedge products. $\square$

The operation constructed in Theorem 8.2, sending $\Lambda^k(M) \times \Lambda^\ell(M)$ to $\Lambda^{k+\ell}(M)$, is denoted $\wedge$ and is called the *wedge product*. We place the operation in between the elements it acts on, just like other multiplication functions in mathematics. So for any $\omega \in \Lambda^k(M)$ and $\eta \in \Lambda^\ell(M)$ we have an element $\omega \wedge \eta \in \Lambda^{k+\ell}(M)$, and this operation is bilinear in ω and η. The formula in Theorem 8.2 on two elementary wedge products looks like this:

$$(m_1 \wedge \cdots \wedge m_k) \wedge (m_1' \wedge \cdots \wedge m_\ell') = m_1 \wedge \cdots m_k \wedge m_1' \wedge \cdots \wedge m_\ell'.$$

Notice we have given a *new* meaning to the notation $\wedge$ and to the terminology "wedge product," which we have until now used in a purely formal way always in the phrase "elementary wedge product." This new operational meaning of the wedge product is *consistent* with the old formal one, *e.g.* the (new) wedge product operation $\wedge\colon \Lambda^1(M) \times \Lambda^1(M) \to \Lambda^2(M)$ sends (m, m') to $m \wedge m'$ (old notation). This is very much like the definition of $R[T]$ as formal finite sums $\sum_i a_i T^i$ where, after the ring operations are defined, the symbol T^i is recognized as the i-fold product of the element T.

We have defined a wedge product $\Lambda^k(M) \times \Lambda^\ell(M) \to \Lambda^{k+\ell}(M)$ when k and ℓ are positive. What if one of them is 0? Recall (by definition) $\Lambda^0(M) = R$. Theorem 8.2 extends to the case $k = 0$ or $\ell = 0$ if we let the maps $\Lambda^0(M) \times \Lambda^\ell(M) \to \Lambda^\ell(M)$ and $\Lambda^k(M) \times \Lambda^0(M) \to \Lambda^0(M)$ be scalar multiplication: $r \wedge \eta = r\eta$ and $\omega \wedge r = r\omega$.

Now we can think about the old notation $m_1 \wedge \cdots \wedge m_k$ in an operational way: it is the result of applying the wedge product operation k times with elements from the module $\Lambda^1(M) = M$. Since we are now able to speak about a wedge product $\omega_1 \wedge \cdots \wedge \omega_k$ where the ω_i's lie in exterior power modules $\Lambda^{k_i}(M)$, the elementary wedge products $m_1 \wedge \cdots \wedge m_k$, where the factors are in M, are merely special cases of wedge products.

Actually, to speak of $\omega_1 \wedge \cdots \wedge \omega_k$ unambiguously we need $\wedge$ to be associative! So let's check that.

Theorem 8.3. *The wedge product is associative: for $a, b, c \geq 0$, if $\omega \in \Lambda^a(M)$, $\eta \in \Lambda^b(M)$, and $\xi \in \Lambda^c(M)$, then $(\omega \wedge \eta) \wedge \xi = \omega \wedge (\eta \wedge \xi)$ in $\Lambda^{a+b+c}(M)$.*

Proof. This is easy if a, b, or c is zero (then one of ω, η, and ξ is in R, and wedging with R is just scalar multiplication), so we can assume a, b, and c are all positive.

Let f and g be the functions from $\Lambda^a(M) \times \Lambda^b(M) \times \Lambda^c(M)$ to $\Lambda^{a+b+c}(M)$ given by both choices of parentheses:

$$f(\omega, \eta, \xi) = (\omega \wedge \eta) \wedge \xi, \quad g(\omega, \eta, \xi) = \omega \wedge (\eta \wedge \xi).$$

(Note $\omega \wedge \eta \in \Lambda^{a+b}(M)$ and $\eta \wedge \xi \in \Lambda^{b+c}(M)$.) Since the wedge product on two exterior power modules is bilinear, f and g are both trilinear functions. Therefore to show $f = g$ it suffices to verify equality on triples of elementary wedge products, since they are spanning sets of $\Lambda^a(M)$, $\Lambda^b(M)$, and $\Lambda^c(M)$. On such elements the equality is obvious by the definition of the wedge product operation, so we are done. $\square$

The following theorem puts associativity of the wedge product to work.

Theorem 8.4. *Let M be an R-module. If $\Lambda^i(M)$ is finitely generated for some $i \geq 1$, then $\Lambda^j(M)$ is finitely generated for all $j \geq i$.*

The important point here is that we are not assuming M is finitely generated, only that some exterior power is finitely generated. Because exterior powers are not defined recursively, this theorem is not a tautology.[6]

Proof. It suffices to show that if $\Lambda^i(M)$ is finitely generated then $\Lambda^{i+1}(M)$ is finitely generated. Our argument is a simplification of [3, Lemma 2.1].

The module $\Lambda^i(M)$ has a finite spanning set, which we can take to be a set of elementary wedge products. Let $x_1, \dots, x_p \in M$ be the terms appearing in those elementary wedge products, so $\Lambda^i(M)$ is spanned by the i-fold wedges of $x_1, \dots, x_p$.

We will show $\Lambda^{i+1}(M)$ is spanned by the $(i+1)$-fold wedges of $x_1, \dots, x_p$. It suffices to show the span of these wedges contains all the elementary wedge products in $\Lambda^{i+1}(M)$. Choose $(i+1)$-fold elementary wedge product, say

$$y_1 \wedge \cdots \wedge y_i \wedge y_{i+1} = (y_1 \wedge \cdots \wedge y_i) \wedge y_{i+1}.$$

Since $y_1 \wedge \cdots \wedge y_i$ is in $\Lambda^i(M)$, it is an R-linear combination of i-fold wedges of $x_1, \dots, x_p$. Therefore $y_1 \wedge \cdots \wedge y_i \wedge y_{i+1}$ is an R-linear combination of expressions

$$(x_{j_1} \wedge \cdots \wedge x_{j_i}) \wedge y_{i+1} = x_{j_1} \wedge (x_{j_2} \wedge \cdots \wedge x_{j_i} \wedge y_{i+1}),$$

where the equation uses associativity of the wedge product. Since $x_{j_2} \wedge \cdots \wedge x_{j_i} \wedge y_{i+1}$ is in $\Lambda^i(M)$, it is an R-linear combination of i-fold wedges of $x_1, \dots, x_p$. □

Theorem 8.5. *If M is spanned as an R-module by $x_1, \dots, x_d$ for $d \geq 1$, then for $1 \leq k \leq d$ every element of $\Lambda^k(M)$ is a sum*

$$\omega_1 \wedge x_1 + \omega_2 \wedge x_2 + \cdots + \omega_d \wedge x_d$$

for some $\omega_i \in \Lambda^{k-1}(M)$.

We don't consider $k > d$ in this theorem since $\Lambda^k(M) = 0$ for such k by Theorem 4.1.

Proof. The result is clear when $k = 1$ since $\Lambda^0(M) = R$ by definition, so we can assume $k \geq 2$. From the beginning of Section 4, $\Lambda^k(M)$ is spanned as an R-module by the k-fold elementary wedge products $x_{i_1} \wedge \cdots \wedge x_{i_k}$ where $1 \leq i_1 < \cdots < i_k \leq d$. Using the wedge product multiplication $\Lambda^{k-1}(M) \times M \to \Lambda^k(M)$, we can write

$$x_{i_1} \wedge \cdots \wedge x_{i_k} = \eta \wedge x_{i_k},$$

where $\eta = x_{i_1} \wedge \cdots \wedge x_{i_{k-1}} \in \Lambda^{k-1}(M)$. Therefore every element of $\Lambda^k(M)$ is an R-linear combination $r_1(\eta_1 \wedge x_1) + \cdots + r_d(\eta_d \wedge x_d)$, where $r_i \in R$ and $\eta_i \in \Lambda^{k-1}(M)$. Since $r_i(\eta_i \wedge x_i) = (r_i\eta_i) \wedge x_i$ we can set $\omega_i = r_i\eta_i$ and we're done. □

Here is an analogue of Theorem 4.8 for linear maps.

Theorem 8.6. *Let $\varphi \colon M \to N$ be a linear map of R-modules. If $\wedge^i(\varphi) = 0$ for some $i \geq 1$, then $\wedge^j(\varphi) = 0$ for all $j \geq i$.*

Theorem 4.8 is the special case when $N = M$ and $\varphi = \mathrm{id}_M$ (why?).

[6]The theorem doesn't go backwards down to M, *i.e.*, if some $\Lambda^i(M)$ is finitely generated this does not imply M is finitely generated. For example, $\Lambda^2(\mathbf{Q}/\mathbf{Z}) = 0$ as a $\mathbf{Z}$-module, but $\mathbf{Q}/\mathbf{Z}$ is not finitely generated as a $\mathbf{Z}$-module.

Proof. It suffices to show $\wedge^{i+1}(\varphi) = 0$.

Since $\wedge^i(\varphi) = 0$ for any $m_1, \ldots, m_i$ in M we have

$$0 = \wedge^i(\varphi)(m_1 \wedge \cdots \wedge m_i) = \varphi(m_1) \wedge \cdots \wedge \varphi(m_i)$$

in $\Lambda^i(N)$. For any $m_1, \ldots, m_{i+1}$ in M, in $\Lambda^{i+1}(N)$ we have

$$\begin{aligned} \wedge^{i+1}(\varphi)(m_1 \wedge \cdots \wedge m_i \wedge m_{i+1}) &= \varphi(m_1) \wedge \cdots \wedge \varphi(m_i) \wedge \varphi(m_{i+1}) \\ &= (\varphi(m_1) \wedge \cdots \wedge \varphi(m_i)) \wedge \varphi(m_{i+1}) \\ &= 0 \wedge \varphi(m_{i+1}) \\ &= 0. \end{aligned}$$

Such terms span the image of $\wedge^{i+1}(\varphi)$, so $\wedge^{i+1}(\varphi) = 0$. □

In an elementary wedge product $m_1 \wedge \cdots \wedge m_k$, where the factors are in $M = \Lambda^1(M)$, transposing two of them introduces a sign change. What is the sign-change rule for transposing factors in a wedge product $\omega_1 \wedge \cdots \wedge \omega_k$? By associativity, we just need to understand how a single wedge product $\omega \wedge \eta$ changes when the factors are reversed.

