# Appendix A

# Multilinear algebra and index notation

## Contents



If linear algebra is the study of vector spaces and linear maps, then *multilinear* algebra is the study of tensor products and the natural generalizations of linear maps that arise from this construction. Such concepts are extremely useful in differential geometry but are essentially *algebraic* rather than geometric; we shall thus introduce them in this appendix using only algebraic notions. We'll see finally in §A.10 how to apply them to tangent spaces on manifolds and thus recover the usual formalism of tensor fields and differential forms. Along the way, we will explain the conventions of "upper" and "lower" index notation and the Einstein summation convention, which are standard among physicists but less familiar in general to mathematicians.

## A.1 Vector spaces and linear maps

We assume the reader is somewhat familiar with linear algebra, so at least most of this section should be review—its main purpose is to establish notation that is used in the rest of the notes, as well as to clarify the relationship between *real* and *complex* vector spaces.

Throughout this appendix, let $\mathbb{F}$ denote either of the fields $\mathbb{R}$ or $\mathbb{C}$; we will refer to elements of this field as *scalars*. Recall that a *vector space over* $\mathbb{F}$ (or simply a *real/complex vector space*) is a set $V$ together with two algebraic operations:

- (VECTOR ADDITION) $V \times V \to V : (v, w) \mapsto v + w$
- (SCALAR MULTIPLICATION) $\mathbb{F} \times V \to V : (\lambda, v) \mapsto \lambda v$

One should always keep in mind the standard examples $\mathbb{F}^n$ for $n \geq 0$; as we will recall in a moment, every finite dimensional vector space is isomorphic to one of these. The operations are required to satisfy the following properties:

- (ASSOCIATIVITY) $(u + v) + w = u + (v + w)$.
- (COMMUTATIVITY) $v + w = w + v$.
- (ADDITIVE IDENTITY) There exists a *zero vector* $0 \in V$ such that $0 + v = v$ for all $v \in V$.
- (ADDITIVE INVERSE) For each $v \in V$ there is an *inverse* element $-v \in V$ such that $v + (-v) = 0$. (This is of course abbreviated $v - v = 0$.)
- (DISTRIBUTIVITY) For scalar multiplication, $(\lambda + \mu)v = \lambda v + \mu v$ and $\lambda(v + w) = \lambda v + \lambda w$.
- (SCALAR ASSOCIATIVITY) $\lambda(\mu v) = (\lambda\mu)v$.
- (SCALAR IDENTITY) $1v = v$ for all $v \in V$.

Observe that every complex vector space can also be considered a real vector space, though the reverse is not true. That is, in a complex vector space, there is automatically a well defined notion of multiplication by *real* scalars, but in real vector spaces, one has no notion of “multiplication by $i$”. As is also discussed in Chapter 2, such a notion can sometimes (though not always) be defined as an extra piece of structure on a real vector space.

For two vector spaces $V$ and $W$ over the same field $\mathbb{F}$, a map

$$A : V \to W : v \mapsto Av$$

is called *linear* if it *respects* both vector addition and scalar multiplication, meaning it satisfies the relations $A(v + w) = Av + Aw$ and $A(\lambda v) = \lambda(Av)$

for all $v, w \in V$ and $\lambda \in \mathbb{F}$. Linear maps are also sometimes called *vector space homomorphisms*, and we therefore use the notation

$$\mathrm{Hom}(V, W) := \{A : V \to W \mid A \text{ is linear}\}.$$

The symbols $L(V, W)$ and $\mathcal{L}(V, W)$ are also quite common but are not used in these notes. When $\mathbb{F} = \mathbb{C}$, we may sometimes want to specify that we mean the set of *real* or *complex* linear maps by defining:

$$\begin{aligned} \mathrm{Hom}_{\mathbb{R}}(V, W) &:= \{A : V \to W \mid A \text{ is } \textit{real} \text{ linear}\} \\ \mathrm{Hom}_{\mathbb{C}}(V, W) &:= \mathrm{Hom}(V, W). \end{aligned}$$

The first definition treats both $V$ and $W$ as *real* vector spaces, reducing the set of scalars from $\mathbb{C}$ to $\mathbb{R}$. The distinction is that a real linear map on a complex vector space need not satisfy $A(\lambda v) = \lambda(Av)$ for all $\lambda \in \mathbb{C}$, but rather for $\lambda \in \mathbb{R}$. Thus every complex linear map is also real linear, but the reverse is not true: there are many more real linear maps in general. An example is the operation of *complex conjugation*

$$\mathbb{C} \to \mathbb{C} : x + iy \mapsto \overline{x + iy} = x - iy.$$

Indeed, we can consider $\mathbb{C}$ as a real vector space via the one-to-one correspondence

$$\mathbb{C} \to \mathbb{R}^2 : x + iy \mapsto (x, y).$$

Then the map $z \mapsto \bar{z}$ is equivalent to the linear map $(x, y) \mapsto (x, -y)$ on $\mathbb{R}^2$; it is therefore real linear, but it does not respect multiplication by complex scalars in general, e.g. $\overline{iz} \neq i\bar{z}$. It does however have another nice property that deserves a name: for two complex vector spaces $V$ and $W$, a map $A : V \to W$ is called *antilinear* (or *complex antilinear*) if it is real linear and also satisfies

$$A(iv) = -i(Av).$$

Equivalently, such maps satisfy $A(\lambda v) = \bar{\lambda} v$ for all $\lambda \in \mathbb{C}$. The canonical example is complex conjugation in $n$ dimensions:

$$\mathbb{C}^n \to \mathbb{C}^n : (z_1, \ldots, z_n) \mapsto (\bar{z}_1, \ldots, \bar{z}_n),$$

and one obtains many more examples by composing this conjugation with any complex linear map. We denote the set of complex antilinear maps from $V$ to $W$ by

$$\overline{\mathrm{Hom}}_{\mathbb{C}}(V, W).$$

When the domain and target space are the same, a linear map $V \to V$ is sometimes called a *vector space endomorphism*, and we therefore use the notation

$$\mathrm{End}(V) := \mathrm{Hom}(V, V),$$

with corresponding definitions for $\mathrm{End}_{\mathbb{R}}(V)$, $\mathrm{End}_{\mathbb{C}}(V)$ and $\overline{\mathrm{End}}_{\mathbb{C}}(V)$.

Observe that all these sets of linear maps are themselves also vector spaces in a natural way: simply define $(A+B)v := Av + Bv$ and $(\lambda A)v := \lambda(Av)$.

Given a vector space $V$, a *subspace* $V' \subset V$ is a subset which is *closed* under both vector addition and scalar multiplication, i.e. $v + w \in V'$ and $\lambda v \in V'$ for all $v, w \in V'$ and $\lambda \in \mathbb{F}$. Every linear map $A \in \mathrm{Hom}(V, W)$ gives rise to important subspaces of $V$ and $W$: the *kernel*

$$\ker A = \{v \in V \mid Av = 0\} \subset V$$

and *image*

$$\mathrm{im}\, A = \{w \in W \mid w = Av \text{ for some } v \in V\} \subset W.$$

We say that $A \in \mathrm{Hom}(V, W)$ is *injective* (or *one-to-one*) if $Av = Aw$ always implies $v = w$, and *surjective* (or *onto*) if every $w \in W$ can be written as $Av$ for some $v \in V$. It is useful to recall the basic algebraic fact that $A$ is injective if and only if its kernel is the trivial subspace $\{0\} \subset V$. (Prove it!)

An *isomorphism* between $V$ and $W$ is a linear map $A \in \mathrm{Hom}(V, W)$ that is both injective and surjective: in this case it is *invertible*, i.e. there is another map $A^{-1} \in \mathrm{Hom}(W, V)$ so that the compositions $A^{-1}A$ and $AA^{-1}$ are the identity map on $V$ and $W$ respectively. Two vector spaces are *isomorphic* if there exists an isomorphism between them. When $V = W$, isomorphisms $V \to V$ are also called *automorphisms*, and the space of these is denoted by

$$\mathrm{Aut}(V) = \{A \in \mathrm{End}(V) \mid A \text{ is invertible}\}.$$

This is not a vector space since the sum of two invertible maps need not be invertible. It is however a *group*, with the natural "multiplication" operation defined by composition of linear maps:

$$AB := A \circ B.$$

As a special case, for $V = \mathbb{F}^n$ one has the *general linear group* $\mathrm{GL}(n, \mathbb{F}) := \mathrm{Aut}(\mathbb{F}^n)$. This and its subgroups are discussed in some detail in Appendix B.

## A.2 Bases, indices and the summation convention

A *basis* of a vector space $V$ is a set of vectors $e_{(1)}, \ldots, e_{(n)} \in V$ such that every $v \in V$ can be expressed as

$$v = \sum_{j=1}^{n} c_j e_{(j)}$$

for some unique set of scalars $c_1, \dots, c_n \in \mathbb{F}$. If a basis of $n$ vectors exists, then the vector space $V$ is called *$n$-dimensional*. Observe that the map

$$\mathbb{F}^n \to V : (c_1, \dots, c_n) \mapsto \sum_{j=1}^{n} c_j e_{(j)}$$

is then an isomorphism, so every $n$-dimensional vector space over $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$. Not every vector space is $n$-dimensional for some $n \geq 0$: there are also *infinite dimensional* vector spaces, e.g. the set of continuous functions $f : [0,1] \to \mathbb{R}$, with addition and scalar multiplication defined by $(f+g)(x) := f(x)+g(x)$ and $(\lambda f)(x) := \lambda f(x)$. Such spaces are interesting, but beyond the scope of the present discussion: for the remainder of this appendix, we restrict attention to *finite dimensional* vector spaces.

It is time now to begin explaining the index notation that is ubiquitous in the physics literature and in more classical treatments of differential geometry. Given an $n$-dimensional vector space $V$ and a basis $e_{(1)}, \dots, e_{(n)} \in V$, any vector $v \in V$ can be written as

$$v = v^j e_{(j)}, \tag{A.1}$$

where the numbers $v^j \in \mathbb{F}$ for $j = 1, \dots, n$ are called the *components* of $v$, and there is an *implied summation*: one would write (A.1) more literally as

$$v = \sum_{j=1}^{n} v^j e_{(j)}.$$

The shorthand version we see in (A.1) makes use of the *Einstein summation convention*, in which a summation is implied whenever one sees a pair of matching *upper* and *lower* indices. Moreover, the choice of upper and lower is not arbitrary: we intentionally assigned a lower index to the basis vectors, so that the components could have an upper index. This is a matter of well established convention.

In physicists' terminology, a vector whose components are labelled with upper indices is called a *contravariant* vector; there are also *covariant* vectors, whose components have lower indices—these are in fact slightly different objects, the *dual vectors* to be discussed in §A.3.

