1. Volume Functions

Determinants are a crucial concept in linear algebra and every other field relying on linear algebra. However, they are notoriously tricky to define in a way that gives intuition about what is going on. Our textbook attempts to avoid determinants as long as possible and at all costs. As a result, determinants are relegated to the last section, and the exposition in that section relies upon heavy theory developed throughout the main body of the book. Meanwhile, Axler goes to great lengths to avoid any mention of determinants in the earlier material, even where their use is both standard and helpful.

Instead, we are going to discuss determinants at this point, trying our best to motivate them by the concept of the volume (area) spanned by a pair of vectors from \( \mathbb{R}^2 \).

1.1. Motivation: volumes in \( \mathbb{R}^2 \). One way to think about determinants is that they are all about volumes. For instance, given vectors \( u, v \in \mathbb{R}^2 \), what is the volume of the parallelogram spanned by \( u \) and \( v \)? We could calculate a formula in various ways, but let us first consider some properties of the result as a function of \( u \) and \( v \).

Remark 1.1.1. Let \( V = \mathbb{R}^2 \) and \( \text{vol}: V \times V \to \mathbb{R} \) be the function whose value \( \text{vol}(u, v) \) is the area of the parallelogram spanned by \( u \) and \( v \). Then \( \text{vol} \) satisfies the following properties.

(a) First, \( \text{vol} \) is bilinear. This means that it is a function linear with respect to each individual argument separately:

\[
\text{vol}(u_1 + u_2, v) = \text{vol}(u_1, v) + \text{vol}(u_2, v)
\]

\[
\text{vol}(\lambda u, v) = \lambda \text{vol}(u, v)
\]

\[
\text{vol}(u, v_1 + v_2) = \text{vol}(u, v_1) + \text{vol}(u, v_2)
\]

\[
\text{vol}(u, \lambda v) = \lambda \text{vol}(u, v).
\]

(b) Second, we have \( \text{vol}(u, u) = 0 \) (the spanned area is zero) for any \( u \).

Actually, these statements only make sense for a “signed” volume function: for example, we may consider the volume to be negative for left-handed parallelograms. This, however, makes sense as a mean of indicating the orientation of \( u \) and \( v \) as the sides of the parallelogram additionally to its nonnegative area \( |\text{vol}(u, v)| \).

As we did with fields and vector spaces, we extract the most essential mathematical properties desired of an abstract volume function from a particular, tangible volume function (the area of the parallelogram built on any two given vectors) and use them to define an abstract volume function.

Definition 1.1.2. Let us say that a function \( \alpha: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \) is a volume function if it satisfies the following requirements:

1. bilinearity:

\[
\alpha(u + \lambda v, w) = \alpha(u, v) + \lambda \alpha(u, w),
\]

\[
\alpha(u, v + \mu w) = \alpha(u, v) + \mu \alpha(u, w)
\]

for all \( u, v, w \in \mathbb{R}^2 \) and \( \lambda, \mu \in \mathbb{R} \).
(2) alternation: $\alpha(u, u) = 0$ for all $u \in \mathbb{R}^2$.

With $\text{Vol}(\mathbb{R}^2)$ we will denote the set of all volume functions.

It turns out that these properties are almost sufficient to identify what any $\alpha \in \text{Vol}(\mathbb{R}^2)$ is. Specifically, we can always take another volume function and rescale it by a suitable constant (this makes sense: there is no standard choice of volume units). This can be formalized as follows.

**Proposition 1.1.3.** The set $\text{Vol}$ is a one-dimensional real vector space with respect to the pointwise operations of addition and multiplication by a scalar. Equivalently, there exists a nonzero volume function $\beta$ such that any other volume function $\alpha$ satisfies $\alpha = \kappa \beta$ with some $\kappa \in \mathbb{R}$.

**Proof.** Let $\alpha$ be a volume function. We first note that

$$\alpha(x, y) = -\alpha(y, x)$$

for any $x, y \in \mathbb{R}^2$. Indeed, this is because

$$\alpha(x + y, x + y) = \alpha(x + x, y) + \alpha(x, y + y)$$

$$= \alpha(x, x) + \alpha(y, x + y) + \alpha(y, y)$$

and $\alpha(x, x) = \alpha(y, y) = \alpha(x + y, x + y) = 0$ by the alternation requirement.

