TENSORS
(DRAFT COPY)

LARRY SUSANKA

Abstract. The purpose of this note is to define tensors with respect to a fixed finite dimensional real vector space and indicate what is being done when one performs common operations on tensors, such as contraction and raising or lowering indices. We include discussion of relative tensors, inner products, symplectic forms, interior products, Hodge duality and the Hodge star operator and the Grassmann algebra.

All of these concepts and constructions are extensions of ideas from linear algebra including certain facts about determinants and matrices, which we use freely. None of them requires additional structure, such as that provided by a differentiable manifold.

Sections 2 through 11 provide an introduction to tensors. In sections 12 through 25 we show how to perform routine operations involving tensors. In sections 26 through 28 we explore additional structures related to spaces of alternating tensors.

Our aim is modest. We attempt only to create a very structured development of tensor methods and vocabulary to help bridge the gap between linear algebra and its (relatively) benign notation and the vast world of tensor applications. We (attempt to) define everything carefully and consistently, and this is a concise repository of proofs which otherwise may be spread out over a book (or merely referenced) in the study of an application area.

Many of these applications occur in contexts such as solid-state physics or electrodynamics or relativity theory. Each subject area comes equipped with its own challenges: subject-specific vocabulary, traditional notation and other conventions. These often conflict with each other and with modern mathematical practice, and these conflicts are a source of much confusion.

Having read these notes the vocabulary, notational issues and constructions of raw tensor methods, at least, will likely be less of an obstacle.

Date: September 6, 2018.

Contents

1. Some Notation

In these notes we will be working with a few sets and functions repeatedly, so we lay out these **critters** up front. The page number references their introduction.

$\mathbb{R}^n$ all $n \times 1$ "column matrices" with real entries (*Page 6.*)
Standard ordered basis: $\mathbf{e}$ given by $e_1, \ldots, e_n$

$\mathbb{R}^{n*}$ all $1 \times n$ "row matrices" with real entries (*Page 7.*)
Standard dual ordered basis: $\mathbf{e}^*$ given by $e^1, \ldots, e^n$

The standard inner product on $\mathbb{R}^n$ generates the **Euclidean geometry** on $\mathbb{R}^n$.
Denoted $x \cdot y$ or $\langle x, y \rangle$ or $x^t y$ for $x, y \in \mathbb{R}^n$ ("t" indicates **transpose**.)

V generic real n-dimensional vector space
with two ordered bases $\mathbf{a}$ given by $a_1, \ldots, a_n$ and $\mathbf{b}$ given by $b_1, \ldots, b_n$

$A \colon V \to \mathbb{R}^n$ defined by $v = \sum_{i=1}^n A^i(v) a_i$ for any $v \in V$ (*Page 8.*)
$B \colon V \to \mathbb{R}^n$ defined by $v = \sum_{i=1}^n B^i(v) b_i$ for any $v \in V$ (*Page 10.*)

Matrix M has ijth entry M^i_j. M itself can be denoted $\left(M^i_j\right)$.

Define matrix $\mathcal{A}$ by $\mathcal{A}^i_j = A^i(b_j)$. Note $b_j = \sum_{i=1}^n A^i(b_j) a_i = \sum_{i=1}^n \mathcal{A}^i_j a_i$.
Define matrix $\mathcal{B}$ by $\mathcal{B}^i_j = B^i(a_j)$. Note $a_j = \sum_{i=1}^n B^i(a_j) b_i = \sum_{i=1}^n \mathcal{B}^i_j b_i$.
$\mathcal{A}$ and $\mathcal{B}$ are called **matrices of transition**. (*Page 10.*)

We calculate that $\mathcal{B}\mathcal{A} = \mathcal{I} = \left(\delta^i_j\right)$, the $n \times n$ identity matrix. (*Page 10.*)

Suppose $v = \sum_{i=1}^n x^i a_i = \sum_{i=1}^n y^i b_i \in V$. (*Page 11.*)
Then $x^i = \sum_{j=1}^n y^j \mathcal{A}^i_j = \sum_{j=1}^n y^j A^i(b_j)$
and $y^i = \sum_{j=1}^n x^j \mathcal{B}^i_j = \sum_{j=1}^n x^j B^i(a_j)$.