Theorem 8.7. *For $\omega \in \Lambda^k(M)$ and $\eta \in \Lambda^\ell(M)$, where $k, \ell \geq 0$, $\omega \wedge \eta = (-1)^{k\ell}\eta \wedge \omega$.*

Proof. This is trivial if $k = 0$ or $\ell = 0$ (in which case the wedge product is simply scaling and $(-1)^{k\ell} = 1$), so we can take k and ℓ positive.

Both sides of the desired equation are bilinear functions of ω and η, so to verify equality for all ω and η it suffices to do so on spanning sets of the modules. Thus we can take $\omega = m_1 \wedge \cdots \wedge m_k$ and $\eta = m_1' \wedge \cdots \wedge m_\ell'$, so we want to show

$$m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell' = (-1)^{k\ell}(m_1' \wedge \cdots \wedge m_\ell' \wedge m_1 \wedge \cdots \wedge m_k). \tag{8.1}$$

Starting with the expression on the left, we successively move $m_1', m_2', \ldots, m_\ell'$ to the front. First move m_1' past each of the m_i's, which is a total of k swaps, so

$$m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell' = (-1)^k m_1' \wedge m_1 \wedge \cdots \wedge m_k \wedge m_2' \wedge \cdots \wedge m_\ell'.$$

Now move m_2' past every m_i, introducing another set of k sign changes:

$$m_1 \wedge \cdots \wedge m_k \wedge m_1' \wedge \cdots \wedge m_\ell' = (-1)^{2k} m_1' \wedge m_2' \wedge m_1 \wedge \cdots \wedge m_k \wedge m_3' \wedge \cdots \wedge m_\ell'.$$

Repeat until m_ℓ' has been moved past every m_i. In all, there are $k\ell$ swaps, so the overall sign at the end is $(-1)^{k\ell}$. □

Theorem 8.8. *For odd k and $\omega \in \Lambda^k(M)$, $\omega \wedge \omega = 0$.*

Proof. There is a quick proof using Theorem 8.7 when $2 \in R^\times$: $\omega \wedge \omega = (-1)^{k^2}(\omega \wedge \omega) = -(\omega \wedge \omega)$, so $2(\omega \wedge \omega) = 0$, so $\omega \wedge \omega = 0$.

To handle the general case when 2 may not be a unit, write

$$\omega = \sum_{i_1, \ldots, i_k} c_{i_1, \ldots, i_k} m_{i_1} \wedge \cdots \wedge m_{i_k} = \sum_I c_I \omega_I,$$

where $I = (i_1, \ldots, i_k)$ is a multi-index of k integers and $\omega_I = \omega_{(i_1, \ldots, i_k)}$ is an abbreviation for $m_{i_1} \wedge \cdots \wedge m_{i_k}$. Then, by the bilinearity of the wedge product,

$$\omega \wedge \omega = \sum_I c_I m_I \wedge \sum_J c_J \omega_J = \sum_{I,J} c_I c_J \omega_I \wedge \omega_J, \tag{8.2}$$

where I and J run over the same set of multi-indices. Note each $\omega_I \wedge \omega_J$ is an elementary wedge product with $2k$ factors.

The multi-indices I and J could be equal. In that case $\omega_I \wedge \omega_J = 0$ since it is a $2k$-fold elementary wedge product with repeated factors. When $I \neq J$ the double sum in (8.2) contains

$$c_I c_J \omega_I \wedge \omega_J + c_J c_I \omega_J \wedge \omega_I = c_I c_J(\omega_I \wedge \omega_J + \omega_J \wedge \omega_I).$$

Since ω_I and ω_J are in $\Lambda^k(M)$, $\omega_J \wedge \omega_I = (-1)^{k^2}(\omega_I \wedge \omega_J) = -(\omega_I \wedge \omega_J)$, so $c_I c_J(\omega_I \wedge \omega_J + \omega_J \wedge \omega_I) = 0$. □

What about Theorem 8.8 when k is even? If $\omega = m_1 \wedge \cdots \wedge m_k$ is an *elementary* wedge product in $\Lambda^k(M)$ then $\omega \wedge \omega$ vanishes since it is an elementary wedge product with a repeated factor from M. But it is not generally true that $\omega \wedge \omega = 0$ for all $\omega \in \Lambda^k(M)$.

Example 8.9. For $k \geq 2$, let M be finite free with linearly independent subset $e_1, \ldots, e_{2k}$. Set $\omega = e_1 \wedge \cdots \wedge e_k + e_{k+1} \wedge \cdots \wedge e_{2k} \in \Lambda^k(M)$. This is a sum of two elementary wedge products, and

$$\omega \wedge \omega = 2e_1 \wedge \cdots \wedge e_k \wedge e_{k+1} \wedge \cdots \wedge e_{2k} \in \Lambda^{2k}(M).$$

By Corollary 5.12 and Theorem 7.8, $e_1 \wedge \cdots \wedge e_k \wedge e_{k+1} \wedge \cdots \wedge e_{2k}$ is torsion-free in $\Lambda^{2k}(M)$, so when $2 \neq 0$ in R we have $\omega \wedge \omega \neq 0$. Elementary wedge products always "square" to 0, so ω is *not* an elementary wedge product when $2 \neq 0$ in R.

To get practice computing in an exterior power module, we look at the equation $v \wedge \omega = 0$ in $\Lambda^{k+1}(V)$, where V is a vector space and $\omega \in \Lambda^k(V)$.

Theorem 8.10. *Let V be a vector space. For $k \geq 0$ and nonzero $\omega \in \Lambda^k(V)$,*

$$\dim(\{v \in V : v \wedge \omega = 0\}) \leq k,$$

with equality if and only if ω is an elementary wedge product.

Proof. The result is obvious if $k = 0$, so we can suppose $k \geq 1$. Let $v_1, \ldots, v_d$ be linearly independent vectors in V that each wedge ω to 0. We want to show $d \leq k$. Since V might be infinite-dimensional, we first create a suitable finite-dimensional subspace W in which we can work. (If you want to assume V is finite-dimensional, set $W = V$ and skip the rest of this paragraph.) Let W be the span of $v_1, \ldots, v_d$ and the nonzero vectors appearing in some fixed representation of ω as a finite sum of elementary wedge products in $\Lambda^k(V)$. Then W is finite-dimensional. There's a natural embedding $W \hookrightarrow V$ and we get an embedding $\Lambda^\ell(W) \hookrightarrow \Lambda^\ell(V)$ in a natural way for all ℓ by Corollary 5.9. If we view each v_i in W and ω in $\Lambda^k(W)$ then the condition $v_i \wedge \omega = 0$ in $\Lambda^{k+1}(V)$ implies $v_i \wedge \omega = 0$ in $\Lambda^{k+1}(W)$.

Set $n = \dim W$, so obviously $d \leq n$. If $k \geq n$ then obviously $d \leq k$, so we can suppose $k \leq n - 1$.

Extend $\{v_1, \ldots, v_d\}$ to a basis $\{v_1, \ldots, v_n\}$ of W. Using this basis we can write ω in $\Lambda^k(W)$ as a finite sum of linearly independent elementary wedge products:

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} c_{i_1,\ldots,i_k} v_{i_1} \wedge \cdots \wedge v_{i_k}$$

and some coefficient is not 0. Fix i between 1 and d, and compute $v_i \wedge \omega$ using this formula:

$$\begin{aligned} 0 &= v_i \wedge \omega \\ &= \sum_{1 \le i_1 < \cdots < i_k \le n} c_{i_1,\dots,i_k} v_i \wedge v_{i_1} \wedge \cdots \wedge v_{i_k} \\ &= \sum_{\substack{1 \le i_1 < \cdots < i_k \le n \\ i_1,\dots,i_k \neq i}} c_{i_1,\dots,i_k} v_i \wedge v_{i_1} \wedge \cdots \wedge v_{i_k}. \end{aligned}$$

This equation is taking place in $\Lambda^{k+1}(W)$, where the $(k+1)$-fold wedges $v_i \wedge v_{i_1} \wedge \cdots \wedge v_{i_k}$ for $i \notin \{i_1, \dots, i_k\}$ are linearly independent. Therefore the coefficients here are all 0:

$$i \notin \{i_1, \dots, i_k\} \Rightarrow c_{i_1,\dots,i_k} = 0.$$

Here i was any number from 1 to d, so

$$c_{i_1,\dots,i_k} \neq 0 \Rightarrow \{1, \dots, d\} \subset \{i_1, \dots, i_k\}.$$

There is at least one nonzero coefficient, so we must have $\{1, \dots, d\} \subset \{i_1, \dots, i_k\}$ for some k-tuple of indices. Counting the two sets, $d \le k$.