Further, let $e_1, e_2$ be the standard basis for $\mathbb{R}^2$. Then

$$\alpha(ae_1 + be_2, ce_1 + de_2) = c\alpha(ae_1 + be_2, e_1) + d\alpha(ae_1 + be_2, e_2)$$

$$= ac\alpha(e_1, e_1) + bc\alpha(e_2, e_1) + ad\alpha(e_1, e_2) + bd\alpha(e_2, e_2)$$

$$= (ad - bc)\alpha(e_1, e_2).$$

On the other hand, we can check directly that the function $\beta: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\beta(ae_1 + be_2, ce_1 + de_2) = ad - bc$$

for all $a, b, c, d \in \mathbb{R}$

is indeed a volume function. The above calculation shows that for any $\alpha \in \text{Vol}(\mathbb{R}^2)$, we have $\alpha = \kappa \beta$ with $\kappa = \alpha(e_1, e_2)$.

So, just from straightforward properties of volume functions, we have just recovered the standard formula for the determinant of a $2 \times 2$ matrix, which is indeed a volume function. We have not yet talked about determinants of matrices, but we will soon.

Our plan is to extend our definition of (signed) volume functions to higher dimensions and to identify an arbitrary volume function (find a formula) just like above.

Also, note that we never actually used $k = \mathbb{R}$ above: for any field, our construction retains its mathematical meaning, even though the term “volume” may lose its common meaning.

### 1.2. Volume functions in general.

**Definition 1.2.1.** Let $V$ be a vector space over $k$ of dimension $n \in \mathbb{N}$. A **volume function** on $V$ is a function

$$\alpha: V^n \to k$$

(which means, a function that takes $n$ arguments, each a vector from $V$) satisfying the following three conditions:

(i) multilinearity:

$$\alpha(v_1, \ldots, v_{i-1}, \lambda v_i + \mu w, v_{i+1}, \ldots, v_n)$$

$$= \lambda \alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + \mu \alpha(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n)$$

for all $v_1, \ldots, v_n, w \in V$ and $\lambda, \mu \in k$;
(ii) alternation:
\[ \alpha(v_1, \ldots, v_n) = 0 \]
whenever \( v_1, \ldots, v_n \in V \) are not distinct: there exist \( i, j \in \{1, \ldots, n\} \) such that \( i \neq j \) and \( v_i = v_j \).

Again, we define \( \text{Vol}(V) \) as the set of all volume functions on \( V \).

You can check that this is the natural generalization for \( \mathbb{R}^3 \): in the case of \( k = \mathbb{R} \) and \( n = 2 \), Definition 1.2.1 agrees with Definition 1.1.2.

**Remark 1.2.2.** Again, it is easy to check that linear combinations of volume functions are volume functions, and the zero function is a volume function. So \( \text{Vol}(V) \) is, by its nature, a vector space over \( k \). As before, our goal is to show it has dimension one, i.e., essentially (up tp rescaling) there is only one volume function. We would also like to have a formula for some non-zero volume function.

Let us first generalize the observation that any volume function is antisymmetric, made in the course of proving Proposition 1.1.3.

**Lemma 1.2.3.** Let \( V \) be a vector space of dimension \( n \in \mathbb{N} \) and \( v_1, \ldots, v_n \in V \). Assume that \( \alpha \) is a volume function on \( V \). Then, for any \( i, j \in \{1, \ldots, n\} \) such that \( j > i \),
\[
\alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) = -\alpha(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n).
\]
That is, swapping two vectors changes the sign of the volume.

**Proof.** Let us define \( \hat{\alpha} : V \times V \to k \) by setting,
\[
\hat{\alpha}(u, w) = \alpha(v_1, \ldots, v_{i-1}, u, v_i+1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n)
\]
for all \( u, w \in V \). Applying Definition 1.2.1 to \( \alpha \), we obtain that the function \( \hat{\alpha} \) is bilinear and alternating. In particular,
\[
\hat{\alpha}(v_i + v_j, v_i + v_j) = \hat{\alpha}(v_i, v_i + v_j) + \hat{\alpha}(v_j, v_i) = \hat{\alpha}(v_i, v_i) + \hat{\alpha}(v_i, v_j) + \hat{\alpha}(v_j, v_j) + \hat{\alpha}(v_j, v_j) = 0.
\]
and
\[
\hat{\alpha}(v_i + v_j, v_i + v_j) = \hat{\alpha}(v_i, v_i) = \hat{\alpha}(v_j, v_j) = 0.
\]
Combining these two facts, we obtain
\[
\hat{\alpha}(v_i, v_j) + \hat{\alpha}(v_j, v_i) = 0.
\]
Recalling the definition of \( \hat{\alpha} \), we find that this is the claimed equality. \( \square \)

**2. PERMUTATIONS**

The above result analyzes the situation when two arguments of a volume function are swapped. More generally, we would like to know what happens when we re-order the arguments of a volume function. That is, how
\[
\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(n)})
\]
is related to \( \alpha(v_1, \ldots, v_n) \) when
\[
\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}
\]
is a permutation, i.e. a bijection from a finite set to itself. In other words, \((\sigma(1), \ldots, \sigma(n))\) are the numbers \( \{1, \ldots, n\} \) in some arbitrary order.