Now that bases have entered the discussion, it becomes convenient to describe linear maps via *matrices*. In principle, this is the same thing as using basis vectors and components for the vector space $\mathrm{Hom}(V, W)$. Indeed, given bases $e_{(1)}, \dots, e_{(n)} \in V$ and $f_{(1)}, \dots, f_{(m)} \in W$, we obtain a natural basis

$$\{\mathbf{a}_{(i)}{}^{(j)}\}_{i=1,\dots,m}^{j=1,\dots,n}$$

of $\mathrm{Hom}(V, W)$ by defining $\mathbf{a}_{(i)}{}^{(j)}(e_{(j)}) = f_{(i)}$ and $\mathbf{a}_{(i)}{}^{(j)}(e_{(k)}) = 0$ for $k \neq j$. To see that this is a basis, note that for any $A \in \mathrm{Hom}(V, W)$, the fact that

$f_{(1)}, \dots, f_{(m)}$ is a basis of $W$ implies there exist unique scalars $A^i{}_j \in \mathbb{F}$ such that

$$Ae_{(j)} = A^i{}_j f_{(i)},$$

where again summation over $i$ is implied on the right hand side. Then for any $v = v^j e_{(j)} \in V$, we exploit the properties of linearity and find[1]

$$\begin{aligned}(A^i{}_j \mathbf{a}_{(i)}{}^{(j)})v &= (A^i{}_j \mathbf{a}_{(i)}{}^{(j)})v^k e_{(k)} = A^i{}_j v^k \mathbf{a}_{(i)}{}^{(j)} e_{(k)} \\ &= (A^i{}_j v^j) f_{(i)} = v^j A^i{}_j f_{(i)} = v^j A e_{(j)} = A(v^j e_{(j)}) = Av.\end{aligned} \tag{A.2}$$

Thus $A = A^i{}_j \mathbf{a}_{(i)}{}^{(j)}$, and we've also derived the standard formula for matrix-vector multiplication:

$$(Av)^i = A^i{}_j v^j.$$

**Exercise A.1.** If you're not yet comfortable with the summation convention, rewrite the derivation (A.2) including all the summation signs. Most terms should contain two or three; two of them contain only one, and only the last has none.

**Exercise A.2.** If $B : V \to X$ and $A : X \to W$ are linear maps and $(AB)^i{}_j$ are the components of the composition $AB : V \to W$, derive the standard formula for matrix-matrix multiplication:

$$(AB)^i{}_j = A^i{}_k B^k{}_j.$$

It should be emphasized at this point that our choice of upper and lower indices in the symbol $A^i{}_j$ is not arbitrary: the placement is selected specifically so that the Einstein summation convention can be applied, and it is tied up with the fact that $A$ is a linear map from one vector space to another. In the following we will see other matrices for which one uses either two upper or two lower indices—the reason is that such matrices play a different role algebraically, as something other than linear maps.

**Exercise A.3** (Change of basis)**.** If $e_{(1)}, \dots, e_{(n)}$ and $\hat{e}_{(1)}, \dots, \hat{e}_{(n)}$ are two bases of $V$, we can write each of the vectors $e_{(i)}$ as linear combinations of the $\hat{e}_{(j)}$: this means there are unique scalars $S^j{}_i$ for $i, j = 1, \dots, n$ such that $e_{(i)} = \hat{e}_{(j)} S^j{}_i$. Use this to derive the formula

$$\hat{v}^i = S^i{}_j v^j$$

relating the components $v^i$ of any vector $v \in V$ with respect to $\{e_{(i)}\}$ to its components $\hat{v}^i$ with respect to $\{\hat{e}_{(i)}\}$. Note that if we define vectors $\mathbf{v} = (v^1, \dots, v^n)$ and $\hat{\mathbf{v}} = (\hat{v}^1, \dots, \hat{v}^n) \in \mathbb{F}^n$ and regard $S^i{}_j$ as the components of an $n$-by-$n$ invertible matrix $\mathbf{S}$, this relation simply says

$$\hat{\mathbf{v}} = \mathbf{S}\mathbf{v}.$$

[1] A reminder: any matching pair of upper and lower indices implies a summation, so some terms in (A.2) have as many as *three* implied summations.

## A.3 Dual spaces

Any $n$-dimensiional vector space $V$ has a corresponding *dual space*

$$V^* := \mathrm{Hom}(V, \mathbb{F}),$$

whose elements are called *dual vectors*, or sometimes *covectors*, or 1-*forms*; physicists also favor the term *covariant* (as opposed to *contravariant*) vectors. The spaces $V$ and $V^*$ are closely related and are in fact isomorphic, though it's important to observe that there is no *canonical* isomorphism between them. Isomorphisms between $V$ and $V^*$ do arise naturally from various types of extra structure we might add to $V$: the simplest of these is a basis. Indeed, if $e_{(1)}, \ldots, e_{(n)}$ is a basis of $V$, there is a corresponding *dual basis* $\theta^{(1)}, \ldots, \theta^{(n)}$ of $V^*$, defined by the condition

$$\theta^{(i)}(e_{(j)}) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Extending the definition of $\theta^{(i)}$ by linearity to a map $V \to \mathbb{F}$, we see that for any $v = v^j e_{(j)} \in V$,

$$\theta^{(i)}(v^j e_{(j)}) = v^i.$$

Notice that we've chosen an *upper* index for the dual basis vectors, and we will correspondingly use a lower index for components in $V^*$:

$$\alpha = \alpha_j \theta^{(j)} \in V^*.$$

This choice is motivated by the fact that dual vectors can naturally be *paired* with vectors, giving rise to an implied summation:

$$\alpha(v) = \alpha_j \theta^{(j)}(v^i e_{(i)}) = \alpha_j v^i \theta^{(j)}(e_{(i)}) = \alpha_j v^j \in \mathbb{F}. \tag{A.3}$$

When working in a basis, it often makes sense to think of vectors as *column vectors* in $\mathbb{F}^n$, and dual vectors as *row vectors*, i.e.

$$\mathbf{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \qquad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \cdots \alpha_n \end{pmatrix},$$

so that in terms of matrix multiplication, (A.3) becomes

$$\alpha(v) = \boldsymbol{\alpha}\mathbf{v}.$$

There are situations in which the choice to use lower indices for components of dual vectors might not make sense. After all, $V^*$ is itself a vector space, and independently of its association with $V$, we could simply choose an arbitrary basis $\theta_{(1)}, \ldots, \theta_{(n)}$ of $V^*$ and write dual vectors as $\alpha = \alpha^j \theta_{(j)}$. The difference is one of perspective rather than reality. Whenever we wish to view elements of $V^*$ specifically as linear maps $V \to \mathbb{F}$, it is customary and appropriate to use lower indices for components.

While the isomorphism between $V$ and $V^*$ is generally dependent on a choice, it should be noted that the dual space of $V^*$ itself is *naturally* isomorphic to $V$. Indeed, an isomorphism $\Phi : V \to V^{**}$ is defined by setting

$$\Phi(v)(\alpha) := \alpha(v)$$

for any $\alpha \in V^*$. It is therefore often convenient to blur the distinction between $V$ and $V^{**}$, using the same notation for elements of both.

**Exercise A.4.** Verify that the map $\Phi : V \to V^{**}$ defined above is an isomorphism. *Note: this is not always true in infinite dimensional vector spaces.*

**Exercise A.5.** Referring to Exercise A.3, assume $e_{(1)}, \ldots, e_{(n)}$ is a basis of $V$ and $\hat{e}_{(1)}, \ldots, \hat{e}_{(n)}$ is another basis, related to the first by $e_{(i)} = \hat{e}_{(j)} S^j{}_i$ where $S^i{}_j \in \mathbb{F}$ are the components of an invertible $n$-by-$n$ matrix $\mathbf{S}$. Denote the components of $\mathbf{S}^{-1}$ by $(S^{-1})^i{}_j$, and show that the corresponding dual bases are related by

$$\theta^{(i)} = (S^{-1})^i{}_j \hat{\theta}^{(j)},$$

while the components of a dual vector $\alpha = \alpha_i \theta^{(i)} = \hat{\alpha}_i \hat{\theta}^{(i)}$ transform as

$$\hat{\alpha}_i = \alpha_j (S^{-1})^j{}_i.$$

In particular, putting these components together as row vectors, we have

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} \mathbf{S}^{-1}.$$

## A.4 Inner products, raising and lowering indices

On a real vector space $V$, an *inner product* is a pairing $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ that has the following properties:

- (BILINEAR) For any fixed $v_0 \in V$, the maps $V \to \mathbb{R} : v \mapsto \langle v_0, v \rangle$ and $v \mapsto \langle v, v_0 \rangle$ are both linear.
- (SYMMETRIC) $\langle v, w \rangle = \langle w, v \rangle$.
- (POSITIVE) $\langle v, v \rangle \geq 0$, with equality if and only if $v = 0$.[2]

[2] As we'll discuss at the end of this section, is is sometimes appropriate to relax the positivity condition—this is particularly important in the geometric formulation of relativity.

In the complex case we instead consider a pairing $\langle\ ,\ \rangle : V \times V \to \mathbb{C}$ and generalize the first two properties as follows:

- (SESQUILINEAR) For any fixed $v_0 \in V$, the maps $V \to \mathbb{C} : v \mapsto \langle v_0, v\rangle$ and $v \mapsto \langle v, v_0\rangle$ are linear and antilinear respectively.
- (SYMMETRY) $\langle v, w\rangle = \overline{\langle w, v\rangle}$.

The standard models of inner products are the dot product for vectors $\mathbf{v} = (v^1, \ldots, v^n)$ in Euclidean $n$-space,

$$\mathbf{v} \cdot \mathbf{w} = \sum_{j=1}^{n} v^j w^j, \tag{A.4}$$

and its complex analogue in $\mathbb{C}^n$,

$$\mathbf{v} \cdot \mathbf{w} = \sum_{j=1}^{n} \bar{v}^j w^j. \tag{A.5}$$

In both cases, one interprets

$$|\mathbf{v}| := \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\sum_j |v^j|^2}$$

as the length of the vector $\mathbf{v}$, and in the real case, one can also compute the angle $\theta$ between vectors $\mathbf{v}$ and $\mathbf{w}$ via the formula $\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}||\mathbf{w}| \cos\theta$. Inner products on real vector spaces are always understood to have this geometric interpretation.

In some sense, (A.4) and (A.5) describe all possible inner products. Certainly, choosing a basis $e_{(1)}, \ldots, e_{(n)}$ of any vector space $V$, one can write vectors in components $v = v^j e_{(j)}$ and use (A.4) or (A.5) to define an inner product. In this case the chosen basis turns out to be an *orthonormal* basis, meaning

$$\langle e_{(i)}, e_{(j)}\rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Conversely, one can show that any inner product $\langle\ ,\ \rangle$ admits an orthonormal basis,[3] in which case a quick computation gives (A.4) or (A.5) as the formula for $\langle\ ,\ \rangle$ in components.