If $d = k$ then $\{1, \dots, k\} = \{i_1, \dots, i_k\}$, which allows just one nonzero term and $\omega = c_{1,\dots,k} v_1 \wedge \cdots \wedge v_k$, which is an elementary wedge product. Conversely, if ω is a nonzero elementary wedge product in $\Lambda^k(V)$ then $\{v \in V : \omega \wedge v = 0\}$ has dimension k by Corollary 7.3 and associativity of the wedge product. □

Theorem 8.11. *Let V be a vector space and $k \ge 1$. For nonzero $v \in V$ and $\omega \in \Lambda^k(V)$, $v \wedge \omega = 0$ if and only if $\omega = v \wedge \eta$ for some $\eta \in \Lambda^{k-1}(V)$.*

Proof. By associativity, if $\omega = v \wedge \eta$ then $v \wedge \omega = v \wedge (v \wedge \eta) = (v \wedge v) \wedge \eta = 0$. The point of the theorem is that the converse direction holds: if $v \wedge \omega = 0$ then we can write $\omega = v \wedge \eta$ for some η.

As in the proof of the previous theorem, we can reduce to the finite-dimensional case (details left to the reader), so we'll just take V to be a finite-dimensional vector space. Set $n = \dim V \ge 1$. Since $\omega \in \Lambda^k(V)$, if $k > n$ then $\omega = 0$ and we can trivially write $\omega = v \wedge 0$. So we may suppose $k \le n$. If $k = 0$ then $\omega \in \Lambda^0(V)$ is a scalar and wedging with ω is scalar multiplication, so the conditions $v \neq 0$ and $v \wedge \omega = 0$ imply $\omega = 0$. Thus again $\omega = v \wedge 0$.

Now suppose $1 \le k \le n$. Extend v to a basis of V, say $v_1, \dots, v_n$ where $v = v_1$. If $k = n$ the condition $v \wedge \omega = 0$ is automatic since $\Lambda^{n+1}(V) = 0$. And it is also automatic that ω is "divisible" by v if $k = n$: $\Lambda^n(V)$ has basis $v_1 \wedge \cdots \wedge v_n$, so $\omega = c(v_1 \wedge \cdots \wedge v_n)$ for some scalar c. Then $\omega = v \wedge \eta$ where $\eta = cv_2 \wedge \cdots \wedge v_n$.

We now assume $1 \le k \le n-1$ (so $n \ge 2$). Using the basis of $\Lambda^k(V)$ coming from our chosen basis of V that includes v as the first member v_1,

$$\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} c_{i_1,\dots,i_k} v_{i_1} \wedge \cdots \wedge v_{i_k} \tag{8.3}$$

with scalar coefficients. Since $v_1 = v$, $v \wedge v_1 = 0$ so

$$v \wedge \omega = \sum_{2 \le i_1 < \cdots < i_k \le n} c_{i_1,\dots,i_k} v \wedge v_{i_1} \wedge \cdots \wedge v_{i_k}. \tag{8.4}$$

The elementary wedge products on the right are part of a basis of $\Lambda^{k+1}(V)$, so from $v \wedge \omega = 0$ we see all the coefficients in (8.4) vanish. Thus in (8.3) the only nonzero terms are among those with $i_1 = 1$, so we can pull out $v_1 = v$:

$$\omega = v \wedge \sum_{1<i_2<\cdots<i_k\leq n} c_{1,i_2,\ldots,i_k} v_{i_2} \wedge \cdots \wedge v_{i_k}.$$

Let η be the large sum here, so $\omega = v \wedge \eta$. $\square$

9. The Künneth formula and modules over a PID

In this section we describe an important formula expressing an exterior power of a direct sum $M \oplus N$ in terms of exterior powers of M and N. Here's that formula.

Theorem 9.1. *Let M and N be R-modules. For $k \geq 0$, there is an R-module isomorphism*

$$\Lambda^k(M \oplus N) \cong \bigoplus_{i=0}^{k} (\Lambda^i(M) \otimes_R \Lambda^{k-i}(N)). \tag{9.1}$$

The isomorphism in this theorem is called the Künneth formula (for exterior powers). It formally resembles the binomial theorem. When M and N are free R-modules of respective ranks a and b, where $a+b \geq k$, all the exterior powers in the Künneth formula are free with rank given by a binomial coefficient (see Theorem 4.2), and computing the ranks for both sides tells us $\binom{a+b}{k} = \sum_{i=0}^{k} \binom{a}{i}\binom{b}{k-i}$, which is called the Vandermonde convolution formula for binomial coefficients.

Example 9.2. When k is 2 and 3, the Künneth formula says (since $\Lambda^0(M) = \Lambda^0(N) = R$)

$$\begin{aligned} \Lambda^2(M \oplus N) &\cong \Lambda^2(M) \oplus (M \otimes_R N) \oplus \Lambda^2(N), \\ \Lambda^3(M \oplus N) &\cong \Lambda^3(M) \oplus (\Lambda^2(M) \otimes_R N) \oplus (M \otimes_R \Lambda^2(N)) \oplus \Lambda^3(N). \end{aligned} \tag{9.2}$$

To appreciate the Kunneth formula (9.1), we'll apply it before proving it. For the application we have in mind, we need to extend (9.1) to direct sums of more than two modules.

Corollary 9.3. *For $k \geq 0$ and R-modules $M_1, \ldots, M_\ell$ there is an R-module isomorphism*

$$\Lambda^k(M_1 \oplus \cdots \oplus M_\ell) \cong \bigoplus_{i_1+\cdots+i_\ell=k} \Lambda^{i_1}(M_1) \otimes_R \cdots \otimes_R \Lambda^{i_\ell}(M_\ell),$$

where the direct sum runs over ℓ-tuples of nonnegative integers with sum k.

Proof. We induct on ℓ. The case $\ell = 1$ is trivial (both sides are $\Lambda^k(M_1)$), and the case $\ell = 2$ is (9.1). If $\ell \geq 3$ and the corollary is proved for every direct sum of $\ell - 1$ R-modules, then for a direct sum of ℓ R-modules $M_1, \ldots, M_\ell$ we have (using induction in the third step)

$$\begin{aligned} \Lambda^k(M_1 \oplus \cdots \oplus M_\ell) &= \Lambda^k((M_1 \oplus \cdots \oplus M_{\ell-1}) \oplus M_\ell) \\ &\cong \bigoplus_{i=0}^{k} \Lambda^i(M_1 \oplus \cdots \oplus M_{\ell-1}) \otimes_R \Lambda^{k-i}(M_\ell) \quad \text{by (9.1)} \\ &\cong \bigoplus_{i=0}^{k} \left(\bigoplus_{i_1+\cdots+i_{\ell-1}=i} \Lambda^{i_1}(M_1) \otimes_R \cdots \otimes_R \Lambda^{i_{\ell-1}}(M_{\ell-1}) \right) \otimes_R \Lambda^{k-i}(M_\ell) \\ &\cong \bigoplus_{i_1+\cdots+i_\ell=k} \Lambda^{i_1}(M_1) \otimes_R \cdots \otimes_R \Lambda^{i_{\ell-1}}(M_{\ell-1}) \otimes_R \Lambda^{i_\ell}(M_\ell) \ \text{ with } i_\ell = k - i. \square \end{aligned}$$

The next theorem is an application of the Künneth formula to the structure of finitely generated torsion modules over a PID. For a PID R and nonzero finitely generated torsion R-module M, the classification of finitely generated modules over a PID tells us that

$$M \cong R/(a_1) \oplus \cdots \oplus R/(a_\ell) \quad \text{where } (a_1) \subset (a_2) \subset \cdots \subset (a_\ell) \tag{9.3}$$

for $a_1, \ldots, a_\ell$ in R that are not 0 or units. In terms of divisibility, $a_\ell \mid \cdots \mid a_2 \mid a_1$.[7] Each of the modules $R/(a_1), \ldots, R/(a_\ell)$ is killed from multiplication by a_1 since $a_1, \ldots, a_\ell$ divide a_1, and $R/(a_1)$ is only killed by multiples of a_1, so $\boxed{\mathrm{Ann}_R(M) = (a_1)}$. For $j > 1$, the next theorem gives a formula for the ideal (a_j) as the annihilator ideal of an exterior power of M. (Similar results are in [4, Sect. 8] and [5, Prop. 4.20, Chap. XIII].)

Theorem 9.4. *For a* PID R *and nonzero finitely generated torsion* R*-module* M *decomposed into cyclic torsion modules as in* (9.3), $\boxed{\mathrm{Ann}_R(\Lambda^j(M)) = (a_j)}$ *for* $j = 1, \ldots, \ell$.[8]

Since M in (9.3) has a ℓ-element spanning set, $\Lambda^j(M) = 0$ for $j > \ell$ by Theorem 4.1. We have $\Lambda^j(M) \neq 0$ for $j \leq \ell$ since $\mathrm{Ann}_R((0)) = R$ while $\mathrm{Ann}_R(\Lambda^j(M)) = (a_j) \neq R$. Thus the number ℓ of cyclic summands for M in (9.3) is the largest $j \geq 1$ such that $\Lambda^j(M) \neq 0$. (To make that true if $M = (0)$ too, use the largest $j \geq 0$: $\Lambda^0((0)) = R$, $\Lambda^j((0)) = 0$ for $j > 0$.)

Proof. By (9.3) and Corollary 9.3,

$$\Lambda^j(M) \cong \bigoplus_{i_1+\cdots+i_\ell=j} \Lambda^{i_1}(R/(a_1)) \otimes_R \cdots \otimes_R \Lambda^{i_\ell}(R/(a_\ell)). \tag{9.4}$$

The exterior powers of a cyclic torsion module $R/(a)$ (for $a \neq 0$) are quite limited. Since $R/(a)$ has a 1-element spanning set as an R-module, $\Lambda^j(R/(a)) = 0$ for $j > 1$ by Theorem 4.1. Therefore the term in the direct sum (9.4) associated to $(i_1, \ldots, i_\ell)$ is nonzero if and only $i_1, \ldots, i_\ell$ are *at most* 1: some are 0 and some are 1. That requires $j = i_1 + \cdots + i_\ell \leq \ell$.