It is intuitively clear that:
- we can rearrange the vectors by repeatedly swapping pairs of them;
- every time we do a swap, we pick up a factor of \((-1)\);
so, the answer is that the two expressions are equal if the number of swaps we need is even
and have opposite signs if it is odd.

Unfortunately, there is a difficulty with this reasoning: it is not so clear whether it is impossible to
construct the same permutation \((\sigma(1), \ldots, \sigma(n))\) in one way using an even number of swaps and in a
different way using an odd number of swaps. If that were possible, we would be in trouble: among
other things, there would be no nonzero volume functions. So, in the first place, we have to rule out
this possibility.

Our goal is to find some definition of the “sign” of a permutation, which measures whether it is
constructed from an odd or even number of swaps, that is obviously well-defined, i.e. doesn’t depend
on how the permutation is constructed from swaps. There are several ways to do this; here is one.

**Definition 2.0.1.** For any permutation \(\sigma\) of \(\{1, \ldots, n\}\), define
\[
\text{inv}(\sigma) = \# \{(i, j) \mid 1 \leq i < j \leq n \text{ and } \sigma(i) > \sigma(j)\},
\]
the **number of inversions** of \(\sigma\). We set
\[
\text{sign}(\sigma) = (-1)^{\text{inv}(\sigma)},
\]
the **sign** of \(\sigma\), which equals 1 if \(\text{inv}(\sigma)\) is even and \(-1\) if it is odd. Conventionally, we extend the defi-
nition of sign to all functions \(\{1, \ldots, n\} \to \{1, \ldots, n\}\) by setting \(\text{sign}(\sigma) = 0\) if \(\sigma\) is not a permutation.

We now show that this odd definition really does count the parity of the number of swaps.

**Proposition 2.0.2.** Let \(\sigma\) be a permutation and \(\tau\) be obtained from \(\sigma\) by swapping the values at
positions \(a\) and \(b\) (where \(1 \leq a < b \leq n\)):
\[
\tau(r) = \begin{cases} 
\sigma(b) & \text{if } r = a \\
\sigma(a) & \text{if } r = b \\
\sigma(r) & \text{otherwise}
\end{cases}
\]
for each \(r = 1, \ldots, n\). Then \(\text{sign}(\tau) = -\text{sign}(\sigma)\).

**Proof.** We first prove the claim in the case \(b = a + 1\), when we swap two adjacent values. We claim
\[
\text{inv}(\tau) = \begin{cases} 
\text{inv}(\sigma) + 1 & \text{if } \sigma(a) < \sigma(b) \\
\text{inv}(\sigma) - 1 & \text{if } \sigma(a) > \sigma(b)
\end{cases}
\]
Indeed, swapping \(a\) and \(a + 1\) leaves every \((i, j)\) pair in the definition of inv unaffected, except for
\((a, a + 1)\) itself, which changes from in-order to out-of-order or vice versa. In either case, it follows that
\(\text{sign}(\tau) = -\text{sign}(\sigma)\).

Now consider swapping any \(a < b\). We claim this can be performed by combining \(2(b - a) - 1\) adjacent swaps. First, we sweep the element at position \(a\) through positions \(a + 1, \ldots, b - 1\):

- swap \((a, a + 1)\);
- swap \((a + 1, a + 2)\);
- \ldots;
- swap \((b - 2, b - 1)\);
- swap \((b - 1, b)\).

These \(b - a\) swaps transform the permutation
\[
\sigma = (\sigma(1), \ldots, \sigma(a - 1), \sigma(a), \sigma(a + 1), \ldots, \sigma(b - 1), \sigma(b), \sigma(b + 1), \ldots, \sigma(n))
\]
into
\[
(\sigma(1), \ldots, \sigma(a - 1), \sigma(a + 1), \ldots, \sigma(b - 1), \sigma(b), \sigma(a), \sigma(b + 1), \ldots, \sigma(n)).
\]
Second, we sweep the element at position \(b - 1\) through positions \(b - 2, b - 3, \ldots, a + 1\):
• swap \((b - 2, b - 1)\);
• \(\ldots\);
• swap \((a + 1, a + 2)\);
• swap \((a, a + 1)\).