Given any basis $e_{(1)}, \ldots, e_{(n)}$ of $V$, not necessarily orthonormal, $\langle\ ,\ \rangle$ is fully determined by the set of scalars

$$g_{ij} := \langle e_{(i)}, e_{(j)}\rangle \in \mathbb{F},$$

[3] Such a basis is constructed by the *Gram-Schidt orthogonalization* procedure, see for instance [Str80].

for $i, j \in \{1, \dots, n\}$. Indeed, we compute

$$\langle v, w \rangle = \langle v^i e_{(i)}, w^j e_{(j)} \rangle = \bar{v}^i w^j \langle e_{(i)}, e_{(j)} \rangle = g_{ij} \bar{v}^i w^j. \tag{A.6}$$

(This is the complex case; the real case is the same except we can ignore complex conjugation.) Notice how the choice of two lower indices in $g_{ij}$ makes sense in light of the summation convention. The $n$-by-$n$ matrix $\mathbf{g}$ with entries $g_{ij}$ is symmetric in the real case, and Hermitian in the complex case, i.e. it satisfies $\mathbf{g}^\dagger := \overline{\mathbf{g}}^T = \mathbf{g}$. Then in matrix notation, treating $v^i$ and $w^j$ as the entries of column vectors $\mathbf{v}$ and $\mathbf{w}$, we have

$$\langle v, w \rangle = \overline{\mathbf{v}}^T \mathbf{g} \mathbf{w} = \mathbf{v}^\dagger \mathbf{g} \mathbf{w},$$

or simply $\mathbf{v}^T \mathbf{g} \mathbf{w}$ in the real case.

An inner product can be used to "raise" or "lower" indices, which is an alternative way to say that it determines a natural isomorphism between $V$ and its dual space. For simplicity, assume for the remainder of this section that $V$ is a *real* vector space (most of what we will say can be generalized to the complex case with a little care). Given an inner product on $V$, there is a homomorphism

$$V \to V^* : v \mapsto v^\flat$$

defined by setting $v^\flat(w) = \langle v, w \rangle$.[4] The positivity of $\langle\ ,\ \rangle$ implies that $v \mapsto v^\flat$ is an injective map, and it is therefore also surjective since $V$ and $V^*$ have the same dimension. The inverse map is denoted by

$$V^* \to V : \alpha \mapsto \alpha^\sharp,$$

and the resulting identification of $V$ with $V^*$ is called a *musical isomorphism*. We can now write the pairing $\langle v, w \rangle$ alternatively as either $v^\flat(w)$ or $w^\flat(v)$. In index notation, the convention is that given a vector $v = v^j e_{(j)} \in V$, we denote the corresponding dual vector

$$v^\flat = v_j \theta^{(j)},$$

i.e. the components of $v^\flat$ are labelled with the same letter but a lowered index. It is important to remember that the objects labelled by components $v^j$ and $v_j$ are not the same, but they are closely related: the danger of confusion is outweighed by the convenience of being able to express the inner product in shorthand form as

$$\langle v, w \rangle = v^\flat(w) = v_j w^j.$$

Comparing with (A.6), we find

$$v_i = g_{ij} v^j, \tag{A.7}$$

[4]In the complex case the map $v \mapsto v^\flat$ is not linear, but *antilinear*.

or in matrix notation,

$$\mathbf{v}^\flat = \mathbf{v}^T \mathbf{g}.$$

It's clear from this discussion that $\mathbf{g}$ must be an invertible matrix; its inverse will make an appearance shortly.

One can similarly "raise" the index of a dual vector $\alpha = \alpha_j \theta^{(j)}$, writing $\alpha^\sharp = \alpha^j e_{(j)}$. To write $\alpha^j$ in terms of $\alpha_j$, it's useful first to observe that there is an induced inner product on $V^*$, defined by

$$\langle \alpha, \beta \rangle := \langle \alpha^\sharp, \beta^\sharp \rangle$$

for any dual vectors $\alpha, \beta \in V^*$. Define $g^{ij} = \langle \theta^{(i)}, \theta^{(j)} \rangle$, so the same argument as in (A.6) gives

$$\langle \alpha, \beta \rangle = g^{ij} \alpha_i \beta_j.$$

This is of course the same thing as $\beta(\alpha^\sharp) = \beta_j \alpha^j$, thus

$$\alpha^i = g^{ij} \alpha_j. \tag{A.8}$$

In light of (A.7), we see now that $g^{ij}$ are precisely the entries of the inverse matrix $\mathbf{g}^{-1}$. This fact can be expressed in the form

$$g_{ij} g^{jk} = \delta_i^{\ k},$$

where the right hand side is the *Kronecker delta*,

$$\delta_i^{\ j} := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

In some situations, notably in *Lorentzian* geometry (the mathematical setting for General Relativity), one prefers to use inner products that are not necessarily positive but satisfy a weaker requirement:

- (NONDEGENERATE) There is no $v_0 \in V$ such that $\langle v_0, v \rangle = 0$ for all $v \in V$.

An example is the *Minkowski* inner product, defined for *four-vectors* $v = v^\mu e_{(\mu)} \in \mathbb{R}^4$, $\mu = 0, \dots, 3$ by

$$\langle v, w \rangle = v^0 w^0 - \sum_{j=1}^{3} v^j w^j.$$

This plays a crucial role in relativity: though one can no longer interpret $\sqrt{\langle v, v \rangle}$ as a length, the product contains information about the geometry of three-dimensional space while treating time (the "zeroth" dimension) somewhat differently.

All of the discussion above is valid for this weaker notion of inner products as well. The crucial observation is that nondegeneracy guarantees that the homomorphism $V \to V^* : v \mapsto v^\flat$ be injective, and therefore still an isomorphism—then the same prescription for raising and lowering indices still makes sense. So for instance, using the summation convention we can write the Minkowski inner product as $\langle v, w \rangle = v_\mu w^\mu = \eta_{\mu\nu} v^\mu w^\nu$, where $\eta_{\mu\nu}$ are the entries of the matrix

$$\boldsymbol{\eta} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

**Exercise A.6.** If $\langle\ ,\ \rangle$ is the standard inner product on $\mathbb{R}^n$ and $X = (X^1, \ldots, X^n) \in \mathbb{R}^n$ is a vector, show that the components $X_j$ of $X^\flat \in (\mathbb{R}^n)^*$ satisfy $X_j = X^j$. Show however that this is *not* true if $\langle\ ,\ \rangle$ is the Minkowski inner product on $\mathbb{R}^4$.

## A.5 Direct sums

The direct sum of two vector spaces $V$ and $W$ is the vector space $V \oplus W$ consisting of pairs $(v, w) \in V \times W$, with vector addition and scalar multiplication defined by

$$(v, w) + (v', w') = (v + v', w + w'),$$
$$\lambda(v, w) = (\lambda v, \lambda w).$$

As a set, $V \oplus W$ is the same as the Cartesian product $V \times W$, but the "sum" notation is more appropriate from a linear algebra perspective since $\dim(V \oplus W) = \dim V + \dim W$.

One can easily extend the definition of a direct sum to more than two vector spaces: in particular the direct sum of $k$ copies of $V$ itself is sometimes denoted by

$$V^k = V \oplus \ldots \oplus V.$$

Both $V$ and $W$ are naturally subspaces of $V \oplus W$ by identifying $v \in V$ with $(v, 0) \in V \oplus W$ and so forth; in particular then, $V$ and $W$ are transverse subspaces with trivial intersection. Given bases $e_{(1)}, \ldots, e_{(m)} \in V$ and $f_{(1)}, \ldots, f_{(n)} \in W$, we naturally obtain a basis of $V \oplus W$ in the form

$$e_{(1)}, \ldots, e_{(m)}, f_{(1)}, \ldots, f_{(n)} \in V \oplus W.$$

Moreover if both spaces have inner products, denoted $\langle\ ,\ \rangle_V$ and $\langle\ ,\ \rangle_W$ respectively, an inner product on the direct sum is naturally defined by

$$\langle (v, w), (v', w') \rangle_{V \oplus W} = \langle v, v' \rangle_V + \langle w, w' \rangle_W.$$

In terms of components, if $\langle\ ,\ \rangle_V$ and $\langle\ ,\ \rangle_W$ are described by matrices $g^V_{ij}$ and $g^W_{ij}$ respectively, then the matrix $g^{V\oplus W}_{ij}$ for $\langle\ ,\ \rangle_{V\oplus W}$ has the form

$$\mathbf{g}^{V\oplus W} = \begin{pmatrix} \mathbf{g}^V & \\ & \mathbf{g}^W \end{pmatrix}.$$

**Exercise A.7.** Show that the spaces $(V \oplus W)^*$ and $V^* \oplus W^*$ are naturally isomorphic.

## A.6 Tensors and multilinear maps

We now begin the generalization from linear to *multi*linear algebra. We've already seen one important example of a multilinear map, namely the inner product on a real vector space $V$, which gives a *bilinear* transformation $V \times V \to \mathbb{R}$. More generally, given vector spaces $V_1, \ldots, V_k$ and $W$, a map

$$T : V_1 \times \ldots \times V_k \to W$$

is called *multilinear* if it is separately linear on each factor, i.e. for each $m = 1, \ldots, k$, fixing $v_j \in V_j$ for $j = 1, \ldots, m-1, m+1, \ldots, k$, the map

$$V_m \to W : v \mapsto T(v_1, \ldots, v_{m-1}, v, v_{m+1}, \ldots, v_k)$$

is linear.

**Definition A.8.** For an $n$-dimensional vector space $V$ and nonnegative integers $k$ and $\ell$, define the vector space $V^k_\ell$ to consist of all multilinear maps

$$T : \underbrace{V \times \ldots \times V}_{\ell} \times \underbrace{V^* \times \ldots \times V^*}_{k} \to \mathbb{F}.$$

These are called *tensors of type* $(k, \ell)$ over $V$.

Thus tensors $T \in V^k_\ell$ act on sets of $\ell$ vectors and $k$ dual vectors, and by convention $V^0_0 = \mathbb{F}$. A choice of basis $e_{(1)}, \ldots, e_{(n)}$ for $V$, together with the induced dual basis $\theta^{(1)}, \ldots, \theta^{(n)}$ for $V^*$, determines a natural basis for $V^k_\ell$ defined by setting

$$\mathbf{a}_{(i_1)\ldots(i_k)}{}^{(j_1)\ldots(j_\ell)}(e_{(j_1)}, \ldots, e_{(j_\ell)}, \theta^{(i_1)}, \ldots, \theta^{(i_k)}) = 1$$

and requiring that $\mathbf{a}_{(i_1)\ldots(i_k)}{}^{(j_1)\ldots(j_\ell)}$ vanish on any other combination of basis vectors and basis dual vectors. Here the indices $i_k$ and $j_k$ each vary from 1 to $n$, thus $\dim V^k_\ell = n^{k+\ell}$.