Since $\Lambda^0(R/(a)) = R$, $\Lambda^1(R/(a)) = R/(a)$, and tensoring R-modules with R has no effect up to isomorphism, if $j \leq \ell$ and the tensor product in (9.4) associated to $(i_1, \ldots, i_\ell)$ is nonzero then that tensor product uses j cyclic modules among $R/(a_1), \ldots, R/(a_\ell)$. What could such an j-fold tensor product look like?

Since $R/I \otimes_R R/J \cong R/(I + J)$ as R-modules, extending that to more than two cyclic modules shows the summand in (9.4) where $i_1, \ldots, i_j = 1$ (later i's are 0) is isomorphic to

$$R/(a_1) \otimes_R R/(a_2) \otimes_R \cdots \otimes_R R/(a_j) \cong R/(a_1, \ldots, a_j) = R/(a_j) \tag{9.5}$$

from $a_j \mid \cdots \mid a_2 \mid a_1$. For other j-fold tensor products among $R/(a_1), \ldots, R/(a_\ell)$, we need to use some $R/(a_i)$ where $i \geq j$ (why?). This a_i divides a_j, so each j-fold tensor product among $R/(a_1), \ldots, R/(a_\ell)$ is isomorphic to some $R/(b)$ where $b \mid a_j$. Since $a_j(R/(b)) = 0$ when $b \mid a_j$, a_j kills each summand in (9.4) and thus a_j kills $\Lambda^j(M)$. Hence $(a_j) \subset \mathrm{Ann}_R(\Lambda^j(M))$. The j-fold tensor product in (9.5) is a summand in (9.4) and is killed precisely by the multiples of a_j, so $\mathrm{Ann}_R(\Lambda^j(M)) = (a_j)$. $\square$

In case the notation in the previous proof was a bit hard to parse in places, here is an example of the main ideas with $\ell = 3$.

[7]Pay attention to the order of divisibility of the a_i, which may be reversed from the typical ordering.

[8]If we indexed the terms in (9.3) to have $a_1 \mid a_2 \mid \cdots \mid a_\ell$ then $\mathrm{Ann}_R(\Lambda^j(M)) = (a_{\ell+1-j})$, which is clunky.

Example 9.5. Suppose $M \cong R/(a_1) \oplus R/(a_2) \oplus R/(a_3)$ with $a_3 \mid a_2 \mid a_1$, or equivalently $(a_1) \subset (a_2) \subset (a_3)$. Then $\mathrm{Ann}_R(\Lambda^1(M)) = \mathrm{Ann}_R(M) = (a_1)$. We'll show $\mathrm{Ann}_R(\Lambda^2(M)) = (a_2)$ and $\mathrm{Ann}_R(\Lambda^3(M)) = (a_3)$.

Since $\Lambda^j(R/(a)) = 0$ for $j > 1$ and $\Lambda^1(R/(a)) = R/(a)$, the Künneth formula for $\Lambda^2(M)$ says this module is isomorphic to a direct sum of three nonzero tensor products:

$$\Lambda^2(M) \cong (R/(a_1) \otimes_R R/(a_2)) \oplus (R/(a_1) \otimes_R R/(a_3)) \oplus (R/(a_2) \otimes_R R/(a_3)).$$

Using $R/I \otimes_R R/J \cong R/(I + J)$,

$$\Lambda^2(M) \cong R/(a_1, a_2) \oplus R/(a_1, a_3) \oplus R/(a_2, a_3).$$

From the relations $a_3 \mid a_2 \mid a_1$, $\Lambda^2(M) \cong R/(a_2) \oplus R/(a_3) \oplus R/(a_3)$ and this has annihilator ideal (a_2) because $a_3 \mid a_2$. Thus $\mathrm{Ann}_R(\Lambda^2(M)) = (a_2)$.

The Künneth formula for $\Lambda^3(M)$ says this module is isomorphic to the single nonzero tensor product $R/(a_1) \otimes_R R/(a_2) \otimes_R R/(a_3) \cong R/(a_1, a_2, a_3)$. Since $a_3 \mid a_2 \mid a_1$, the ideal (a_1, a_2, a_3) equals (a_3), so $\Lambda^3(M) \cong R/(a_3)$. Thus $\mathrm{Ann}_R(\Lambda^3(M)) = (a_3)$. That concludes this example.

Now that we've seen an interesting use of Theorem 9.1 (the Künneth formula), let's prove the theorem. The details in the proof below are rather long (it is all we do for the rest of this section), so we will prove the case $k = 2$ first, where many of the ideas for general k can be seen but with less burdensome notation. Another proof could show the module $\bigoplus_{i=0}^k \Lambda^i(M) \otimes_R \Lambda^{k-i}(N)$ and an alternating multilinear map to it from $(M \oplus N)^k$ fits the universal mapping property of $\Lambda^k(M \oplus N)$, so the isomorphism in Theorem 9.1 is then immediate from the isomorphism of two objects satisfying the same universal mapping property. Yet another approach to proving Theorem 9.1 is indicated at the very end of the final section.

Proof. At $k = 0$ both sides are R and at $k = 1$ both sides are isomorphic to $M \oplus N$, so we can suppose $k \geq 2$.

Step 1: The case $k = 2$. The right side of the isomorphism for $k = 2$ was presented in simplified form in (9.2). Referring to that, we want to build inverse R-linear maps

$$f \colon \Lambda^2(M \oplus N) \to \Lambda^2(M) \oplus (M \otimes_R N) \oplus \Lambda^2(N) \tag{9.6}$$

and

$$g \colon \Lambda^2(M) \oplus (M \otimes_R N) \oplus \Lambda^2(N) \to \Lambda^2(M \oplus N). \tag{9.7}$$

The motivation for the way we'll define f and g is based on the expansion of elementary wedge products in $\Lambda^2(M \oplus N)$ using bilinearity and alternatingness (is that a word?):

$$\begin{aligned}(m_1, n_1) \wedge (m_2, n_2) &= ((m_1, 0) + (0, n_1)) \wedge ((m_2, 0) + (0, n_2)) \\ &= (m_1, 0) \wedge (m_2, 0) + (m_1, 0) \wedge (0, n_2) + (0, n_1) \wedge (m_2, 0) + (0, n_1) \wedge (0, n_2) \\ &= (m_1, 0) \wedge (m_2, 0) + (m_1, 0) \wedge (0, n_2) - (m_2, 0) \wedge (0, n_1) + (0, n_1) \wedge (0, n_2) \quad (9.8)\end{aligned}$$

In (9.8), the first term resembles $m_1 \wedge m_2$ (element of $\Lambda^2(M)$), the last term resembles $n_1 \wedge n_2$ (element of $\Lambda^2(N)$), and the middle two terms look like formal products of something in M and something in N, which is what elements of $M \otimes_R N$ are. With that in mind, in order to *define* f in (9.6), we start with an alternating bilinear function $(M \oplus N)^2 \to \Lambda^2(M) \oplus (M \otimes_R N) \oplus \Lambda^2(N)$ having values resembling the right side of (9.8).

Let $(M \oplus N)^2 \to \Lambda^2(M) \oplus (M \otimes_R N) \oplus \Lambda^2(N)$ by the rule

$$((m_1, n_1), (m_2, n_2)) \mapsto (m_1 \wedge m_2, m_1 \otimes n_2 - m_2 \otimes n_1, n_1 \wedge n_2), \tag{9.9}$$

where the three components on the right are all based on (9.8). This function on $(M \oplus N)^2$ is bilinear: it is linear in (m_1, n_1) when (m_2, n_2) is fixed and *vice versa*. It is also alternating: if $(m_1, n_1) = (m_2, n_2)$ then the value in (9.9) is 0. Therefore we get a linear map f in (9.6) where on elementary wedge products values are described by (9.9):

$$f((m_1, n_1) \wedge (m_2, n_2)) = (m_1 \wedge m_2, m_1 \otimes n_2 - m_2 \otimes n_1, n_1 \wedge n_2). \tag{9.10}$$

To build g in (9.7) we will define three linear maps from $\Lambda^2(M)$, $M \otimes_R N$, and $\Lambda^2(N)$ to $\Lambda^2(M \oplus N)$ and then add their values out of a direct sum. Define

$$M \times M \to \Lambda^2(M \oplus N), \quad M \times N \to \Lambda^2(M \oplus N), \quad N \times N \to \Lambda^2(M \oplus N)$$

by the rules

$$(m_1, m_2) \mapsto (m_1, 0) \wedge (m_2, 0), \quad (m, n) \mapsto (m, 0) \wedge (0, n), \quad (n_1, n_2) \mapsto (0, n_1) \wedge (0, n_2).$$

All three maps are bilinear and the first and third are alternating, so we get linear maps

$$\Lambda^2(M) \to \Lambda^2(M \oplus N), \quad M \otimes_R N \to \Lambda^2(M \oplus N), \quad \Lambda^2(N) \to \Lambda^2(M \oplus N)$$

that on elementary wedge products or elementary tensors have the effects

$$m_1 \wedge m_2 \mapsto (m_1, 0) \wedge (m_2, 0), \quad m \otimes n \mapsto (m, 0) \wedge (0, n), \quad n_1 \wedge n_2 \mapsto (0, n_1) \wedge (0, n_2).$$

Package these together out of a direct sum to define g in (9.7): on two elementary wedge products and an elementary tensor (which as a triple span the direct sum),

$$g(m_1 \wedge m_2, m \otimes n, n_1 \wedge n_2) = (m_1, 0) \wedge (m_2, 0) + (m, 0) \wedge (0, n) + (0, n_1) \wedge (0, n_2). \tag{9.11}$$

This is linear since the map on each component in $\Lambda^2(M) \oplus (M \otimes_R N) \oplus \Lambda^2(N)$ is linear.