These \(b - a - 1\) swaps transform the permutation

\[
(\sigma(1), \ldots, \sigma(a - 1), \sigma(a + 1), \ldots, \sigma(b - 1), \sigma(b), \sigma(a), \sigma(b + 1), \ldots, \sigma(n))
\]

into

\[
(\sigma(1), \ldots, \sigma(a - 1), \sigma(b), \sigma(a + 1), \ldots, \sigma(b - 1), \sigma(a), \sigma(b + 1), \ldots, \sigma(n)) = \tau.
\]

So, the sign changes at each step (by the adjacent case, which we considered first), and there are an odd number \((2(b - a) - 1)\) of steps, so again \(\text{sign}(\tau) = -\text{sign}(\sigma)\).

\[\square\]

**Corollary 2.0.3.** Let \(V\) be a vector space of dimension \(n \in \mathbb{N}\), \(v_1, \ldots, v_n \in V\), \(\alpha\) be a volume function on \(V\) and \(\sigma: \{1, \ldots, n\} \to \{1, \ldots, n\}\) be arbitrary. Then

\[
\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \text{sign}(\sigma) \alpha(v_1, \ldots, v_n).
\]

**Proof.** If \(\sigma\) is not a permutation, it is not a bijection on \(\{1, \ldots, n\}\): there exist \(i, j \in \{1, \ldots, n\}\) such that \(i \neq j\) and \(\sigma(i) = \sigma(j)\). Then \(v_{\sigma(i)} = v_{\sigma(j)}\) for these distinct \(i\) and \(j\), so the left-hand side is zero by Definition 1.2.1. On the other hand, the right-hand side is zero by Definition 2.0.1.

Assume that \(\sigma\) is a permutation. Then we will firstly prove that \(\sigma\) can be constructed by repeated swaps. Let us consider \(k \in \{2, \ldots, n\}\) and assume that \(\sigma_k\) is a permutation of \(\{1, \ldots, n\}\) such that \(\sigma_k(j) = j\) for each \(j \in \{k + 1, \ldots, n\}\) (in particular, when \(k = n\) the permutation \(\sigma_k\) can be arbitrary).

For this permutation, we consider the unique \(i \in \{1, \ldots, k\}\) such that \(\sigma_k(i) = k\), i.e., \(i = \pi_k^{-1}(k)\). Since \(\sigma_k\) does not permute the positions \(k + 1, \ldots, n\), we can be sure that \(i \in \{1, \ldots, k\}\). When \(i \neq k\), the permutation \(\sigma_k\) can be written explicitly as follows:

\[
\sigma_k = (\sigma_k(1), \ldots, \sigma_k(i - 1), \underline{\sigma_k(i)}, \sigma_k(i + 1), \ldots, \sigma_k(k - 1), \underline{\sigma_k(k)}, k + 1, \ldots, n)
\]

Let us define \(\pi_k\) as the identity permutation id if \(i = k\) and as the permutation of \(\{1, \ldots, n\}\) that only swaps the elements \(k\) and \(i = \pi_k^{-1}(k)\) if \(i < k\). Let us also set \(\sigma_{k-1} = \pi_k \circ \sigma_k\). When \(i \neq k\), we have

\[
\sigma_{k-1} = \pi_k \circ \sigma_k = \pi_k(\sigma_k(1), \ldots, \sigma_k(i - 1), \underline{\sigma_k(i)}, \sigma_k(i + 1), \ldots, \sigma_k(k - 1), \underline{\sigma_k(k)}, k + 1, \ldots, n)
\]

\[
= (\sigma_k(1), \ldots, \sigma_k(i - 1), \underline{\sigma_k(i)}, \sigma_k(i + 1), \ldots, \sigma_k(k - 1), \underline{\sigma_k(i)}, k + 1, \ldots, n)
\]

where we underline the elements that are swapped. Note that, regardless of whether \(i = k\), we have \(\sigma_{k-1}(j) = j\) for each \(j \in \{k, \ldots, n\}\).

Applying this argument iteratively with \(k = n, n - 1, \ldots, 2\), we obtain that there exist pairwise-swap permutations \(\pi_n, \pi_{n-1}, \ldots, \pi_2\) such that the permutation \(\sigma_1 = \pi_2 \circ \ldots \circ \pi_n \circ \sigma_n\) does not permute positions 2 to \(n\), i.e., \(\sigma_1 = \text{id}\). The swaps are reversible, and the inverses of pairwise-swap permutations are also pairwise-swap permutations. This gives \(\sigma = \pi_n^{-1} \circ \ldots \circ \pi_2^{-1}\), a representation of \(\sigma\) in the form of a product of \(n - 1\) pairwise-swap permutations.