To any $T \in V^k_\ell$, we assign $k$ upper indices and $\ell$ lower indices $T^{i_1\ldots i_k}{}_{j_1\ldots j_\ell} \in \mathbb{F}$, so that

$$T = T^{i_1\ldots i_k}{}_{j_1\ldots j_\ell}\, \mathbf{a}_{(i_1)\ldots(i_k)}{}^{(j_1)\ldots(j_\ell)}.$$

As one can easily check, it is equivalent to define the components by evaluating $T$ on the relevant basis vectors:

$$T^{i_1\dots i_k}{}_{j_1\dots j_\ell} = T\left(e_{(j_1)},\dots,e_{(j_\ell)},\theta^{(i_1)},\dots,\theta^{(i_k)}\right).$$

The evaluation of $T$ on a general set of vectors $v_{(i)} = v^j_{(i)} e_{(j)}$ and dual vectors $\alpha^{(i)} = \alpha^{(i)}_j \theta^{(j)}$ now takes the form

$$T\left(v_{(1)},\dots,v_{(\ell)},\alpha^{(1)},\dots,\alpha^{(k)}\right) = T^{i_1\dots i_k}{}_{j_1\dots j_\ell}\, v^{j_1}_{(1)}\dots v^{j_\ell}_{(\ell)} \alpha^{(1)}_{i_1}\dots\alpha^{(k)}_{i_k}.$$

We've seen several examples of tensors so far. Obviously

$$V^0_1 = \mathrm{Hom}(V,\mathbb{F}) = V^*,$$

so tensors of type $(0,1)$ are simply dual vectors. Similarly, we have $V^1_0 = \mathrm{Hom}(V^*,\mathbb{F}) = V^{**}$, which, as was observed in §A.3, is naturally isomorphic to $V$. Thus we can think of tensors of type $(1,0)$ as vectors in $V$. An inner product on a real vector space $V$ is a tensor of type $(0,2)$, and the corresponding inner product on $V^*$ is a tensor of type $(2,0)$.[5] Note that our conventions on upper and lower indices for inner products are consistent with the more general definition above for tensors.

Here is a slightly less obvious example of a tensor that we've already seen: it turns out that tensors of type $(1,1)$ can be thought of simply as linear maps $V \to V$. This is suggested already by the observation that both objects have the same pattern of indices: one upper and one lower, each running from 1 to $n$.

**Proposition A.9.** *There is a natural isomorphism $\Phi : \mathrm{End}(V) \to V^1_1$ defined by*

$$\Phi(A)(v,\alpha) = \alpha(Av),$$

*and the components with respect to any basis of $V$ satisfy $A^i{}_j = [\Phi(A)]^i{}_j$.*

*Proof.* One easily checks that $\Phi$ is a linear map and both spaces have dimension $n^2$, thus we only need to show that $\Phi$ is injective. Indeed, if $\Phi(A) = 0$ then $\alpha(Av) = 0$ for all $v \in V$ and $\alpha \in V^*$, implying $A = 0$, so $\Phi$ is in fact an isomorphism. The identification of the components follows now by observing

$$\Phi(A)(v,\alpha) = [\Phi(A)]^i{}_j v^j \alpha_i = \alpha(Av) = \alpha_i A^i{}_j v^j.$$

□

[5] The complex case is slightly more complicated because bilinear does not mean quite the same thing as *sesquilinear*. To treat this properly we would have to generalize our definition of tensors to allow *antilinearity* on some factors. Since we're more interested in the real case in general, we leave further details on the complex case as an exercise to the reader.

**Exercise A.10.** Generalize Prop. A.9 to find a natural isomorphism between $V_k^1$ and the space of multilinear maps $\underbrace{V \times \ldots \times V}_{k} \to V$.[6]

**Exercise A.11.** You should do the following exercise exactly once in your life. Given distinct bases $\{e_{(i)}\}$ and $\{\hat{e}_{(j)}\}$ related by $e_{(i)} = \hat{e}_{(j)} S^j{}_i$ as in Exercises A.3 and A.5, show that the components $T^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell}$ and $\widehat{T}^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell}$ of a tensor $T \in V_\ell^k$ with respect to these two bases are related by

$$\widehat{T}^{i_1 \ldots i_k}{}_{j_1 \ldots j_\ell} = S^{i_1}{}_{p_1} \ldots S^{i_k}{}_{p_k} T^{p_1 \ldots p_k}{}_{q_1 \ldots q_\ell} (S^{-1})^{q_1}{}_{j_1} \ldots (S^{-1})^{q_\ell}{}_{j_\ell}. \tag{A.9}$$

For the case of a type $(1,1)$ tensor $A \in \mathrm{End}(V)$, whose components $A^i{}_j$ and $\hat{A}^i{}_j$ form square matrices $\mathbf{A}$ and $\widehat{\mathbf{A}}$ respectively, the transformation formula (A.9) reduces to

$$\widehat{\mathbf{A}} = \mathbf{S}\mathbf{A}\mathbf{S}^{-1}. \tag{A.10}$$

Formula (A.9) is important for historical reasons: in classical texts on differential geometry, tensors were often defined not directly as multilinear maps but rather as indexed sets of scalars that transform precisely as in (A.9) under a change of basis. In fact, this is still the most common definition in the physics literature. Mathematicians today much prefer the manifestly basis-independent definition via multilinear maps, but (A.9) and (A.10) are nevertheless occasionally useful, as we see in the next result.

**Proposition A.12.** *If $A \in V_1^1$ has components $A^i{}_j$ with respect to any basis of $V$, the scalar $A^i{}_i \in \mathbb{F}$ (note the implied summation!) is independent of the choice of basis.*

*Proof.* In linear algebra terms, $A^i{}_i$ is the *trace* $\operatorname{tr} \mathbf{A}$, so we appeal to the well known fact that traces are unchanged under change of basis. The proof of this is quite simple: it begins with the observation that for any two $n$-by-$n$ matrices $\mathbf{B}$ and $\mathbf{C}$,

$$\operatorname{tr}(\mathbf{BC}) = (BC)^i{}_i = B^i{}_j C^j{}_i = C^i{}_j B^j{}_i = (CB)^i{}_i = \operatorname{tr}(\mathbf{CB}).$$

Thus we can rearrange ordering and compute

$$\operatorname{tr} \widehat{\mathbf{A}} = \operatorname{tr}(\mathbf{SAS}^{-1}) = \operatorname{tr}[(\mathbf{SA})\mathbf{S}^{-1}] = \operatorname{tr}[\mathbf{S}^{-1}(\mathbf{SA})] = \operatorname{tr} \mathbf{A}.$$

□

[6] An important example in the case $k = 3$ appears in Riemannian geometry: the *Riemann tensor*, which carries all information about curvature on a Riemannian manifold $M$, is a tensor field of type $(1,3)$, best interpreted as a trilinear bundle map $TM \oplus TM \oplus TM \to TM$.

This result implies that there is a well defined operation

$$\operatorname{tr} : V_1^1 \to \mathbb{F}$$

which associates to $A \in V_1^1$ the trace $\operatorname{tr} A = A^i{}_i \in \mathbb{F}$ computed with respect to any basis (and independent of the choice). This operation on the tensor $A$ is called a *contraction*. One can generalize Prop. A.12 to define more general contractions

$$V_{\ell+1}^{k+1} \to V_\ell^k : T \mapsto \operatorname{tr} T$$

by choosing any $p \in 1, \ldots, k+1$ and $q \in 1, \ldots, \ell+1$, then computing the corresponding trace of the components $T^{i_1 \ldots i_{\ell+1}}{}_{j_1 \ldots j_{k+1}}$ to define $\operatorname{tr} T$ with components

$$(\operatorname{tr} T)^{i_1 \ldots i_\ell}{}_{j_1 \ldots j_k} = T^{i_1 \ldots i_{q-1} m i_q \ldots i_\ell}{}_{j_1 \ldots j_{p-1} m j_p \ldots j_k}.$$

An important example is the *Ricci curvature* on a Riemannian manifold: it is a tensor field of type $(0, 2)$ defined as a contraction of a tensor field of type $(1, 3)$, namely the Riemann curvature tensor. (See [GHL04] or [Car]).

If $V$ is a real vector space with inner product $\langle\ ,\ \rangle$, the *musical* isomorphisms $V \to V^* : v \mapsto v^\flat$ and $V^* \to V : \alpha \mapsto \alpha^\sharp$ give rise to various isomorphisms

$$V_\ell^k \to V_{\ell-1}^{k+1} \quad \text{and } V_\ell^k \to V_{\ell+1}^{k-1}.$$

For instance, if $T \in V_\ell^k$ with $k \geq 1$, then for any $m = 1, \ldots, k$, we can define a new multlinear map

$$T^\flat : \underbrace{V \times \ldots \times V}_{\ell} \times \underbrace{V^* \times \ldots \times V^*}_{m-1} \times V \times \underbrace{V^* \times \ldots \times V^*}_{k-m} \to \mathbb{R}$$

by

$$\begin{aligned} T^\flat(v_{(1)}, \ldots, v_{(\ell)}, \alpha^{(1)}, \ldots, \alpha^{(m-1)}, v, \alpha^{(m+1)}, \ldots, \alpha^{(k)}) \\ = T(v_{(1)}, \ldots, v_{(\ell)}, \alpha^{(1)}, \ldots, \alpha^{(m-1)}, v^\flat, \alpha^{(m+1)}, \ldots, \alpha^{(k)}). \end{aligned}$$

Choosing a basis, we denote the components of the inner product by $g_{ij}$ and recall the relation $v_i = g_{ij} v^j$ between the components of $v^\flat$ and $v$ respectively. Then we find that $T^\flat$ has components

$$T^{i_1 \ldots i_{m-1}}{}_r{}^{i_{m+1} \ldots i_k}{}_{j_1 \ldots j_\ell} = g_{rs} T^{i_1 \ldots i_{m-1} s i_{m+1} \ldots i_k}{}_{j_1 \ldots j_\ell}.$$

By reordering the factors slightly, we can regard $T^\flat$ naturally as a tensor in $V_{\ell+1}^{k-1}$. This operation $T \mapsto T^\flat$ is often referred to as using the inner product to *lower an index* of $T$. Indices can similarly be raised, giving isomorphisms $V_\ell^k \to V_{\ell-1}^{k+1} : T \mapsto T^\sharp$. Observe that by definition, the inner product $g^{ij}$ on $V^*$ is itself a tensor of type $(2, 0)$ that we obtain from the inner product $g_{ij}$ on $V$ by raising both indices:

$$g^{ij} = g^{ik} g^{j\ell} g_{k\ell}.$$

This implies again the fact that $g^{ij}$ and $g_{ij}$ are inverse matrices.