It remains to check that the maps in (9.6) and (9.7) are inverses. Because of spanning sets built from elementary wedge products and elementary tensors, it suffices to check the linear maps f and g fix the spanning sets. Applying g to both sides of (9.10),

$$\begin{aligned} g(f((m_1, n_1) \wedge (m_2, n_2))) &= g(m_1 \wedge m_2, m_1 \otimes n_2 - m_2 \otimes n_1, n_1 \wedge n_2) \\ &= (m_1, 0) \wedge (m_2, 0) + (m_1, 0) \wedge (0, n_2) - (m_2, 0) \wedge (0, n_1) + (0, n_1) \wedge (0, n_2) \\ &= (m_1, 0) \wedge (m_2, n_2) + (0, n_1) \wedge (m_2, n_2) \\ &= (m_1, n_1) \wedge (m_2, n_2), \end{aligned}$$

and applying f to both sides of (9.11),

$$\begin{aligned} f(g(m_1 \wedge m_2, m \otimes n, n_1 \wedge n_2)) &= f((m_1, 0) \wedge (m_2, 0)) + f((m, 0) \wedge (0, n)) + f((0, n_1) \wedge (0, n_2)) \\ &= (m_1 \wedge m_2, 0, 0) + (0, m \otimes n, 0) + (0, 0, n_1 \wedge n_2) \text{ by (9.10)} \\ &= (m_1 \wedge m_2, m \otimes n, n_1 \wedge n_2). \end{aligned}$$

That proves f and g in (9.6) and (9.7) are inverses, so the case $k = 2$ is settled!

Now we turn to the general case of the theorem. We will construct inverse R-linear maps

$$f \colon \Lambda^k(M \oplus N) \to \bigoplus_{i=0}^{k} (\Lambda^i(M) \otimes_R \Lambda^{k-i}(N)), \tag{9.12}$$

and

$$g\colon \bigoplus_{i=0}^{k}(\Lambda^i(M)\otimes_R \Lambda^{k-i}(N)) \to \Lambda^k(M\oplus N). \tag{9.13}$$

Before we get into the details, let's think informally about what is going on. All we are doing is rewriting elementary wedge products in $\Lambda^k(M\oplus N)$ to put $(m,0)$ terms in front of $(0,n)$ terms, and if such a wedge product has i terms $(m,0)$ and $k-i$ terms $(0,n)$ then it should behave like a "formal product" of something in $\Lambda^i(M)$ and something in $\Lambda^{k-i}(N)$. Since "formal products" in modules means "tensor product", we should formalize the idea of a wedge product of $(m,0)$'s and $(0,n)$'s as an elementary tensor in $\Lambda^i(M)\otimes_R \Lambda^{k-i}(N)$. Adding such things up over all i amounts to taking a direct sum as i runs from 0 to k. If you keep this basic intuition in mind then it can help you wade through the tedious bookkeeping that is coming up.

Since it is simpler to construct g first, that is our next step.

Step 2: Build g.

To map *out* of a direct sum, we'll make a linear map $g_i\colon \Lambda^i(M)\otimes_R\Lambda^{k-i}(N)\to\Lambda^k(M\oplus N)$ for $0\le i\le k$ and then define g on the direct sum to be the sum of g_i on the ith component of the direct sum:

$$g(\omega_0,\omega_1,\dots,\omega_k) = g_0(\omega_0)+\cdots+g_k(\omega_k). \tag{9.14}$$

The standard inclusions $M\to M\oplus N$ and $N\to M\oplus N$, where $m\mapsto(m,0)$ and $n\mapsto(0,n)$, lead to linear maps $\Lambda^i(M)\to\Lambda^i(M\oplus N)$ and $\Lambda^{k-i}(N)\to\Lambda^{k-i}(M\oplus N)$ on exterior powers (Theorem 5.1). These can be described on elementary wedge products as

$$m_1\wedge\cdots\wedge m_i\mapsto (m_1,0)\wedge\cdots\wedge(m_i,0)$$

and

$$n_1\wedge\cdots\wedge n_{k-i}\mapsto (0,n_1)\wedge\cdots\wedge(0,n_{k-i}),$$

where $i>0$ for the first map and $i<k$ for the second map. (If $i=0$ then the first map is the identity $R\to R$ and if $i=k$ then the second map is the identity $R\to R$.) Tensor these two linear maps $\Lambda^i(M)\to\Lambda^i(M\oplus N)$ and $\Lambda^{k-i}(N)\to\Lambda^{k-i}(M\oplus N)$ to get a linear map

$$\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)\to\Lambda^i(M\oplus N)\otimes_R\Lambda^{k-i}(M\oplus N). \tag{9.15}$$

When $i=0$ this is a linear map $\Lambda^k(N)\to\Lambda^k(M\oplus N)$, and when $i=k$ this is a linear map $\Lambda^k(M)\to\Lambda^k(M\oplus N)$, which are the kth exterior powers of the standard inclusion maps of M and N into $M\oplus N$.

The wedge product $\Lambda^i(M\oplus N)\times\Lambda^{k-i}(M\oplus N)\to\Lambda^k(M\oplus N)$ in Theorem 8.2 is R-bilinear, so it induces an R-linear map $\Lambda^i(M\oplus N)\otimes_R\Lambda^{k-i}(M\oplus N)\to\Lambda^k(M\oplus N)$. The composition of this R-linear map with the R-linear map in (9.15) gives us an R-linear map

$$\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)\to\Lambda^i(M\oplus N)\otimes_R\Lambda^{k-i}(M\oplus N)\to\Lambda^k(M\oplus N)$$

and this composition is the definition of g_i. Tracing through this definition, the effect of g_i on an elementary tensor of elementary wedge products is

$$g_i((m_1\wedge\cdots\wedge m_i)\otimes(n_1\wedge\cdots\wedge n_{k-i})) = (m_1,0)\wedge\cdots\wedge(m_i,0)\wedge(0,n_1)\wedge\cdots\wedge(0,n_{k-i})$$

for $0<i<k$. If $i=0$ or $i=k$ then $g_0\colon\Lambda^k(N)\to\Lambda^k(M\oplus N)$ and $g_k\colon\Lambda^k(M)\to\Lambda^k(M\oplus N)$ are the kth exterior powers of the standard inclusions of N and M into $M\oplus N$:

$$g_0(n_1\wedge\cdots\wedge n_k) = (0,n_1)\wedge\cdots\wedge(0,n_k),\quad g_k(m_1\wedge\cdots\wedge m_k) = (m_1,0)\wedge\cdots\wedge(m_k,0).$$

Step 3: Build f for the general case.

We will create a linear map $f\colon \Lambda^k(M\oplus N) \to \bigoplus_{i=0}^k \Lambda^i(M)\otimes_R \Lambda^{k-i}(N)$ that is an inverse to $g\colon \bigoplus_{i=0}^k \Lambda^i(M)\otimes_R \Lambda^{k-i}(N) \to \Lambda^k(M\oplus N)$. The map f has components $f_0,\ldots,f_k$: $f(\omega) = (f_i(\omega))$ for linear maps $f_i\colon \Lambda^k(M\oplus N)\to \Lambda^i(M)\otimes_R \Lambda^{k-i}(N)$.

The way we figure out a definition of $f_i\colon \Lambda^k(M\oplus N) \to \Lambda^i(M)\otimes_R \Lambda^{k-i}(N)$ on an elementary wedge product

$$\omega = (m_1,n_1)\wedge\cdots\wedge(m_k,n_k) \tag{9.16}$$

in $\Lambda^k(M\oplus N)$, is to write (m_j,n_j) as $(m_j,0)+(0,n_j)$ and then expand ω into a sum of 2^k elementary wedge products that each involve terms of the form $(m,0)$ and $(0,n)$. (The case $k=2$ is shown in (9.8).) We want $f_i(\omega)$ to be modeled by the sum of the elementary wedge products in the expansion of ω that have i factors of the form $(m,0)$ followed by $k-i$ factors of the form $(0,n)$ (rearranging to put each $(0,n)$ after each $(m,0)$ can introduce signs, as shown with one term in (9.8)). We'll make $f_i(\omega)$ a sum of $\binom{k}{i}$ terms in $\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)$ by turning $(m_1,0)\wedge\cdots(m_i,0)\wedge(0,n_{i+1})\wedge\cdots\wedge(0,n_k)$ into $(m_1\wedge\cdots\wedge m_i)\otimes(n_{i+1}\wedge\cdots\wedge n_k)$, with the edge cases $i=0$ and $i=k$ treated separately.

We can't really *define* a linear mapping on $\Lambda^k(M\oplus N)$ by saying where it sends elementary wedge products, due to the massive redundancies in writing elements of $\Lambda^k(M\oplus N)$ as a sum of elementary wedge products. To define a linear map $\Lambda^k(M\oplus N)\to\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)$, we will start by defining a k-multilinear map $(M\oplus N)^k\to\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)$ and check it is alternating. Then it automatically leads to a linear map $\Lambda^k(M\oplus N)\to\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)$ by the universal mapping property of exterior powers.

Break up $\{1,2,\ldots,k\}$ into a disjoint union $T\cup T'$ where $|T|=i$ and $|T'|=k-i$. This can be done in $\binom{k}{i}$ ways. Write elements of T in increasing order as $j_1<j_2<\ldots<j_i$ and elements of T' in increasing order as $j'_1<j'_2<\ldots<j'_{k-i}$. Using the notation from the right side of (9.16), define

$$m_T^\wedge := (m_{j_1},0)\wedge\cdots\wedge(m_{j_i},0)\in\Lambda^i(M\oplus N),\quad n_{T'}^\wedge := (0,n_{j'_1})\wedge\cdots\wedge(0,n_{j'_{k-i}})\in\Lambda^{k-i}(M\oplus N).$$

Define $\varepsilon_{T,T'}=\pm1$ to be the sign of the permutation sending the ordered set $\{1,2,\ldots,k\}$ to the set T put in increasing order followed by T' put in increasing order. (For example, if $k=4$, $T=\{1,3\}$, and $T'=\{2,4\}$, then $\varepsilon_{T,T'}=\operatorname{sign}\binom{1234}{1324}=-1$.) Symbolically, we could say $\varepsilon_{T,T'}=\operatorname{sgn}\binom{12\cdots k}{T\ T'}$ where T and T' are each written out in increasing order. If $i=0$ or $i=k$ then T or T' is empty and the other is $\{1,\ldots,k\}$; here we define $\varepsilon_{T,T'}=1$.