For the sake of brevity, let us call those of the pairwise swaps \(\pi_2, \ldots, \pi_n\) that are different from id **proper swaps**. Obviously, a swap is proper if and only if its inverse is a proper swap.

The permutation id has sign \(+1\), whereas every proper swap has sign \(-1\), as we showed in the proof of Proposition 2.0.2 above. This is equivalent to that, when applied to an arbitrary permutation, id does not change the sign, whereas every proper swap changes the sign, as we showed in Proposition 2.0.2. Indeed, to see the relation for a proper swap, it is sufficient to apply it to id.
So we have \( \text{sign}(\pi_k \circ \tau) = \text{sign}(\pi_k) \cdot \text{sign}(\tau) \) for any permutation \( \tau \) of \( \{1, \ldots, n\} \). Since the identity permutation \( \text{id} \) is obtained from \( \sigma \) by applying the swaps \( \pi_2, \ldots, \pi_n \), we conclude that \( \text{sign}(\sigma) = \text{sign}(\pi_2) \cdots \text{sign}(\pi_n) \).

Applying Proposition 1.2.3 for every proper swap, we conclude that

\[
\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \alpha(v_{\sigma(1)}, \ldots, v_{\sigma_n(n)}) = \text{sign}(\sigma) \alpha(v_{\sigma^{-1}(1)}, \ldots, v_{\sigma^{-1}_n(n)}) = \cdots = \text{sign}(\pi_2) \cdots \text{sign}(\pi_n) \alpha(v_{\sigma_1(1)}, \ldots, v_{\sigma_1(n)})
\]

\[
= \text{sign}(\pi_2) \cdots \text{sign}(\pi_n) \alpha(v_{\text{id}(1)}, \ldots, v_{\text{id}(n)}) = \text{sign}(\pi_2) \cdots \text{sign}(\pi_n) \alpha(v_1, \ldots, v_n)
\]

\[
= \text{sign}(\sigma) \alpha(v_1, \ldots, v_n).
\]

\[\square\]

The following result relates the volume spanned by \( n \) vectors, given in a basis, to the volume spanned by the \( n \) vectors of the basis. Naturally, this relation is given in terms of the coefficients of the former \( n \) vectors with respect to the basis.

**Proposition 2.0.4.** Let \( V \) be a vector space of dimension \( n \in \mathbb{N} \) over a field \( F \) and \( b_1, \ldots, b_n \in V \) be a basis for \( V \). Let \( v_1, \ldots, v_n \in V \) be given by

\[
v_j = \sum_{i=1}^{n} a_{ij} b_i
\]

in terms of coefficients \( a_{ij} \in F \) with \( i, j = 1, \ldots, n \). Then if \( \alpha \) is a volume function on \( V \), i.e., \( \alpha \in \text{Vol}(V) \), then

\[
\alpha(v_1, v_2, \ldots, v_n) = \alpha(b_1, b_2, \ldots, b_n) \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) \prod_{k=1}^{n} a_{\sigma(k)k}.
\]

In other words, the value of \( \alpha \) at any tuple of vectors is determined by its value at the basis \( b_1, b_2, \ldots, b_n \) and by the coefficients of the vectors with respect to the basis.

Conversely, the function \( \beta : V^n \to F \) given as follows is a volume function on \( V \):

\[
\beta(v_1, v_2, \ldots, v_n) = \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) \prod_{k=1}^{n} a_{\sigma(k)k}
\]

for all \( v_1, \ldots, v_n \in V \), where \( a_{ij} \in F \) with \( i, j = 1, \ldots, n \) are uniquely defined by (1).

**Proof.** Let \( \alpha \) be an arbitrary volume function on \( V \) and \( v_1, \ldots, v_n \) be given by (1) in terms of the basis \( b_1, \ldots, b_n \). Then

\[
\alpha(v_1, \ldots, v_n) = \alpha \left( \sum_{i=1}^{n} a_{i1} b_1, \ldots, \sum_{i=1}^{n} a_{in} b_n \right).
\]

Using linearity in each argument, we can take each summation and each scalar outside:

\[
\alpha(v_1, \ldots, v_n) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} \alpha(b_{i_1}, \ldots, b_{i_n}) \prod_{k=1}^{n} a_{i_kk}.
\]