## A.7 The tensor product

The $n^{k+\ell}$-dimensional vector space $V^k_\ell$ can be thought of in a natural way as a "product" of $k+\ell$ vector spaces of dimension $n$, namely $k$ copies of $V$ and $\ell$ copies of $V^*$. To make this precise, we must define the *tensor product* $V \otimes W$ of two vector spaces $V$ and $W$. This is a vector space whose dimension is the product of $\dim V$ and $\dim W$, and it comes with a natural bilinear "product" operation $\otimes : V \times W \to V \otimes W : (v, w) \mapsto v \otimes w$.

There are multiple ways to define the tensor product, with a varying balance between concreteness and abstract simplicity: we shall begin on the more concrete end of the spectrum by defining the bilinear operation

$$\otimes : V^k_\ell \times V^p_q \to V^{k+p}_{\ell+q} : (S, T) \mapsto S \otimes T,$$

$$\begin{aligned}(S \otimes T)&(v_{(1)}, \dots, v_{(\ell)}, w_{(1)}, \dots, w_{(q)}, \alpha^{(1)}, \dots, \alpha^{(k)}, \beta^{(1)}, \dots, \beta^{(p)}) \\ &:= S(v_{(1)}, \dots, v_{(\ell)}, \alpha^{(1)}, \dots, \alpha^{(k)}) \cdot T(w_{(1)}, \dots, w_{(q)}, \beta^{(1)}, \dots, \beta^{(p)}).\end{aligned}$$

This extends naturally to an associative multilinear product for any number of tensors on $V$. In particular, choosing a basis $e_{(1)}, \dots, e_{(n)}$ of $V = V^{**}$ and corresponding dual basis $\theta^{(1)}, \dots, \theta^{(n)}$ of $V^*$, one checks easily that the naturally induced basis of $V^k_\ell$ described in the previous section consists of the tensor products

$$\mathbf{a}_{(i_1)\dots(i_k)}{}^{(j_1)\dots(j_\ell)} = \theta^{(j_1)} \otimes \dots \otimes \theta^{(j_\ell)} \otimes e_{(i_1)} \otimes \dots \otimes e_{(i_k)}.$$

The infinite direct sum

$$\mathcal{T}(V) = \bigoplus_{k,\ell} V^k_\ell,$$

with its bilinear product operation $\otimes : \mathcal{T}(V) \times \mathcal{T}(V) \to \mathcal{T}(V)$ is called the *tensor algebra over $V$*.

The above suggests the following more general definition of a tensor product. Recall that any finite dimensional vector space $V$ is naturally isomorphic to $V^{**}$, the dual of its dual space, and thus every vector $v \in V$ can be identified with the linear map $V^* \to \mathbb{R} : \alpha \mapsto \alpha(v)$. Now for any two finite dimensional vector spaces $V$ and $W$, define $V \otimes W$ to be the vector space of bilinear maps $V^* \times W^* \to \mathbb{R}$; we then have a natural product operation $\otimes : V \times W \to V \otimes W$ such that

$$(v \otimes w)(\alpha, \beta) = \alpha(v)\beta(w)$$

for any $\alpha \in V^*$, $\beta \in W^*$. Extending the product operation in the obvious way to more than two factors, one can then define the $k$-fold tensor product of $V$ with itself,

$$\otimes^k V = \bigotimes_{j=1}^{k} V = \underbrace{V \otimes \dots \otimes V}_{k}.$$

There is now a natural isomorphism

$$V_\ell^k = \left(\otimes^k V^*\right) \otimes \left(\otimes^\ell V\right).$$

**Exercise A.13.** If $e_{(1)}, \ldots, e_{(m)}$ is a basis of $V$ and $f_{(1)}, \ldots, f_{(n)}$ is a basis of $W$, show that the set of all products of the form $e_{(i)} \otimes f_{(j)}$ gives a basis of $V \otimes W$. In particular, $\dim(V \otimes W) = mn$.

We now give an equivalent definition which is more abstract but has the virtue of not relying on the identification of $V$ with $V^{**}$. If $X$ is any set, denote by $\mathcal{F}(X)$ the *free vector space generated by* $X$, defined as the set of all formal sums

$$\sum_{x \in X} a_x x$$

with $a_x \in \mathbb{F}$ and only finitely many of the coefficients $a_x$ nonzero. Addition and scalar multiplication on $\mathcal{F}(X)$ are defined by

$$\sum_{x \in X} a_x x + \sum_{x \in X} b_x x = \sum_{x \in X} (a_x + b_x) x,$$
$$c \sum_{x \in X} a_x x = \sum_{x \in X} c a_x x.$$

Note that each element of $X$ can be considered a vector in $\mathcal{F}(X)$, and unless $X$ is a finite set, $\mathcal{F}(X)$ is infinite dimensional.

Setting $X = V \times W$, there is an equivalence relation $\sim$ on $\mathcal{F}(V \times W)$ generated by the relations

$$(v + v', w) \sim (v, w) + (v', w), \qquad (v, w + w') \sim (v, w) + (v, w'),$$
$$(cv, w) \sim c(v, w) \sim (v, cw)$$

for all $v, v' \in V$, $w, w' \in W$ and $c \in \mathbb{F}$. We then define

$$V \otimes W = \mathcal{F}(V \times W) / \sim,$$

and denoting by $[x]$ the equivalence class represented by $x \in V \times W$,

$$v \otimes w := [(v, w)].$$

The definition of our equivalence relation is designed precisely so that this tensor product operation should be bilinear. It follows from Exercises A.17 and A.18 below that our two definitions of $V \otimes W$ are equivalent.

**Exercise A.14.** Show that $V \otimes W$ as defined above has a well defined vector space structure induced from that of $\mathcal{F}(V \times W)$, and that $\otimes$ is then a bilinear map $V \times W \to V \otimes W$.

**Exercise A.15.** Show that if $e_{(1)}, \ldots, e_{(m)}$ is a basis of $V$ and $f_{(1)}, \ldots, f_{(n)}$ a basis of $W$, a basis of $V \otimes W$ (according to the new definition) is given by

$$\{e_{(i)} \otimes f_{(j)}\}_{i=1,\ldots,m,\ j=1,\ldots,n}.$$

Moreover if $v = v^i e_{(i)} \in V$ and $w = w^i f_{(i)} \in W$ then $v \otimes w = (v \otimes w)^{ij} e_{(i)} \otimes f_{(j)}$ where the components of the product are given by

$$(v \otimes w)^{ij} = v^i w^j.$$

Observe that elements of $V \otimes W$ can often be written in many different ways, for example $2(v \otimes w) = 2v \otimes w = v \otimes 2w$, and $0 = 0 \otimes w = v \otimes 0$ for any $v \in V$, $w \in W$. It is also important to recognize that (in contrast to the direct sum $V \oplus W$) not every vector in $V \otimes W$ can be written as a product $v \otimes w$, though everything is a *sum* of such products. The following exercise gives an illustrative example.

**Exercise A.16.** Denote by $\mathbf{e}_{(j)}$ the standard basis vectors of $\mathbb{R}^n$, regarded as column vectors. Show that there is an isomorphism $\mathbb{R}^m \otimes \mathbb{R}^n \cong \mathbb{R}^{m \times n}$ that maps $\mathbf{e}_{(i)} \otimes \mathbf{e}_{(j)}$ to the $m$-by-$n$ matrix $\mathbf{e}_{(i)} {\mathbf{e}_{(j)}}^{\mathrm{T}}$. The latter has 1 in the $i$th row and $j$th column, and zero everywhere else.

**Exercise A.17.** For any vector spaces $V_1, \ldots, V_k$, find a natural isomorphism $(V_1 \otimes \ldots \otimes V_k)^* = V_1^* \otimes \ldots \otimes V_k^*$.

**Exercise A.18.** For any vector spaces $V_1, \ldots, V_k$ and $W$, show that there is a natural isomorphism between $\mathrm{Hom}(V_1 \otimes \ldots \otimes V_k, W)$ and the space of multilinear maps $V_1 \times \ldots \times V_k \to W$.

**Exercise A.19.** Use the second definition of the tensor product to show that the following spaces are all naturally isomorphic:

(i) $V_\ell^k$
(ii) $\left(\otimes^\ell V^*\right) \otimes \left(\otimes^k V\right)$
(iii) $\mathrm{Hom}\left(\otimes^\ell V, \otimes^k V\right)$

If $V$ and $W$ are spaces of dimension $m$ and $n$ equipped with inner products $\langle\ ,\ \rangle_V$ and $\langle\ ,\ \rangle_W$ respectively, then there is a natural inner product $\langle\ ,\ \rangle_{V \otimes W}$ on $V \otimes W$ such that

$$\langle v \otimes w, v' \otimes w' \rangle_{V \otimes W} = \langle v, v' \rangle_V \cdot \langle w, w' \rangle_W.$$

This product is extended uniquely to all pairs in $V \otimes W$ by bilinearity, though the reader should take a moment to check that the resulting construction is well defined. Recall from §A.4 that an inner product on $V$ also gives rise naturally to an inner product on $V^*$. In this way, one also obtains natural inner products on the tensor spaces $V_\ell^k$. For example on $\otimes^k V$, the product $\langle\ ,\ \rangle_{\otimes^k V}$ has the property that if $e_{(1)}, \ldots, e_{(n)}$ is an orthonormal basis of $V$, then the basis of $\otimes^k V$ defined by all products of the form

$$e_{(i_1)} \otimes \ldots \otimes e_{(i_k)}$$

is also orthonormal.

## A.8 Symmetric and exterior algebras

For an $n$-dimensional vector space $V$, we now single out some special subspaces of the $k$-fold tensor product $\otimes^k V$. These are simplest to understand when $V$ is given as a dual space, since $\otimes^k V^*$ is equivalent to the space of $k$-multilinear maps $V \times \ldots \times V \to \mathbb{F}$. We examine this case first.

Recall that a *permutation* of $k$ elements is by definition a bijective map $\sigma$ of the set $\{1, \ldots, k\}$ to itself. There are $k!$ distinct permutations, which form the *symmetric group* $S_k$. It is generated by a set of simple permutations $\sigma_{ij}$ for which $\sigma(i) = j$, $\sigma(j) = i$ and $\sigma$ maps every other number to itself. We call such a permutation a *flip*. In general, any $\sigma \in S_k$ is called *odd* (*even*) if it can be written as a composition of an odd (even) number of flips. We define the *parity* of $\sigma$ by

$$|\sigma| = \begin{cases} 0 & \text{if } \sigma \text{ is even,} \\ 1 & \text{if } \sigma \text{ is odd.} \end{cases}$$

The parity usually appears in the form of a sign $(-1)^{|\sigma|}$, thus one sometimes also refers to odd or even permutations as *negative* or *positive* respectively.