How ω in (9.16) looks when expanded out into a sum of 2^k elementary wedge products is

$$\omega = \sum_{i=0}^k \sum_{\substack{T\subset\{1,\ldots,k\}\\ |T|=i}} \varepsilon_{T,T'}\, m_T^\wedge\wedge n_{T'}^\wedge. \tag{9.17}$$

For each i, the inner sum in (9.17) over all T with $|T|=i$ is a sum of k-fold elementary wedge products with i parts of the form $(m,0)$ and $k-i$ parts of the form $(0,n)$ The ith inner sum will lead to $f_i(\omega)$ by mapping the ith inner sum to $\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)$. To achieve that, in the definitions of $m_T^\wedge$ and $n_{T'}^\wedge$ above replace $(m,0)$ with m and $(0,n)$ with n: define $\varphi_i\colon (M\oplus N)^k\to\Lambda^i(M)\otimes_R\Lambda^{k-i}(N)$ by

$$\varphi_i((m_1,n_1),\ldots,(m_k,n_k)) = \sum_{\substack{T\subset\{1,\ldots,k\}\\ |T|=i}} \varepsilon_{T,T'}\, m_T\otimes n_{T'}, \tag{9.18}$$

where $T = \{j_1 < \ldots < j_i\}$ and $T' = \{j'_1 < \cdots j'_{k-i}\}$ in increasing order and

$$(9.19) \qquad m_T := m_{j_1} \wedge \cdots \wedge m_{j_i} \in \Lambda^i(M), \quad n_{T'} := n_{j'_1} \wedge \cdots \wedge n_{j'_{k-i}} \in \Lambda^{k-i}(N).$$

Because m_T and $n_{T'}$ are each multilinear in the pieces that make them up, φ_i is a k-multilinear mapping (it is important here that T and T' have union $\{1, 2, \ldots, k\}$ and are disjoint). For example, say $k = 4$, $T = \{1, 3\}$, and $T' = \{2, 4\}$. We have $\varepsilon_{T,T'} = -1$, and if $(m_1, n_1) = (m'_1, n'_1) + (m''_1, n''_1)$, then

$$\begin{aligned} -m_T \otimes n_{T'} &= -((m'_1 + m''_1) \wedge m_3) \otimes (n_2 \wedge n_4) \\ &= -(m'_1 \wedge m_3) \otimes (n_2 \wedge n_4) - (m''_1 \wedge m_3) \otimes (n_2 \wedge n_4) \end{aligned}$$

and these two terms are what we get by using (m'_1, n'_1) and (m''_1, n''_1) separately in place of (m_1, n_1) in $-m_T \otimes n_{T'}$. The expression $\varepsilon_{T,T'} m_T \otimes n_T$ in general is additive in (m_1, n_1) because for each T either $1 \in T$ or $1 \in T'$ but not both, so either m_T contains m_1 as a wedge factor or $n_{T'}$ contains n_1 as a wedge factor but not both. Thus, if $(m_1, n_1) = (m'_1, n'_1) + (m''_1, n''_1)$ then either m_T contains $m'_1 + m''_1$ or $n_{T'}$ contains $n'_1 + n''_1$ but not both, so $m_T \otimes n_{T'}$ breaks up into a sum of two terms (not 4 mixed terms). The sign $\varepsilon_{T,T'}$ is carried along through the computation without changing. Thus φ_i is R-multilinear.

We will now show φ_i is an alternating mapping: if $(m_\alpha, n_\alpha) = (m_\beta, n_\beta)$ for distinct indices α and β in $\{1, \ldots, k\}$ then the total sum in (9.18) is 0. In fact many individual terms in the sum are 0 and the other terms can be paired together and are negatives.

Case 1: α and β are in T. Here $m_T = 0$ (an elementary wedge product with two equal parts), so $m_T \otimes n_{T'} = 0$.

Case 2: α and β are in T'. Here $n_{T'} = 0$, so $m_T \otimes n_{T'} = 0$.

Case 3: one of α and β is in T and the other is in T'. Let's consider together $\varepsilon_{T,T'}\, m_T \otimes n_{T'}$ and $\varepsilon_{\widetilde{T},\widetilde{T}'}\, m_{\widetilde{T}} \otimes n_{\widetilde{T}'}$, where $\widetilde{T}$ denotes T with α or β swapped out and the other index swapped in (and likewise for $\widetilde{T}'$ and T'). Since $m_\alpha = m_\beta$ and $n_\alpha = n_\beta$, $m_T \otimes n_{T'} = m_{\widetilde{T}} \otimes n_{\widetilde{T}'}$. But the signs $\varepsilon_{T,T'}$ and $\varepsilon_{\widetilde{T},\widetilde{T}'}$ are opposite since the permutations defining them differ by the transposition $(\alpha\beta)$. Thus the sum $\varepsilon_{T,T'}\, m_T \otimes n_{T'} + \varepsilon_{\widetilde{T},\widetilde{T}'}\, m_{\widetilde{T}} \otimes n_{\widetilde{T}}$ is 0.

Having shown $\varphi_i \colon (M \oplus N)^k \to \Lambda^i(M) \otimes_R \Lambda^{k-i}(N)$ to be an alternating k-multilinear mapping, we automatically get a linear mapping $f_i \colon \Lambda^k(M \oplus N) \to \Lambda^i(M) \otimes_R \Lambda^{k-i}(N)$ where $f_i((m_1, n_1) \wedge \cdots \wedge (m_k, n_k))$ equals the right side of (9.18):

$$(9.20) \qquad f_i((m_1, n_1) \wedge \cdots \wedge (m_k, n_k)) = \sum_{\substack{T \subset \{1, \ldots, k\} \\ |T| = i}} \varepsilon_{T,T'}\, m_T \otimes n_{T'}.$$

Define $f \colon \Lambda^k(M \oplus N) \to \bigoplus_{i=0}^k \Lambda^i(M) \otimes_R \Lambda^{k-i}(N)$ by $f(\omega) = (f_i(\omega))$ for ω in $\Lambda^k(M \oplus N)$.

<u>Step 4</u>: Show $f \circ g$ and $g \circ f$ are identity maps.

We want to show $f(g(\omega_0, \ldots, \omega_k)) = (\omega_0, \ldots, \omega_k)$ for all $\omega_i \in \Lambda^i(M) \otimes_R \Lambda^{k-i}(N)$ and $g(f(\omega)) = \omega$ for all $\omega \in \Lambda^k(M \oplus N)$.

<u>$f(g(\omega_0, \ldots, \omega_k)) = (\omega_0, \ldots, \omega_k)$ for all ω_i in $\Lambda^i(M) \otimes_R \Lambda^{k-i}(N)$</u>. In this equation both sides are k-multilinear, so it suffices to check it when all but one of $\omega_0, \ldots, \omega_k$ is 0. For i from 0 to k, suppose $\omega_j = 0$ for $j \neq i$: we want to show $f(g(\ldots, 0, \omega_i, 0, \ldots)) = (\ldots, 0, \omega_i, 0, \ldots)$ where all coordinates but the ith one are 0. Moreover, both sides of this new identity are additive in ω_i, so it suffices to check this new identity when

$$\omega_i := (m_1 \wedge \cdots \wedge m_i) \otimes (n_{i+1} \wedge \cdots \wedge n_k)$$

because each element of $\Lambda^i(M)\otimes_R \Lambda^{k-i}(N)$ is a sum of these particular types of elementary tensors.

By the definition of g as a sum of values of g_j on $\Lambda^j(M)\otimes_R \Lambda^{k-j}(N)$,

$$g(\ldots,0,\omega_i,0,\ldots)) = \sum_{j\neq i} g_j(0) + g_i(\omega_i) = g_i((m_1\wedge\cdots\wedge m_i)\otimes(n_{i+1}\wedge\cdots\wedge n_k)),$$

and from the end of Step 2, this g_i-value is $\omega' := (m_1,0)\wedge\cdots\wedge(m_i,0)\wedge(0,n_{i+1})\wedge\cdots\wedge(0,n_k)$. Thus

$$f(g(\ldots,0,\omega_i,0,\ldots)) = f(\omega') = (f_0(\omega'),\ldots,f_k(\omega')).$$

The value of $f_i(\omega')$ is given by (9.20) when $n_1=\cdots=n_i=0$ and $m_{i+1}=\cdots=m_k=0$. When forming m_T and n_T from that k-tuple, one of them is 0 (because some elementary wedge factor in m_T or $n_{T'}$ is 0), so $m_T\otimes n_{T'}=0$, *except* when $T=\{1,\ldots,i\}$ and $T'=\{i+1,\ldots,k\}$. Then $\varepsilon_{T,T'}=1$, $m_T=m_1\wedge\cdots\wedge m_i$, and $n_{T'}=n_{i+1}\wedge\cdots\wedge n_k$, so

$$f_i(\omega') = m_T\otimes n_{T'} = (m_1\wedge\cdots\wedge m_i)\otimes(n_{i+1}\wedge\cdots\wedge n_k) = \omega_i.$$