Each \( (i_1, \ldots, i_n) \in \{1, \ldots, n\}^n \) corresponds to a unique \( \sigma \in \Sigma_n \), which is given by \( \sigma(1) = i_1, \ldots, \sigma(n) = i_n \). Then :

\[
\alpha(v_1, \ldots, v_n) = \sum_{\sigma \in \Sigma_n} \alpha(b_{\sigma(1)}, \ldots, b_{\sigma(n)}) \prod_{k=1}^{n} a_{\sigma(k)k} = \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) \alpha(b_1, \ldots, b_n) \prod_{k=1}^{n} a_{\sigma(k)k},
\]

which proves the first part of the claim.
Now we consider the function \( \beta : V^n \to F \) given by (2). For any \( v_1, \ldots, v_n \) given by (1) in terms of the basis \( b_1, \ldots, b_n \), every \( i \in \{1, \ldots, n\} \), \( v = \sum_{i=1}^{n} c_i b_i \) with \( c_1, \ldots, c_n \in F \) and any \( \lambda \in F \):

\[
\beta(v_1, \ldots, v_{i-1}, \lambda v_i + w, v_{i+1}, \ldots, v_n) = \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) \left( \lambda a_{\sigma(i)} + c_{\sigma(i)} \right) \prod_{k \in \{1, \ldots, n\} \setminus \{i\}} a_{\sigma(k)k}
\]

\[
= \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) \prod_{k \in \{1, \ldots, n\} \setminus \{i\}} a_{\sigma(k)k} + \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) c_{\sigma(i)} \prod_{k \in \{1, \ldots, n\} \setminus \{i\}} a_{\sigma(k)k}
\]

\[
= \lambda \beta(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + \beta(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n).
\]

This proves the multilinearity of \( \beta \).

Let \( v_1, \ldots, v_n \) be given by (1) in terms of the basis \( b_1, \ldots, b_n \) and suppose that \( v_i = v_j \) for some \( i, j \in \{1, \ldots, n\} \) such that \( i < j \). Then, for any \( \sigma \in \Sigma_n \), let \( S\sigma \in \Sigma_n \) be the function obtained by swapping the values at \( i \) and \( j \). This defines a bijection \( S : \Sigma_n \to \Sigma_n \). Then, using that \( \text{sign}(\sigma) = 0 \) for any \( \sigma \in \Sigma_n \) that is not a permutation of \( \{1, \ldots, n\} \) and, in particular, when \( \sigma(i) = \sigma(j) \), we obtain

\[
\beta(v_1, \ldots, v_n) = \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) \prod_{k=1}^{n} a_{\sigma(k)k}
\]

since each term is individually zero, where we used the fact that \( a_{k\ell} = a_{\ell k} \) for all \( k = 1, \ldots, n \), since the vectors \( v_i \) and \( v_j \) are equal. \( \square \)

We will now start reaping the benefits of this theory.

**Corollary 2.0.5.** Let \( V \) be a vector space of dimension \( n \in \mathbb{N} \) and \( \alpha \), a nonzero volume function on \( V \). Then any \( b_1, \ldots, b_n \in V \) form a basis for \( V \) if and only if \( \alpha(b_1, \ldots, b_n) \neq 0 \).

**Proof.** Assume that \( b_1, \ldots, b_n \) is a basis. Then, by Proposition 2.0.4, the function \( \beta \) given by (2) is a volume function on \( V \). Note that \( \beta(b_1, \ldots, b_n) = 1 \): indeed, for any permutation \( \sigma \in \Sigma_n \) different form the identity, every term of the right-hand side of (2) contains at least one zero factor, so

\[
\beta(b_1, b_2, \ldots, b_n) = 1 \cdot 1 = 1.
\]
By the same proposition, we have $\alpha = \kappa \beta$ with $\kappa = \alpha(b_1, \ldots, b_n) \in F$. Since $\alpha$ is not the zero function, we have $\kappa \neq 0$.

Conversely, suppose that $b_1, \ldots, b_n$ is not a basis. Then these vectors are linearly dependent so, for some $i \in \{1, \ldots, n\}$, we have

$$b_i = r_1 b_1 + \cdots + r_{i-1} b_{i-1}$$

with some coefficients $r_1, \ldots, r_{i-1} \in F$. Applying the linearity and alternation properties of volume functions, we obtain

$$\alpha(b_1, \ldots, b_n) = \alpha \left( b_1, \ldots, b_{i-1}, \sum_{j=1}^{i-1} r_j b_j, b_{i+1}, \ldots, b_n \right) = \sum_{j=1}^{i-1} r_j \alpha(b_1, \ldots, b_{i-1}, b_j, b_{i+1}, \ldots, b_n) = 0,$$

where each term of the latter sum is zero since the argument vectors are not distinct.