Regarding $\otimes^k V^*$ as a space of multilinear maps on $V$, an element $T \in \otimes^k V^*$ is called *symmetric* if $T(v_1, \ldots, v_k)$ is always unchanged under exchange of any two of the vectors $v_i$ and $v_j$. Similarly we call $T$ *antisymmetric* (or sometimes *skew-symmetric* or *alternating*) if $T(v_1, \ldots, v_k)$ changes sign under every such exchange. Both definitions can be rephrased in terms of permutations by saying that $T$ is symmetric if for all $v_1, \ldots, v_k \in V$ and any $\sigma \in S_k$,

$$T(v_1, \ldots, v_k) = T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),$$

while $T$ is antisymmetric if

$$T(v_1, \ldots, v_k) = (-1)^{|\sigma|} T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

The sets of symmetric and antisymmetric tensors are clearly linear subspaces of $\otimes^k V^*$, which we denote by

$$S^k V^* \quad \text{and} \quad \Lambda^k V^*$$

respectively.

Define the *symmetric projection* $\mathrm{Sym} : \otimes^k V^* \to \otimes^k V^*$ by

$$(\mathrm{Sym}\, T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),$$

and the *antisymmetric* (or *alternating*) *projection* $\mathrm{Alt} : \otimes^k V^* \to \otimes^k V^*$,

$$(\mathrm{Alt}\, T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^{|\sigma|} T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

Both are linear maps.

**Exercise A.20.** Show that

(i) $\mathrm{Sym} \circ \mathrm{Sym} = \mathrm{Sym}$ and $\mathrm{Alt} \circ \mathrm{Alt} = \mathrm{Alt}$.

(ii) A tensor $T \in \otimes^k V^*$ is in $S^k V^*$ if and only if $\mathrm{Sym}(T) = T$, and $T \in \Lambda^k V^*$ if and only if $\mathrm{Alt}(T) = T$.

The subspaces $S^k V, \Lambda^k V \subset \otimes^k V$ can be defined via the recipe above if we treat $V$ as the dual space of $V^*$, but of course this is not the most elegant approach. Instead we generalize the above constructions as follows. Define $\mathrm{Sym} : \otimes^k V \to \otimes^k V$ as the unique linear map which acts on products $v_1 \otimes \ldots \otimes v_k$ by

$$\mathrm{Sym}(v_1 \otimes \ldots \otimes v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)}.$$

Note that this definition is somewhat indirect since not every element of $\otimes^k V$ can be written as such a product; but since every element is a sum of such products, the map Sym is clearly unique if it is well defined. We leave the proof of the latter as an exercise to the reader, with the hint that, for instance in the case $k = 2$, it suffices to prove relations of the form

$$\mathrm{Sym}((v + v') \otimes w) = \mathrm{Sym}(v \otimes w) + \mathrm{Sym}(v' \otimes w).$$

We define $\mathrm{Alt} : \otimes^k V \to \otimes^k V$ similarly via

$$\mathrm{Alt}(v_1 \otimes \ldots \otimes v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^{|\sigma|} v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)}.$$

**Exercise A.21.** Show that the above definitions of Sym and Alt on $\otimes^k V$ are equivalent to our original definitions if $V$ is regarded as the dual space of $V^*$.

It is a straightforward matter to generalize Exercise A.20 and show that Sym and Alt are both projection operators on $\otimes^k V$, that is $\mathrm{Sym} \circ \mathrm{Sym} = \mathrm{Sym}$ and $\mathrm{Alt} \circ \mathrm{Alt} = \mathrm{Alt}$. We now define the *symmetric* and *antisymmetric* subspaces to be the images of these projections:

$$S^k V = \mathrm{Sym}(\otimes^k V), \qquad \Lambda^k V = \mathrm{Alt}(\otimes^k V).$$

Equivalently, $T \in \Lambda^k V$ if and only if $\mathrm{Alt}(T) = T$, and similarly for $S^k V$. The elements of $\Lambda^k V$ are sometimes called *$k$-vectors*.

One can combine the tensor product with the projections above to define product operations that preserve symmetric and antisymmetric tensors. We focus here on the antisymmetric case, since it is of greatest use in differential geometry. The seemingly obvious definition for a product of $\alpha \in \Lambda^k V$ and $\beta \in \Lambda^\ell V$ would be

$$\mathrm{Alt}(\alpha \otimes \beta) \in \Lambda^{k+\ell} V,$$

but this is not quite right. The reason why not is most easily seen in the special case of the dual space $V^*$, where alternating forms in $\Lambda^k V^*$ can be interpreted as computing the signed volumes of parallelopipeds. In particular, assume $V$ and $W$ are real vector spaces of dimension $m$ and $n$ respectively, and $\alpha \in \Lambda^m V$ and $\beta \in \Lambda^n W$ are both nonzero. We can interpret both geometrically by saying for instance that $\alpha(v_1, \dots, v_m) \in \mathbb{R}$ is the signed volume of the parallelopiped in $V$ spanned by $v_1, \dots, v_m$, with the sign corresponding to a choice of orientation on $V$. Now extend $\alpha$ and $\beta$ to define forms on $V \oplus W$ via the natural projections $\pi_V : V \oplus W \to V$ and $\pi_W : V \oplus W \to W$, e.g.

$$\alpha(v_1, \dots, v_m) := \alpha(\pi(v_1), \dots, \pi(v_m))$$

for $v_1, \dots, v_m \in V \oplus W$. Geometrically, one now obtains a natural notion for the signed volume of $(m+n)$-dimensional parallelopipeds in $V \oplus W$, and we wish to define the *wedge product* $\alpha \wedge \beta \in \Lambda^{m+n}((V \oplus W)^*)$ to reflect this. In particular, for any set of vectors $v_1, \dots, v_m \in V$ and $w_1, \dots, w_n \in W$ we must have

$$\begin{aligned}(\alpha \wedge \beta)(v_1, \dots, v_m, w_1, \dots, w_n) &= \alpha(v_1, \dots, v_m) \cdot \beta(w_1, \dots, w_n) \\ &= (\alpha \otimes \beta)(v_1, \dots, v_m, w_1, \dots, w_n).\end{aligned} \tag{A.11}$$

Let us now compute $\mathrm{Alt}(\alpha \otimes \beta)(X_1, \dots, X_{m+n})$ where $X_j = v_j \in V$ for $j = 1, \dots, m$ and $X_{m+j} = w_j \in W$ for $j = 1, \dots, n$. The crucial observation is that only a special subset of the permutations $\sigma \in S^{m+n}$ will matter in this computation: namely,

$$(\alpha \otimes \beta)(X_{\sigma(1)}, \dots, X_{\sigma(m+n)}) = 0$$

unless $\sigma$ preserves the subsets $\{1, \dots, m\}$ and $\{m+1, \dots, m+n\}$. This means that $\sigma$ must have the form

$$\sigma(j) = \begin{cases} \sigma_V(j) & \text{if } j \in \{1, \dots, m\}, \\ \sigma_W(j-m) + m & \text{if } j \in \{m+1, \dots, m+n\} \end{cases}$$

for some pair of permutations $\sigma_V \in S_m$ and $\sigma_W \in S_n$, and in this case $(-1)^{|\sigma|} = (-1)^{|\sigma_V| + |\sigma_W|} = (-1)^{|\sigma_V|}(-1)^{|\sigma_W|}$. Thus we compute:

$$\begin{aligned}
&\operatorname{Alt}(\alpha \otimes \beta)(v_1, \dots, v_m, w_1, \dots, w_n) \\
&\quad = \frac{1}{(m+n)!} \sum_{\sigma \in S_{m+n}} (-1)^{|\sigma|} (\alpha \otimes \beta)(X_{\sigma(1)}, \dots, X_{\sigma(m+n)}) \\
&\quad = \frac{1}{(m+n)!} \sum_{\sigma_V \in S_m} \sum_{\sigma_W \in S_n} (-1)^{|\sigma_V|} \alpha(v_{\sigma_V(1)}, \dots, v_{\sigma_V(m)}) \\
&\qquad\qquad \cdot (-1)^{|\sigma_W|} \beta(w_{\sigma_W(1)}, \dots, w_{\sigma_W(n)}) \\
&\quad = \frac{m!n!}{(m+n)!} \alpha(v_1, \dots, v_m) \cdot \beta(w_1, \dots, w_n),
\end{aligned}$$

where in the last line we use the fact that $\alpha$ and $\beta$ are both alternating. Comparing this with (A.11), we see that in this special case the only geometrically sensible definition of $\alpha \wedge \beta$ satisfies the formula

$$\alpha \wedge \beta = \frac{(m+n)!}{m!n!} \operatorname{Alt}(\alpha \otimes \beta).$$

These considerations motivate the following general definition.

**Definition A.22.** For any $\alpha \in \Lambda^k V$ and $\beta \in \Lambda^\ell W$, the *wedge product* $\alpha \wedge \beta \in \Lambda^{k+\ell} V$ is defined by

$$\alpha \wedge \beta = \frac{(k+\ell)!}{k!\ell!} \operatorname{Alt}(\alpha \otimes \beta).$$

**Exercise A.23.** Show that the wedge product is bilinear and *graded symmetric*; the latter means that for $\alpha \in \Lambda^k V$ and $\beta \in \Lambda^\ell V$,

$$\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha.$$

We've taken the geometric argument above as motivation for the combinatorial factor $\frac{(k+\ell)!}{k!\ell!}$, and further justification is provided by the following result, which depends crucially on this factor:

**Exercise A.24.** Show that the wedge product is associative, i.e. for any $\alpha \in \Lambda^k V$, $\beta \in \Lambda^\ell V$ and $\gamma \in \Lambda^p V$,

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma).$$

The direct sum

$$\Lambda^* V := \bigoplus_{k=0}^{\infty} \Lambda^k V$$

with bilinear operation $\wedge : \Lambda^* V \times \Lambda^* V \to \Lambda^* V$ is called the *exterior algebra* of $V$, with the wedge product also sometimes referred to as the *exterior product*. Note that we include $k = 0$ in this sum: by convention $\Lambda^0 V = \mathbb{F}$, and the wedge product of any $c \in \Lambda^0 V$ with $\alpha \in \Lambda^k V$ is simply $c \wedge \alpha = \alpha \wedge c := c\alpha$.