$\underline{g(f(\omega))=\omega \text{ for all } \omega\in\Lambda^k(M\oplus N)}$. In this equation both sides are additive in ω, so it suffices to check this result on ω that belong to a subset of $\Lambda^k(M\oplus N)$ that spans the module by adding and subtracting. We can use for this purpose the special k-fold elementary wedge products consisting of terms $(m,0)$ followed by terms $(0,n)$, say

$$\omega := (m_1,0)\wedge\cdots\wedge(m_i,0)\wedge(0,n_{i+1})\wedge\cdots\wedge(0,n_k)$$

for some i. Because each elementary wedge factor in ω is a term in $M\oplus N$ with 0 in one coordinate for each pair, the sum in (9.20) has all but one term definitely equal to 0, so

$$f_i(\omega) = (m_1\wedge\cdots\wedge m_i)\otimes(n_{i+1}\wedge\cdots\wedge n_k)$$

and

$$f_j(\omega)=0$$

for $j\neq i$. Thus $f(\omega)=(\ldots,0,f_i(\omega),0,\ldots)$ is concentrated in the ith term of the direct sum of all $\Lambda^j(M)\otimes_R\Lambda^{k-j}(N)$. Therefore

$$g(f(\omega)) = g(\ldots,0,f_i(\omega),0,\ldots) = g_i(f_i(\omega)) = g_i((m_1\wedge\cdots\wedge m_i)\otimes(n_{i+1}\wedge\cdots\wedge n_k))$$

and this g_i-value is $(m_1,0)\wedge\cdots\wedge(m_i,0)\wedge(0,n_{i+1})\wedge\cdots\wedge(0,n_k)$ by the end of Step 2. We have returned to ω, so $g(f(\omega))=\omega$. □

10. The exterior algebra

Since the wedge product operation $\wedge$ from Section 8 moves elements of an exterior power into a higher-degree exterior power, we can view $\wedge$ as multiplication in a noncommutative ring by taking the direct sum of all the exterior powers of a module.

Definition 10.1. For an R-module M, its *exterior algebra* is the direct sum

$$\Lambda(M) = \bigoplus_{k\geq 0}\Lambda^k(M) = R\oplus M\oplus\Lambda^2(M)\oplus\Lambda^3(M)\oplus\cdots,$$

provided with the multiplication rule given by the wedge product from Theorem 8.2, extended distributively to the whole direct sum.

Each $\Lambda^k(M)$ is only an R-module, but their direct sum $\Lambda(M)$ is an R-algebra: it has a multiplication that commutes with scaling by R and has identity $(1, 0, 0, \dots)$. Notice R is a subring of $\Lambda(M)$, embedded as $r \mapsto (r, 0, 0, \dots)$.

When $\Lambda^k(M)$ is viewed in the exterior algebra $\Lambda(M)$, it is called a homogeneous part and its elements are said to be the homogeneous terms of *degree* k in $\Lambda(M)$. For instance, when m_0, m_1, m_2, m_3, and m_4 are in M the sum $m_0 + m_1 \wedge m_2 + m_3 \wedge m_4$ in $\Lambda(M)$ has homogeneous parts m_0 and $m_1 \wedge m_2 + m_3 \wedge m_4$. This is the same terminology used for homogeneous multivariable polynomials, *e.g.*, $X^2 - XY + Y^3$ has homogeneous parts $X^2 - XY$ (of degree 2) and Y^3 (of degree 3).

The wedge product on the exterior algebra $\Lambda(M)$ is bilinear, associative, and distributive, but not commutative (unless $-1 = 1$ in R). The replacement for commutativity in Theorem 8.7 does *not* generalize to a rule between all elements of $\Lambda(M)$. However, if $\omega \in \Lambda^k(M)$ and k is even then ω commutes with every element of $\Lambda(M)$, so more generally the submodule $\bigoplus_{k \text{ even}} \Lambda^k(M)$ lies in the center of $\Lambda(M)$.

A typical element of $\Lambda(M)$ is a sequence $(\omega_k)_{k \geq 0}$ with $\omega_k = 0$ for $k \gg 0$, and we write it as a formal sum $\sum_{k \geq 0} \omega_k$ while keeping the direct sum aspect in mind. In this notation, the wedge product of two elements of $\Lambda(M)$ is

$$\sum_{k \geq 0} \omega_k \wedge \sum_{\ell \geq 0} \eta_\ell = \sum_{p \geq 0} \left(\sum_{k + \ell = p} \omega_k \wedge \eta_\ell \right),$$

where the inner sum on the right is actual addition in $\Lambda^p(M)$ and the outer sum is purely formal (corresponding to the direct sum decomposition defining $\Lambda(M)$). This extension of the wedge product from operations $\Lambda^k(M) \times \Lambda^\ell(M) \to \Lambda^{k+\ell}(M)$ on different exterior power modules to a single operation $\Lambda(M) \times \Lambda(M) \to \Lambda(M)$ is analogous to the way the multiplication rule $(aT^i)(bT^j) = abT^{i+j}$ on monomials, which is associative, can be extended to the usual multiplication between any two polynomials in $R[T] = \bigoplus_{i \geq 0} RT^i$. When M is a finitely generated R-module with n generators, every $\Lambda^k(M)$ is finitely generated as an R-module and $\Lambda^k(M) = 0$ for $k > n$ so $\Lambda(M) = \bigoplus_{k=0}^n \Lambda^k(M)$ is finitely generated as an R-module. Because $\Lambda^k(M)$ is spanned as an R-module by the elementary wedge products, and an elementary wedge product is a wedge product of elements of M, $\Lambda(M)$ is generated as an R-algebra (not as an R-module!) by M.

Theorem 10.2. *When M is a finite free R-module of rank $d \geq 0$, its exterior algebra is a free R-module of rank 2^d.*

Proof. By Theorem 4.2, each exterior power module $\Lambda^k(M)$ is free of rank $\binom{d}{k}$ for $0 \leq k \leq d$ and $\Lambda^k(M) = 0$ for $k > d$, so their direct sum $\Lambda(M)$ is free with rank

$$\sum_{k=0}^{d} \binom{d}{k} = 2^d. \qquad \square$$

Concretely, when M is free with basis $e_1, \dots, e_d$, we can think of $\Lambda(M)$ as an R-algebra generated by the e_i's subject to the relations $e_i^2 = 0$ and $e_i e_j = -e_j e_i$ for $i \neq j$. The construction of $\Lambda(M)$ did not use a basis, but this explicit description when there is a basis is helpful when doing computations.

Example 10.3. Let V be a real vector space of dimension 3 with basis e_1, e_2, e_3. Then $\Lambda^0(V) = \mathbf{R}$, $\Lambda^1(V) = V = \mathbf{R}e_1 \oplus \mathbf{R}e_2 \oplus \mathbf{R}e_3$, $\Lambda^2(V) = \mathbf{R}(e_1 \wedge e_2) \oplus \mathbf{R}(e_1 \wedge e_3) \oplus \mathbf{R}(e_2 \wedge e_3)$, and $\Lambda^3(V) = \mathbf{R}(e_1 \wedge e_2 \wedge e_3)$. The exterior algebra $\Lambda(V)$ is the direct sum of these vector spaces and we can count the dimension as $1 + 3 + 3 + 1 = 8$.

To get practice computing in an exterior algebra, we ask which elements of the exterior algebra of a vector space wedge a given vector to 0.

Theorem 10.4. *Let V be a vector space over the field K. For nonzero $v \in V$ and $\omega \in \Lambda(V)$, $v \wedge \omega = 0$ if and only if $\omega = v \wedge \eta$ for some $\eta \in \Lambda(V)$.*

Proof. The reduction to the case of finite-dimensional V proceeds as in the reduction step of the proof of Theorem 8.11. By associativity, if $\omega = v \wedge \eta$ then $v \wedge \omega = v \wedge (v \wedge \eta) = (v \wedge v) \wedge \eta = 0$. The point of the theorem is that the converse direction holds. First we reduce to the finite-dimensional case. Let W be the span of v and all the nonzero elementary wedge products in an expression for ω. Since $W \hookrightarrow V$, $\Lambda^k(W) \hookrightarrow \Lambda^k(V)$, so $\Lambda(W) \hookrightarrow \Lambda(V)$. From these embeddings, it suffices to prove the theorem with V replaced by W, so we may assume V is finite-dimensional.

Write $\omega = \sum_{k=0}^{n} \omega_k$ where $\omega_k \in \Lambda^k(V)$ is the degree k part of ω. (If you think about direct sums as sequences, $\omega = (\omega_0, \ldots, \omega_n)$.) Then

$$v \wedge \omega = \sum_{k=0}^{n} v \wedge \omega_k.$$

Since $v \wedge \omega_k \in \Lambda^{k+1}(V)$, the terms in the sum are in different homogeneous parts of $\Lambda(V)$, so the vanishing of $v \wedge \omega$ implies $v \wedge \omega_k = 0$ for each k. By Theorem 8.11, $\omega_0 = 0$ and for $k \geq 1$ we have $\omega_k = v \wedge \eta_{k-1}$ for some $\eta_{k-1} \in \Lambda^{k-1}(V)$. Thus $\omega = v \wedge \eta$ where $\eta = \sum_{k=1}^{n} \eta_{k-1} \in \Lambda(V)$. $\square$

While we created the exterior algebra $\Lambda(M)$ as a direct sum of R-modules with a snazzy multiplicative structure, it can be characterized on its own terms among R-algebras by a universal mapping property. Since each $\Lambda^k(M)$ is spanned as an R-module by the elementary wedge products $m_1 \wedge \cdots \wedge m_k$, $\Lambda(M)$ is generated as an R-algebra (using wedge multiplication) by M. Moreover, $m \wedge m = 0$ in $\Lambda(M)$ for all $m \in M$. We now turn this into a universal mapping property.