Corollary 2.0.5 shows that a volume function is, essentially, a multilinear indicator of linear dependence. This agrees with intuition in the case of $V = \mathbb{R}^n$ with $n = 2, 3$: the volume spanned by $n$ vectors is zero if and only if they are linearly dependent.

### 3. Determinants

Let us start off with the determinant of a matrix, which can be defined very concretely.

**Definition 3.0.1.** If $(A_{ij})_{i,j=1}^n$ is an $n \times n$ matrix over a field $k$, we define $\det A \in k$, the determinant of $A$, as follows:

$$\det A = \sum_{\sigma \in \Sigma_n} \text{sign}(\sigma) A_{\sigma(1)1} \cdots A_{\sigma(n)n}.$$

So Proposition 2.0.4 says that if $v_j = \sum_i A_{ij} w_i$ and $\alpha$ is a volume function then

$$\alpha(v_1, \ldots, v_n) = (\det A) \alpha(w_1, \ldots, w_n).$$

**Remark 3.0.2.** Since $\text{sign}(\sigma) = 0$ if $\sigma$ is not a permutation, all these sums over functions from $\Sigma_n$ can be replaced with sums over all permutations of $\{1, \ldots, n\}$.

Since matrices only represent linear maps, our main interest here is in the determinant of a linear map $V \to V$.

**Lemma 3.0.3.** Let $V$ be a vector space of dimension $n \in \mathbb{N}$ over a field $k$. Consider $\phi \in \mathcal{L}(V, V)$ and $\alpha \in \text{Vol}(V)$. Then the function $\Phi_\alpha : V^n \to k$ given by

$$\Phi_\alpha(v_1, \ldots, v_n) = \alpha(\phi(v_1), \ldots, \phi(v_n))$$

is again a volume function on $V$. Moreover, the map $\Phi : \text{Vol}(V) \to \text{Vol}(V)$ defined by $\Phi(\alpha) = \Phi_\alpha$ for all $\alpha \in \text{Vol}(V)$ is a linear map.

It follows that there exists $r \in k$ such that $\Phi_\alpha = r\alpha$ for all $\alpha \in \text{Vol}(V)$.

**Proof.** We verify that $\Phi_\alpha$ is a volume function. Indeed, for any $v_1, \ldots, v_n, w \in V$, $\lambda \in F$ and $i \in \{1, \ldots, n\}$, the multilinearity of $\alpha$ as a volume function and the linearity of $\phi$ as a linear map give

$$\Phi_\alpha(v_1, \ldots, v_{i-1}, \lambda v_i + w, v_{i+1}, \ldots, v_n) = \alpha(\phi(v_1), \ldots, \phi(v_{i-1}), \lambda \phi(v_i) + \phi(w), \phi(v_{i+1}), \ldots, \phi(v_n))$$

$$= \alpha(\phi(v_1), \ldots, \phi(v_{i-1}), \lambda \phi(v_i) + \phi(w), \phi(v_{i+1}), \ldots, \phi(v_n))$$

$$= \lambda \alpha(\phi(v_1), \ldots, \phi(v_{i-1}), \phi(v_i), \phi(v_{i+1}), \ldots, \phi(v_n)) + \alpha(\phi(v_1), \ldots, \phi(v_{i-1}), \phi(w), \phi(v_{i+1}), \ldots, \phi(v_n))$$

$$= \lambda \Phi_\alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + \Phi_\alpha(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n).$$
Finally, if $v_i = v_j$ for some $i \neq j$, then $\phi(v_i) = \phi(v_j)$ and so $\Phi_\alpha(v_1, \ldots, v_n)$ is zero due to the alternation property of $\alpha$ as a volume function.

The fact that $\Phi$ is linear is easy to check using its definition.

Finally, we recall that $\text{Vol}(V)$ is a one-dimensional vector space and that any linear map from a one-dimensional vector space to itself is just multiplication by a scalar. This proves the existence of a scalar $r \in F$ such that $\Phi(\alpha) = r\alpha$ in $\text{Vol}(V)$ for all $\alpha \in \text{Vol}(V)$. □

**Definition 3.0.4.** We call the value $r$ appearing in Lemma 3.0.3 the **determinant** of $\phi$, denoted with $\det \phi$. Equivalently,

$$\det \phi = \frac{\alpha(\phi(v_1), \ldots, \phi(v_n))}{\alpha(v_1, \ldots, v_n)}$$

for any non-zero volume function $\alpha$ and any basis $v_1, \ldots, v_n$ (this definition relies on the last claim of Lemma 3.0.3, which guarantees that the ratio does not depend on the basis).