In light of Exercise A.24, it makes sense to consider wedge products of more than two elements in $\Lambda^* V$, and one verifies by a simple induction argument that for any $v_1, \ldots, v_k \in V$,

$$v_1 \wedge \ldots \wedge v_k = \sum_{\sigma \in S_k} (-1)^{|\sigma|} v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)}. \tag{A.12}$$

This provides a simple way to write down a basis of $\Lambda^k V$ in terms of a given basis $e_{(1)}, \ldots, e_{(n)}$ of $V$. Indeed, recall that any $\omega \in \otimes^k V$ can be written uniquely in terms of components $\omega^{i_1 \ldots i_k} \in \mathbb{F}$ as $\omega = \omega^{i_1 \ldots i_k} e_{(i_1)} \otimes \ldots \otimes e_{(i_k)}$. A formula for these components is obtained by interpreting $\omega$ as a $k$-multilinear map on $V^*$: plugging in the corresponding dual basis vectors $\theta^{(1)}, \ldots, \theta^{(n)}$, we have

$$\begin{aligned}
\omega(\theta^{(i_1)}, \ldots, \theta^{(i_k)}) &= \omega^{j_1 \ldots j_k} (e_{(j_1)} \otimes \ldots \otimes e_{(j_k)})(\theta^{(i_1)}, \ldots, \theta^{(i_k)}) \\
&= \omega^{j_1 \ldots j_k} \theta^{(i_1)}(e_{(j_1)}) \ldots \theta^{(i_k)}(e_{(j_k)}) \\
&= \omega^{i_1 \ldots i_k}.
\end{aligned}$$

It follows that if $\omega \in \Lambda^k V$, the components $\omega^{i_1 \ldots i_k}$ are antisymmetric with respect to permutations of the indices. Then applying (A.12), we have

$$\begin{aligned}
\omega &= \mathrm{Alt}(\omega) = \mathrm{Alt}\left(\omega^{i_1 \ldots i_k} e_{(i_1)} \otimes \ldots \otimes e_{(i_k)}\right) \\
&= \omega^{i_1 \ldots i_k} \,\mathrm{Alt}\left(e_{(i_1)} \otimes \ldots \otimes e_{(i_k)}\right) = \frac{1}{k!} \omega^{i_1 \ldots i_k} e_{(i_1)} \wedge \ldots \wedge e_{(i_k)} \\
&= \sum_{i_1 < \ldots < i_k} \omega^{i_1 \ldots i_k} e_{(i_1)} \wedge \ldots \wedge e_{(i_k)},
\end{aligned}$$

where we must point out that in the last expression we are *not* using the summation convention. This proves:

**Proposition A.25.** *If $V$ is a vector space with basis $e_{(1)}, \ldots, e_{(n)}$, then for $k \in \{1, \ldots, n\}$, $\dim \Lambda^k V = \binom{n}{k} = \frac{n!}{k!(n-k)!}$, and a basis of $\Lambda^k V$ is given by all products of the form*

$$e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}$$

*with $1 \leq i_1 < \ldots < i_k \leq n$.*

An important corollary to the discussion above is that $\Lambda^k V = \{0\}$ whenever $k > n$, thus in contrast to the full tensor algebra $\mathcal{T}(V)$, $\Lambda^* V$ is finite dimensional. We also see that $\dim \Lambda^n V = 1$: in particular, for any basis $e_{(1)}, \ldots, e_{(n)}$, every $\omega \in \Lambda^n V$ is a multiple of $e_{(1)} \wedge \ldots \wedge e_{(n)}$.

**Exercise A.26.** Given a basis $e_{(1)}, \ldots, e_{(n)}$ of $V$ and a set of $n$ vectors $v_{(j)} = v^i_{(j)} e_{(i)}$, show that

$$v_{(1)} \wedge \ldots \wedge v_{(n)} = \det(\mathbf{V})\; e_{(1)} \wedge \ldots \wedge e_{(n)},$$

where $\mathbf{V}$ is the $n$-by-$n$ matrix whose entry in the $i$th row and $j$th column is $v^i_{(j)}$. *Hint:* the basis defines an isomorphism $\mathbb{F}^n \to V$, so that forming the $n$-fold wedge product above leads to an antisymmetric $n$-form on $\mathbb{F}^n$. The determinant, regarded as a multilinear map on the columns of a matrix, is such an $n$-form. How many others are there?

A *volume form* on the $n$-dimensional vector space $V$ is any nonzero element $\mu$ of the 1-dimensional vector space $\Lambda^n V$. As explained in Chapter 2, if $V$ is a real vector space then such a choice defines a notion of signed volume for $n$-dimensional parallelopipeds in $V$. Suppose now that $V$ is equipped with an inner product $\langle\ ,\ \rangle$ and $e_{(1)}, \ldots, e_{(n)}$ is an orthonormal basis. This basis then defines a volume form $e_{(1)} \wedge \ldots \wedge e_{(n)}$, and for any other basis $f_{(1)}, \ldots, f_{(n)}$ there is a number $c \in \mathbb{F}$ such that

$$f_{(1)} \wedge \ldots \wedge f_{(n)} = c\; e_{(1)} \wedge \ldots \wedge e_{(n)}. \tag{A.13}$$

We claim that if $f_{(1)}, \ldots, f_{(n)}$ is also orthonormal, then $|c| = 1$. Indeed, this follows from Exercise A.26 and the observation that the matrix $\mathbf{V}$ with entries $f^j_{(i)}$ is in this case orthogonal (or unitary, in the complex case), so $|\det(\mathbf{V})| = 1$. (See Appendix B for the relevant details on orthogonal and unitary matrices.) This observation is most interesting in the real case, for it says that the two volume forms in (A.13) are equal up to a sign. The sign can be fixed if $V$ is also equipped with an *orientation*, which means every basis of $V$ is labelled positive or negative, two bases always having the same sign if they can be deformed into one another. Indeed, the constant $c \in \mathbb{R}$ in (A.13) is positive if and only if the basis $f_{(1)}, \ldots, f_{(n)}$ can be deformed through a continuous family of bases to $e_{(1)}, \ldots, e_{(n)}$. This proves:

**Proposition A.27.** *Suppose $V$ is an oriented vector space equipped with an inner product $\langle\ ,\ \rangle$. Then there is a unique volume form $\mu \in \Lambda^n V$ such that*

$$\mu = e_{(1)} \wedge \ldots \wedge e_{(n)}$$

*for every positively oriented orthonormal basis $e_{(1)}, \ldots, e_{(n)}$.*

An inner product on $V$ also defines inner products on each of the spaces $\Lambda^k V$ in a natural way. The most obvious definition would arise from the observation that $\Lambda^k V$ is a subspace of $\otimes^k V$, so one could simply restrict $\langle\ ,\ \rangle_{\otimes^k V}$. This turns out to be almost but not quite the most desirable definition.

**Proposition A.28.** *If $V$ has an inner product $\langle\ ,\ \rangle$ then there is a unique inner product $\langle\ ,\ \rangle_{\Lambda^k V}$ on $\Lambda^k V$ such that for every orthonormal basis $e_{(1)}, \ldots, e_{(n)}$ of $V$, the basis of $\Lambda^k V$ consisting of all products of the form*

$$e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}$$

*is also orthonormal.*

*Proof.* If $e_{(1)}, \ldots, e_{(n)}$ is an orthonormal basis of $V$, then recalling from the end of §A.7 the inner product $\langle\ ,\ \rangle_{\otimes^k V}$ on $\otimes^k V$, it is easy to check that

$$\langle e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}, e_{(j_1)} \wedge \ldots \wedge e_{(j_k)} \rangle_{\otimes^k V} = 0$$

unless $e_{(i_1)} \wedge \ldots \wedge e_{(i_k)} = \pm e_{(j_1)} \wedge \ldots \wedge e_{(j_k)}$. Now from (A.12), we find

$$\langle e_{(i_1)} \wedge e_{(i_k)}, e_{(i_1)} \wedge e_{(i_k)} \rangle_{\otimes^k V} = k!,$$

thus the new product

$$\langle\ ,\ \rangle_{\Lambda^k V} := \frac{1}{k!} \langle\ ,\ \rangle_{\otimes^k V}$$

has the desired properties. □

Observe that by the above construction, the special volume form $\mu = e_{(1)} \wedge \ldots \wedge e_{(n)}$ on an oriented inner product space satisfies $\langle \mu, \mu \rangle_{\Lambda^n V} = 1$. Since $\Lambda^n V$ inherits both an inner product *and* a natural orientation from $V$, we could thus have defined $\mu$ as the one element of the unique positively oriented orthonormal basis of $\Lambda^n V$. Isn't that nice?

## A.9 Duality and the Hodge star

One notices immediately from the formula $\dim \Lambda^k V = \frac{n!}{k!(n-k)!}$ that $\Lambda^k V$ and $\Lambda^{n-k} V$ have the same dimension, and are therefore isomorphic. This raises the question of whether there is a *canonical* isomorphism $\Lambda^k V \to \Lambda^{n-k} V$. The answer is no without some additional choices, but we will show that such an isomorphism does arise naturally if $V$ is equipped with an orientation and an inner product. Throughout this section we assume $\mathbb{F} = \mathbb{R}$.

**Lemma A.29.** *The map* $\Phi : \Lambda^{n-k}V \to \mathrm{Hom}(\Lambda^k V, \Lambda^n V)$ *defined by*

$$\Phi(\beta)\alpha = \alpha \wedge \beta$$

*is an isomorphism.*

*Proof.* Since $\dim \Lambda^n V = 1$, we have $\dim \Lambda^{n-k}V = \dim \mathrm{Hom}(\Lambda^k V, \Lambda^n V) = \frac{n!}{k!(n-k)!}$, thus it suffices to check that $\Phi$ is injective. This amounts to the statement that the bilinear pairing $\wedge : \Lambda^k V \times \Lambda^{n-k}V \to \Lambda^n V$ is *nondegenerate*, i.e. for any nonzero $\alpha \in \Lambda^k V$, there exists $\beta \in \Lambda^{n-k}V$ such that $\alpha \wedge \beta \neq 0$.