If x and y in $\mathbb{R}^n$ represent $v \in V$ in bases $\mathbf{a}$ and $\mathbf{b}$ respectively (*Page 11*)
then $y = \mathcal{B}x$ and $\mathcal{B} = \left(\frac{\partial y^i}{\partial x^j}\right)$ and $x = \mathcal{A}y$ and $\mathcal{A} = \left(\frac{\partial x^i}{\partial y^j}\right)$.

The entries of $\mathcal{B}$ represent the rate of change of a $\mathbf{b}$ coordinate with respect to variation of an $\mathbf{a}$ coordinate, when other $\mathbf{a}$ coordinates are left unchanged.

V^* all real linear functionals on V
$\mathbf{a}^*$ given by $a^1, \ldots, a^n$: ordered basis of V^* dual to $\mathbf{a}$
$\mathbf{b}^*$ given by $b^1, \ldots, b^n$: ordered basis of V^* dual to $\mathbf{b}$
We show $a^j = \sum_{i=1}^n A^j(b_i) b^i$ and $b^j = \sum_{i=1}^n B^j(a_i) a^i$ (*Page 12.*)

Suppose $\theta = \sum_{j=1}^n g_j a^j = \sum_{j=1}^n f_j b^j \in V^*$. (*Page 12.*)
Then $g_j = \sum_{i=1}^n f_i \mathcal{B}^i_j = \sum_{i=1}^n f_i B^i(a_j)$ and $f_j = \sum_{i=1}^n g_i \mathcal{A}^i_j = \sum_{i=1}^n g_i A^i(b_j)$.

If τ and σ in $\mathbb{R}^{n*}$ represent $\theta \in V^*$ in bases $\mathbf{a}^*$ and $\mathbf{b}^*$ respectively (*Page 13*)
then $\sigma = \tau\mathcal{A}$ and $\tau = \sigma\mathcal{B}$.

H^* (**pushforward**) and H_* (**pullback**) (*Page 26.*)

Bilinear form G has matrix $\mathbf{G_a} = \left({\mathbf{G_a}}^i_j\right) = (G_{i,j}(\mathbf{a}))$ (*Page 45.*)
If invertible, define $G^{i,j}(\mathbf{a})$ by ${\mathbf{G_a}}^{-1} = \overline{\mathbf{G_a}} = \left({\overline{\mathbf{G_a}}}^{\,i}_{\,j}\right) = \left(G^{i,j}(\mathbf{a})\right)$ (*Page 47.*)

2. $\mathbb{R}^n, \mathbb{R}^{n*}$ and $\mathbb{R}^{n**}$

We presume throughout that the reader has seen the basics of **linear algebra**, including, at least, the concepts of **ordered basis** and **dimension** of a **real vector space** and the fact that a **linear transformation** is determined by its effect on an ordered basis of its domain. Facts about symmetric and skew symmetric matrices and how they can be brought to standard forms and several facts about determinants are referenced and used. Beyond that, the reader need only steel him-or-her-self for the index-fest which (no way around it) ensues.

Rest assured that with some practice you too will be slinging index-laden monstrosities about with the aplomb of a veteran.

We represent members of the real vector space $\mathbb{R}^n$ as $n \times 1$ "column matrices"

$$x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \sum_{i=1}^{n} x^i \, e_i$$

$$\text{where for each } i = 1, \ldots, n \text{ we define } \mathbf{e_i} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow i\text{th row.}$$

A real number "is" a 1×1 matrix.

We abandon the standard practice in basic calculus and algebra textbooks whereby members of $\mathbb{R}^n$ are denoted by rows. The typographical convenience of this practice for the authors of these books is inarguable, saving the reams of paper in every elementary class which would be devoted to whitespace in the textbook using our convention. Our choice is the one most commonly adopted for the subject at hand. Here, rows of numbers will have a different meaning, to be defined shortly.

Superscripts and subscripts abound in the discussion while exponents are scarce, so it should be presumed that a notation such as x^i refers to the ith coordinate or ith instance of x rather than the ith power of x.

Any $x \in \mathbb{R}^n$ has one and only one representation as $\sum_{j=1}^n x^j e_j$ where the x^j are real numbers. The ordered list $e_1, e_2, \ldots, e_n$ is called an **ordered basis** of $\mathbb{R}^n$ by virtue of this property, and is called the **standard ordered basis of** $\mathbf{\mathbb{R}^n}$, denoted $\mathbf{e}$.