In the next proposition we give some properties of the determinant of a linear map.

**Proposition 3.0.5** (Properties of the determinant of a linear map). Assume that $V$ is a vector space of dimension $n \in \mathbb{N}$ over a field $k$.

(i) Let $\phi, \psi \in \mathcal{L}(V, V)$. Then $\det(\psi \circ \phi) = (\det \psi)(\det \phi)$.

(ii) Let $\phi \in \mathcal{L}(V, V)$ and $B = v_1, \ldots, v_n$ be a basis for $V$. Then

$$\det \phi = \det \mathcal{M}(\phi, B, B).$$

(iii) Let $\phi \in \mathcal{L}(V, V)$ be a linear map. Then $\phi$ is invertible if and only if $\det \phi \neq 0$.

**Proof.** For (i), consider the maps $\Phi, \Psi, \Xi : \text{Vol}(V) \to \text{Vol}(V)$ defined at every $\alpha \in \text{Vol}(V)$ by their values $\Phi(\alpha) \in \text{Vol}(V)$, $\Psi(\alpha) \in \text{Vol}(V)$ and $\Xi(\alpha) \in \text{Vol}(V)$ that are the volume function given by the equalities

$$(\Phi(\alpha))(v_1, \ldots, v_n) = \alpha(\phi(v_1), \ldots, \phi(v_n)),$$

$$(\Psi(\alpha))(v_1, \ldots, v_n) = \alpha(\psi(v_1), \ldots, \psi(v_n)),$$

$$(\Xi(\alpha))(v_1, \ldots, v_n) = \alpha((\psi \circ \phi)(v_1), \ldots, (\psi \circ \phi)(v_n))$$

for all $v_1, \ldots, v_n \in V$ (breathe!). These maps are linear, and, moreover, we see immediately from these definitions that, for every $\alpha \in \text{Vol}(V)$ and for $\Phi(\alpha) \in \text{Vol}(V)$, we have

$$(\Phi(\Psi(\alpha)))(v_1, \ldots, v_n) = (\Psi(\alpha))(\phi(v_1), \ldots, \phi(v_n)) = \alpha(\psi(\phi(v_1)), \ldots, \psi(\phi(v_n)))$$

$$= \alpha((\psi \circ \phi)(v_1), \ldots, (\psi \circ \phi)(v_n)) = (\Xi(\alpha))(v_1, \ldots, v_n)$$

for all $v_1, \ldots, v_n \in V$. This means that $\Xi = \Phi \circ \Psi$. Since each of $\Phi$, $\Psi$ and $\Xi$ is just multiplication by the respective of the scalars $\det \phi$, $\det \psi$ and $\det(\psi \circ \phi)$ respectively, we must have $\det(\psi \circ \phi) = (\det \phi)(\det \psi)$ as required.

For (ii), we consider $M = \mathcal{M}(\phi, B, B) \in F^{n \times n}$, then

$$\phi(v_j) = \sum_{i=1}^n M_{ij} v_i.$$ 

Hence,

$$\alpha(\phi(v_1), \ldots, \phi(v_n)) = (\det M) \alpha(v_1, \ldots, v_n)$$

for any volume function $\alpha$. Hence, $\Phi(\alpha) = (\det M) \alpha$ for any $\alpha \in \text{Vol}(V)$. The result follows by the definition of the determinant of $\phi$.

Finally, for (iii), recall that $\phi$ is invertible if and only if it maps bases of $V$ into bases of $V$. Also, $w_1, \ldots, w_n$ is a basis if and only if $\alpha(w_1, \ldots, w_n) \neq 0$ for any non-zero volume function $\alpha$. 
So, if $v_1, \ldots, v_n$ is a basis and $\alpha$ is a nonzero volume function, then $\phi(v_1), \ldots, \phi(v_n)$ is a basis if and only if

$$\alpha(\phi(v_1), \ldots, \phi(v_n)) = (\det \phi) \alpha(v_1, \ldots, v_n) \neq 0.$$ 

Note that $\alpha(v_1, \ldots, v_n) \neq 0$, since $v_1, \ldots, v_n$ is a basis and $\alpha$ is a nonzero volume function, so $\phi(v_1), \ldots, \phi(v_n)$ is a basis if and only if $\det \phi \neq 0$, as required.