We can construct such a $\beta$ explicitly using a basis $e_{(1)}, \ldots, e_{(n)}$ of $V$. A basis of $\Lambda^k V$ is then given by products of the form $e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}$ with $i_1 < \ldots < i_k$, and for each of these there is a unique "dual" product

$$*(e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}) := e_{(j_1)} \wedge \ldots \wedge e_{(j_{n-k})} \in \Lambda^{n-k}V$$

such that

$$e_{(i_1)} \wedge \ldots \wedge e_{(i_k)} \wedge e_{(j_1)} \wedge \ldots \wedge e_{(j_{n-k})} = e_{(1)} \wedge \ldots \wedge e_{(n)}.$$

Extending this by linearity to an isomorphism $* : \Lambda^k V \to \Lambda^{n-k}V$, we can now write $\alpha = \sum_{i_1 < \ldots < i_k} \alpha^{i_1 \ldots i_k} e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}$ and compute

$$\alpha \wedge *\alpha = \sum_{i_1 < \ldots < i_k} \left(\alpha^{i_1 \ldots i_k}\right)^2 e_{(1)} \wedge \ldots \wedge e_{(n)},$$

so $\alpha \wedge *\alpha = 0$ only if $\alpha = 0$. □

The notation $* : \Lambda^k V \to \Lambda^{n-k}V$ used in the proof is a preview of things to come: this isomorphism as we defined it depends on the choice of basis $e_{(1)}, \ldots, e_{(n)}$, but we will see that is no longer the case if the basis is assumed to be orthonormal and positively oriented. Indeed, if $V$ has an inner product and orientation, there is then a natural volume form $\mu \in \Lambda^n V$ and an induced inner product $\langle\ ,\ \rangle$ on $\Lambda^k V$ (see Prop. A.28), giving rise to an isomorphism

$$\Psi : \Lambda^k V \to \mathrm{Hom}(\Lambda^k V, \Lambda^n V)$$
$$\Psi(\beta)\alpha = \langle \alpha, \beta \rangle \mu.$$

We now define the *Hodge star operator*

$$* : \Lambda^k V \to \Lambda^{n-k}V$$

as $* = \Phi^{-1} \circ \Psi$, which is equivalent to saying that for any $\beta \in \Lambda^k V$, $*\beta$ is the unique element of $\Lambda^{n-k}V$ such that

$$\langle \alpha, \beta \rangle\ \mu = \alpha \wedge *\beta. \tag{A.14}$$

for all $\alpha \in \Lambda^k V$.

**Exercise A.30.** Suppose $V$ is an oriented inner product space with positive orthonormal basis $e_{(1)}, \ldots, e_{(n)}$.

(i) Show that for any set of indices $1 \le i_1 < \ldots < i_k \le n$,

$$*(e_{(i_1)} \wedge \ldots \wedge e_{(i_k)}) = e_{(j_1)} \wedge \ldots \wedge e_{(j_{n-k})}$$

with $j_1, \ldots, j_{n-k} \in \{1, \ldots, n\}$ chosen so that $(i_1, \ldots, i_k, j_1, \ldots, j_{n-k})$ is any even permutation of $(1, \ldots, n)$.

(ii) Show that $* : \Lambda^k V \to \Lambda^{n-k} V$ is an *isometry*, i.e. $\langle \alpha, \beta \rangle = \langle *\alpha, *\beta \rangle$ for any $\alpha, \beta \in \Lambda^k V$.

(iii) Show that as an operator on $\Lambda^k V$, $*^2 = (-1)^{k(n-k)}$.

## A.10 Tensor fields and forms on manifolds

Most of the concepts above can be generalized wholesale from a single vector space $V$ to a smooth vector bundle $E \to M$. Instead of a basis $e_{(1)}, \ldots, e_{(n)} \in V$, one then considers a *local frame* $e_{(1)}, \ldots, e_{(n)} \in \Gamma(E|_{\mathcal{U}})$, i.e. a set of sections over some subset $\mathcal{U} \subset M$ which give bases of the fibers $E_p$ for $p \in \mathcal{U}$. (Note that frames over the entirety of $M$ do not exist unless $E$ is globally trivializable.) Any section $v \in \Gamma(E)$ can then be written locally over $\mathcal{U}$ in terms of its *component functions* $v^i : \mathcal{U} \to \mathbb{F}$ as

$$v(p) = v^i(p) e_{(i)}(p),$$

and the same remarks apply to sections of the dual bundle $E^* \to M$, other tensor bundles $E^k_\ell \to M$, direct sums, tensor products and so forth. In particular, tensors now become *tensor fields*, assigning to each $p \in M$ a tensor on the fiber $E_p$ in a manner that varies smoothly with $p$. As important examples, introducing smoothly varying inner products on the fibers gives rise to a *bundle metric* (these always exist), and if $E$ has rank $m$, a *volume form* is now a smooth nowhere zero section of $\Lambda^m E$ (these exist if and only if $E \to M$ is orientable). Operations such as the tensor product, wedge product, tensor contractions, musical isomorphisms and the Hodge star are defined on bundles exactly the same way as on an individual vector space; note that the last two require first fixing a bundle metric, and the last also requires an orientation.

For the tangent bundle $TM \to M$, there is a special class of local frames defined by local coordinate charts. Recall that if $M$ is a smooth $n$-manifold, a *chart* is a diffeomorphism $x : \mathcal{U} \to x(\mathcal{U}) \subset \mathbb{R}^n$ of some open subset $\mathcal{U} \subset M$ to some open subset of $\mathbb{R}^n$, and can be thought of as a set of real-valued functions $x^1, \ldots, x^n : \mathcal{U} \to \mathbb{R}$, the *coordinates*. These define a set of differential operators for smooth functions $\mathcal{U} \to \mathbb{R}$, namely the partial derivatives

$$\partial_j = \frac{\partial}{\partial x^j} : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U}).$$

These are linear operators that satisfy the *Leibnitz rule* $L(fg) = Lf \cdot g + f \cdot Lg$ and are thus *derivations*. Now using the natural identification between vector fields and derivations (see [Spi99]), one can regard $\partial_1, \ldots, \partial_n$ as smooth vector fields over $\mathcal{U}$, i.e. sections of $TM|_{\mathcal{U}}$, which provide a basis of each tangent space $T_pM$ for $p \in \mathcal{U}$. Note that one can also understand these vector fields as the partial derivatives of the inverse map $x^{-1} : x(\mathcal{U}) \hookrightarrow M$, regarded as a smooth function of the $n$ variables $x^1, \ldots, x^n$ on an open subset of $\mathbb{R}^n$. Indeed, for $p \in \mathcal{U}$ one has

$$\partial_j|_p = \frac{\partial x^{-1}}{\partial x^j}(x(p)).$$

In any case, $(\partial_1, \ldots, \partial_n)$ defines a frame for $TM$ over $\mathcal{U}$. The corresponding dual basis at each point $p \in \mathcal{U}$ gives a frame for the cotangent bundle $T^*M$ over $\mathcal{U}$, and it's easy to check that this frame consists precisely of the differentials of the coordinate functions:

$$dx^1, \ldots, dx^j \in \Gamma(T^*M|_{\mathcal{U}}).$$

Any tensor field $T$ of type $(k, \ell)$ over $\mathcal{U}$ can then be written via its component functions ${T^{i_1 \ldots i_k}}_{j_1 \ldots j_\ell} \in C^\infty(\mathcal{U})$ as

$$T = {T^{i_1 \ldots i_k}}_{j_1 \ldots j_\ell}\ dx^{j_1} \otimes \ldots \otimes dx^{j_\ell} \otimes \partial_{i_1} \otimes \ldots \otimes \partial_{i_k},$$

and differential $k$-forms $\omega \in \Omega^k(\mathcal{U}) = \Gamma(\Lambda^k T^*M|_{\mathcal{U}})$ take the form

$$\omega = \sum_{i_1 < \ldots < i_k} \omega_{i_1 \ldots i_k}\ dx^{i_1} \wedge \ldots \wedge dx^{i_k},$$

with smooth component functions $\omega_{i_1 \ldots i_k} \in C^\infty(\mathcal{U})$ that are antisymmetric with respect to exchange of indices.

If $M$ has a Riemannian metric $g$ (i.e. a bundle metric on $TM$) and an orientation, then there is a natural volume form $\mu \in \Omega^n(M) = \Gamma(\Lambda^n T^*M)$ and a Hodge star operator

$$* : \Lambda^k T^*M \to \Lambda^{n-k} T^*M$$

such that using the natural bundle metric on $\Lambda^k T^*M$ induced by $g$, $\alpha \wedge {*\beta} = g(\alpha, \beta)\ \mu$. One can combine this construction with Stokes' theorem to define a *formal adjoint* of the exterior derivative $d : \Omega^k(M) \to \Omega^{k+1}(M)$. This starts from the observation that if $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^{n-k-1}(M)$

are any two forms with compact support, then one can choose a compact $n$-dimensional submanifold $\mathcal{U} \to M$ with boundary such that $\alpha$ and $\beta$ have support in $\mathcal{U}$ and both vanish on $\partial\mathcal{U}$, thus

$$\begin{aligned} 0 &= \int_{\partial\mathcal{U}} \alpha \wedge \beta = \int_{\mathcal{U}} d(\alpha \wedge \beta) = \int_M d(\alpha \wedge \beta) \\ &= \int_M d\alpha \wedge \beta + (-1)^k \int_M \alpha \wedge d\beta. \end{aligned} \tag{A.15}$$

This is essentially the $n$-dimensional manifold version of integration by parts.

**Exercise A.31.** Suppose $M$ is an oriented Riemannian manifold with metric $g$ and induced volume form $\mu$. Defining the operator $\delta : \Omega^k(M) \to \Omega^{k-1}(M)$ by

$$\delta = (-1)^{n(k+1)+1} * d*,$$

use (A.14) and (A.15) to show that for any $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^{k-1}(M)$, both with compact support,

$$\int_M g(\alpha, d\beta)\, \mu = \int_M g(\delta\alpha, \beta)\, \mu.$$

*Remark* A.32. Even though the two bundles $\Lambda^k T^* M$ and $\Lambda^{n-k} T^* M$ have the same rank, it was by no means obvious that they are isomorphic as bundles; rather we've proved this by constructing the Hodge star operator. In fact, it's easy to see that there generally is no such isomorphism if $M$ is not orientable—in particular $\Lambda^0 T^* M$ is always a trivial line bundle by definition, while the line bundle $\Lambda^n T^* M$ is trivial if and only if $M$ is orientable, so if not there can be no bundle isomorphism $\Lambda^0 T^* M \to \Lambda^n T^* M$. This is a major difference between the theories of vector spaces and vector bundles, and the reason why it's often important to observe that certain vector space isomorphisms are "canonical" or "natural" and others are not. As a rule, any given prescription for defining an isomorphism of vector spaces will give a similar isomorphism of vector bundles if and only if the prescription is canonical: that means it can be defined without reference to objects such as local frames, which do not necessarily exist globally on a bundle. So for instance, the existence of bundle metrics, together with our definition of the musical isomorphism $V \to V^* : v \mapsto v^\flat$ on a vector space, gives rise to a bundle isomorphism $E \to E^*$ for any real vector bundle. Note that this does *not* generally work for complex bundles (because the musical isomorphism is not linear but antilinear), and in general complex vector bundles are not isomorphic to their dual bundles.

# References

[Car] S. M. Carroll, *Lecture Notes on General Relativity*. available at http://pancake.uchicago.edu/~carroll/notes/.

[GHL04] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian geometry*, 3rd ed., Springer-Verlag, Berlin, 2004.

[Spi99] M. Spivak, *A comprehensive introduction to differential geometry*, 3rd ed., Vol. 1, Publish or Perish Inc., Houston, TX, 1999.

[Str80] G. Strang, *Linear algebra and its applications*, 2nd ed., Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1980.