Any linear transformation σ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is identified with an $m \times n$ matrix $\left(\sigma^i_j\right)$ with ith row and jth column entry given by $\sigma^{\mathbf{i}}_{\mathbf{j}} = \sigma^i(e_j)$. Then

$$\sigma(x) = \begin{pmatrix} \sigma^1(x) \\ \vdots \\ \sigma^m(x) \end{pmatrix} = \sum_{i=1}^m \sigma^i(x) e_i = \sum_{i=1}^m \sigma^i \left(\sum_{j=1}^n x^j e_j \right) e_i$$
$$= \sum_{i=1}^m \left(\sum_{j=1}^n \sigma^i(e_j) x^j \right) e_i = \left(\sigma^i_j\right) x.$$

So a linear transformation σ "is" left matrix multiplication by the $m \times n$ matrix $\left(\sigma^i_j\right) = \left(\sigma^i(e_j)\right)$. We will usually not distinguish between a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ and this matrix, writing either $\sigma(x)$ or σx for σ evaluated at x.

Be aware, however, that a linear transformation is a *function* while matrix multiplication by a certain matrix is an arrangement of arithmetic operations which can be employed, when done in specific order, to *evaluate* a function at one of its domain members.

With this in mind, the set of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}$, also called the **dual of $\mathbb{R}^{\mathbf{n}}$**, is identified with the set of $1 \times n$ "row matrices," which we will denote $\mathbb{R}^{n*}$. Individual members of $\mathbb{R}^{n*}$ are called **linear functionals on $\mathbb{R}^{\mathbf{n}}$**.

Any linear functional σ on $\mathbb{R}^n$ is identified with the row matrix

$$(\sigma_1, \sigma_2, \ldots, \sigma_n) = (\sigma(e_1), \sigma(e_2), \ldots, \sigma(e_n))$$

acting on $\mathbb{R}^n$ by left matrix multiplication.

$$\sigma(x) = \sum_{j=1}^n \sigma(e_j) x^j = (\sigma_1, \sigma_2, \ldots, \sigma_n) \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \sigma x.$$

The entries of row matrices are distinguished from each other by commas. These commas are not strictly necessary, and can be omitted. But without them there might be no obvious visual cue that you have moved to the next column: $(3 \; -2x \; - 4 - y)$ versus $(3, \; -2x, \; -4 - y)$.

$\mathbb{R}^{n*}$ itself has ordered basis $\mathbf{e}^*$ given by the list $e^1, e^2, \ldots, e^n$ where each $\mathbf{e}^{\mathbf{i}} = (e_i)^t$, the transpose of e_i. The ordered basis $\mathbf{e}^*$ is called the **standard ordered basis of $\mathbb{R}^{\mathbf{n}*}$**.

Any $\sigma \in \mathbb{R}^{n*}$ is $\sum_{j=1}^n \sigma(e_j) e^j$ and this representation is the only representation of σ as a nontrivial linear combination of the e^j.

The **"double dual" of $\mathbb{R}^{\mathbf{n}}$**, denoted $\mathbb{R}^{n**}$, is the set of linear transformations from $\mathbb{R}^{n*}$ to $\mathbb{R}$. These are also called linear functionals on $\mathbb{R}^{n*}$. This set of functionals is identified with $\mathbb{R}^n$ by the **evaluation map $\mathbf{E}\colon \mathbb{R}^{\mathbf{n}} \to \mathbb{R}^{\mathbf{n}**}$** defined for $x \in \mathbb{R}^n$ by

$$E(x)(\sigma) = \sigma(x) = \sigma x \;\; \forall \; \sigma \in \mathbb{R}^{n*}.$$

Note that if $E(x)$ is the 0 transformation then x must be 0, and since $\mathbb{R}^n$ and $\mathbb{R}^{n**}$ have the same dimension E defines an isomorphism onto $\mathbb{R}^{n**}$.

We will not usually distinguish between $\mathbb{R}^n$ and $\mathbb{R}^{n**}$. Functional notations $x(\sigma)$ and $\sigma(x)$ and matrix notation $\sigma\, x$ are used interchangeably for $x \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}^{n*}$.