Theorem 10.5. *Let A be any R-algebra and suppose there is an R-linear map $L\colon M \to A$ such that $L(m)^2 = 0$ for all $m \in M$. Then there is a unique extension of L to an R-algebra map $\widetilde{L}\colon \Lambda(M) \to A$. That is, there is a unique R-algebra map $\widetilde{L}$ making the diagram*

$$\begin{tikzcd}
\Lambda(M) \arrow[dr, dotted, "\widetilde{L}"] & \\
M \arrow[u, hook] \arrow[r, "L"'] & A
\end{tikzcd}$$

commute.

This theorem is saying that any *R-linear* map from M to an R-algebra A such that the image elements square to 0 can always be extended uniquely to an *R-algebra* map from $\Lambda(M)$ to A. Such a universal property determines $\Lambda(M)$ up to R-algebra isomorphism by the usual argument.

Proof. Since M generates $\Lambda(M)$ as an R-algebra, there is at most one R-algebra map $\Lambda(M) \to A$ whose values on M are given by L. The whole problem is to construct such a map.

For any m and m' in M, $L(m)^2 = 0$, $L(m')^2 = 0$, and $L(m+m')^2 = 0$. Expanding $(L(m)+L(m'))^2$ and removing the squared terms leaves

$$0 = L(m)L(m') + L(m')L(m),$$

which shows $L(m)L(m') = -L(m')L(m)$. Therefore any product of L-values on M can be permuted at the cost of an overall sign change. This implies $L(m_1)L(m_2)\cdots L(m_k) = 0$ if two m_i's are equal, since we can permute the terms to bring them together and then use the vanishing of $L(m_i)^2$. This will be used later.

If there is going to be an R-algebra map $f\colon \Lambda(M) \to A$ extending L, then on an elementary wedge product we must have

$$f(m_1 \wedge \cdots \wedge m_k) = f(m_1)\cdots f(m_k) = L(m_1)\cdots L(m_k).$$

since $\wedge$ is the multiplication in $\Lambda(M)$. To show there is such a map, we start on the level of the $\Lambda^k(M)$'s. For $k \geq 0$, let $M^k \to A$ by $(m_1,\ldots,m_k) \mapsto L(m_1)\cdots L(m_k)$. This is multilinear since L is R-linear. It is alternating because $L(m_1)\cdots L(m_k) = 0$ when two m_i's are equal. Hence we obtain an R-linear map $f_k\colon \Lambda^k(M) \to A$ satisfying

$$f_k(m_1 \wedge \cdots \wedge m_k) = L(m_1)\cdots L(m_k)$$

for all elementary wedge products $m_1 \wedge \cdots \wedge m_k$. Define $f\colon \Lambda(M) \to A$ by letting it be f_k on $\Lambda^k(M)$ and extending to the direct sum $\Lambda(M) = \bigoplus_{k\geq 0} \Lambda^k(M)$ by additivity. This function f is R-linear and it is left to the reader to show f is multiplicative. $\square$

Notice the individual $\Lambda^k(M)$'s don't appear in the statement of Theorem 10.5. This theorem describes an intrinsic feature of the full exterior algebra as an R-algebra.

It turns out that the exterior algebra $\Lambda(M)$, *unlike* the individual exterior power modules $\Lambda^k(M)$, behaves very nicely for direct sums: $\Lambda(M \oplus N) \cong \Lambda(M) \otimes_R \Lambda(N)$ as R-algebras. This can be proved without knowing first how exterior powers behave on direct sums, and then by looking at the degree-k pieces on both sides of that isomorphism of R-algebras, we get the Künneth formula for $\Lambda^k(M \oplus N)$.

Appendix A. Linear maps preserving wedge products

Let V be a vector space and $A\colon V \to V$ be a linear map such that $Av \wedge Aw = v \wedge w$ for all v and w in V. When $\dim V = 1$ this is not interesting, since all wedge products on a one-dimensional space are 0. When $\dim V = 2$, $Av \wedge Aw = (\det A)(v \wedge w)$, so taking v and w to be linearly independent implies $\det A$ has to be 1, and conversely all linear maps $A\colon V \to V$ where $\det A = 1$ satisfy $Av \wedge Aw = (\det A)(v \wedge w) = v \wedge w$ for all v and w in V. What happens when $\dim V \geq 3$?[9]

Lemma A.1. *If $\dim V \geq 3$, v and w are linearly independent in V, and $A\colon V \to V$ is linear with $Av \wedge Aw = v \wedge w$, then Av and Aw are in the span of v and w.*

Proof. By the symmetry in v and w, it suffices to show Av is in the span of v and w.

Step 1: If u, v, w are linearly independent in V and there is an $x \in V$ such that $u \wedge x$ and $v \wedge w$ linearly dependent, then $u \wedge x = 0$.

We will show x is a scalar multiple of u, which makes $u \wedge x = 0$.

Since v and w are linearly independent in V, $v \wedge w \neq 0$, so linear dependence of $u \wedge x$ and $v \wedge w$ implies $u \wedge x = c(v \wedge w)$ for some scalar c. Extend the linearly independent subset

[9] Our discussion of this question was inspired by https://math.stackexchange.com/questions/4723915/.

$\{u, v, w\}$ to a basis of the span of u, v, w, x. Call this basis $\{e_1, \ldots, e_d\}$, where $d = 3$ or 4, $e_1 = u$, $e_2 = v$, and $e_3 = w$. Write x in the basis $\{e_i\}$ as $\sum_{j=1}^{d} a_j e_j$. Then

$$u \wedge x = e_1 \wedge x = \sum_{j=2}^{d} a_j (e_1 \wedge e_j).$$

Also $u \wedge x = c(v \wedge w) = c(e_2 \wedge e_3)$, so

$$\sum_{j=2}^{d} a_j (e_1 \wedge e_j) = c(e_2 \wedge e_3). \tag{A.1}$$

The wedge products $\{e_i \wedge e_j\}_{i<j}$ in $\Lambda^2(V)$ are linearly independent, so (A.1) tells us $a_j = 0$ when $j = 2, \ldots, d$ (and $c = 0$). Thus $x = a_1 e_1 = a_1 u$, which is a scalar multiple of u, so $u \wedge x = 0$.

Step 2: If v and w are linearly independent in V, and $A \colon V \to V$ is linear with $Av \wedge Aw = v \wedge w$, then Av is in the span of v and w.

Suppose Av is not in the span of v and w. Then $\{v, w, Av\}$ is linearly independent, so the condition $Av \wedge Aw = v \wedge w$ implies $Av \wedge Aw = 0$ by Step 1. Then $v \wedge w = 0$, which contradicts the linear independence of v and w. Thus Av lies in the span of v and w. □

Theorem A.2. *When* $\dim V \geq 3$, *the only linear maps* $A \colon V \to V$ *satisfying* $Av \wedge Aw = v \wedge w$ *for all* v *and* w *in* V *are* $A = \pm \mathrm{id}_V$.

Proof. If $A = \pm \mathrm{id}_V$, then it's obvious that $Av \wedge Aw = v \wedge w$ for all v and w in V.

Now assume $Av \wedge Aw = v \wedge w$ for all v and w in V. To prove $A = \mathrm{id}_V$ we break up the argument into three steps.

Step 1: For each nonzero $v \in V$ there is a scalar a_v such that $Av = a_v v$.

Since $\dim V \geq 3$, there are w and w' in V such that $\{v, w, w'\}$ is linearly independent.

The condition $Av \wedge Aw = v \wedge w$ implies Av is in the span of v and w by Lemma A.1, while the condition $Av \wedge Aw' = v \wedge w'$ implies Av is in the span of v and w'. Thus

$$Av = av + bw, \quad Av = a'v + b'w'$$

for some scalars a, b, a', and b'. The linear independence of $\{v, w, w'\}$ implies $a = a'$ and $b = b' = 0$, so $Av = av$. The coefficient a may depend on v, so write a as a_v.

Step 2: All the scalars a_v are equal.

Pick nonzero v and w in V, so $Av = a_v v$ and $Aw = a_w w$. We want to show $a_v = a_w$.

If v and w are linearly dependent, say $w = cv$, then $Aw = cA(v) = ca_v v = a_v w$, so $a_w = a_v$.

If v and w are linearly independent, then

$$A(v + w) = a_{v+w}(v + w) = a_{v+w} v + a_{v+w} w \quad \text{and} \quad A(v + w) = Av + Aw = a_v v + a_w w.$$

By linear independence of v and w we have $a_{v+w} = a_v$ and $a_{v+w} = a_w$, so $a_v = a_w$.

Step 3: $A = \pm \mathrm{id}_V$.

By Step 2, there is a common scalar a such that $Av = av$ for all nonzero $v \in V$, so $A = a\,\mathrm{id}_V$. Pick two linearly independent v and w in V, so the condition $Av \wedge Aw = v \wedge w$ implies $a^2(v \wedge w) = v \wedge w$. Since $v \wedge w$ is nonzero, $a^2 = 1$, so $a = \pm 1$. Thus $A = a\,\mathrm{id}_V = \pm \mathrm{id}_V$. □

References

[1] W. C. Brown, "A Second Course in Linear Algebra," J. Wiley & Sons, New York, 1988.
[2] H. Flanders, *Tensor and Exterior Powers*, J. Algebra **7**, 1–24 (1967).
[3] R. Gardner, *Modules Admitting Determinants*, Linear and Multilinear Algebra **3** (1975/76), 209–214.
[4] P. Garrett, Exterior powers. URL http://www-users.math.umn.edu/~garrett/m/algebra/notes/28.pdf.
[5] S. Lang, "Algebra," revised 3rd edition, Springer-Verlag, New York, 2002.
[6] J. Rotman, "Advanced Modern Algebra," Prentice-Hall, Upper Saddle River, NJ, 2002.
[7] F. Warner, "Foundations of Differentiable Manifolds and Lie Groups," Springer-Verlag, New York, 1983.