Note that the matrix product $x\,\sigma$ is an $n \times n$ matrix, not a real number. It will normally be obvious from context if we intend functional evaluation $x(\sigma)$ or this matrix product. Members of $\mathbb{R}^n = \mathbb{R}^{n**}$ "act on" members of $\mathbb{R}^{n*}$ by matrix multiplication *on the right.*

3. V, V^* and V^{**}

If V is a generic n-dimensional real vector space with ordered basis $\mathbf{a}$ given by $a_1, \ldots, a_n$ there is a unique function $\mathbf{A}\colon V \to \mathbb{R}^n$ defined by

$$A(v) = \begin{pmatrix} A^1(v) \\ \vdots \\ A^n(v) \end{pmatrix} = \sum_{i=1}^n A^i(v) e_i \quad \text{when} \quad v = \sum_{i=1}^n A^i(v) a_i.$$

A is called the **coordinate map for the ordered basis a**, and $A(v)$ is said to **represent v in ordered basis a**. The individual entries in $A(v)$ are called "the first coordinate," "the second coordinate" and so on. The word **component** is used synonymously with coordinate.

The function A associates $\sum_{i=1}^n A^i(v) a_i \in V$ with $\sum_{i=1}^n A^i(v) e_i \in \mathbb{R}^n$.

A is an isomorphism of V onto $\mathbb{R}^n$. Denote its inverse by $\overline{\mathbf{A}}$.

$$\overline{A}(x) = \sum_{i=1}^n x^i a_i \in V \quad \text{when} \quad x = \sum_{i=1}^n x^i e_i \in \mathbb{R}^n.$$

A real valued linear transformation on V is called a **linear functional on V**. The set of all these is denoted $\mathbf{V}^*$ and called the **dual of V**. V^* has dimension n and has ordered basis $\mathbf{a}^*$ given by $a^1, \ldots, a^n$ where each a^i is defined by

$$a^j(v) = A^j(v) \quad \text{when} \quad v = \sum_{i=1}^n A^i(v) a_i.$$

In other words, a^j "picks off" the jth coordinate of $v \in V$ when you represent v in terms of the ordered basis $\mathbf{a}$.

This ordered basis is said to be the ordered basis **dual to a**.

If you know that $\mathbf{a_1}$ is a member of an ordered basis, but do not know $\mathbf{a_2}, \ldots, \mathbf{a_n}$ you cannot determine $\mathbf{a^1}$. The coefficient of a_1 in the sum $v = \sum_{i=1}^n A^i(v) a_i$ depends on *every* member of the ordered basis, not just a_1.

If $\sigma \in \mathbb{R}^{n*}$ define $A^*(\sigma) \in V^*$ by $A^*(\sigma)(v) = \sigma(A(v))$.

$$A^*(\sigma) = \sum_{i=1}^n \sigma_i a^i \in V^* \quad \text{when} \quad \sigma = \sum_{i=1}^n \sigma_i e^i \in \mathbb{R}^{n*}.$$

$\mathbf{A}^*$ is an isomorphism of $\mathbb{R}^{n*}$ onto V^*. We will denote its inverse by $\overline{\mathbf{A}^*}$.

$$\overline{A^*}(\theta) = \sum_{i=1}^{n} g_i e^i \quad \text{when} \quad \theta = \sum_{i=1}^{n} g_i a^i .$$

$\overline{A^*}$ is called the **coordinate map** for ordered basis $\mathbf{a}^*$. The row matrix $\overline{A^*}(\theta) = \left(\overline{A^*}_1(\theta), \ldots, \overline{A^*}_n(\theta) \right)$ is said to **represent θ in the dual ordered basis $\mathbf{a}^*$**.

The four isomorphisms described here are represented pictorially below.

$$a_i \leftrightarrow e_i \qquad \begin{tikzcd} V \arrow[d, shift left, "A"] \\ \mathbb{R}^n \arrow[u, shift left, "\overline{A}"] \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} V^* \arrow[d, shift left, "\overline{A^*}"] \\ \mathbb{R}^{n*} \arrow[u, shift left, "A^*"] \end{tikzcd} \qquad a^i \leftrightarrow e^i$$

To keep track of this **small menagerie** of "A"s remember: an even number of "hats" on A sends a member of V or V^* to its representative in $\mathbb{R}^n$ or $\mathbb{R}^{n*}$.

These are not merely isomorphisms. They are isomorphisms coordinated in a way that allows you to evaluate functionals on vectors by moving them down to $\mathbb{R}^n$ and $\mathbb{R}^{n*}$ and using their images under these isomorphisms to perform the calculation.

So for $\theta \in V^*$, $\sigma \in \mathbb{R}^{n*}$, $v \in V$ and $x \in \mathbb{R}^n$:

$$\overline{A^*}(\theta) A(v) = \theta(v) \quad \text{and} \quad A^*(\sigma) \left(\overline{A}(x) \right) = \sigma x .$$

V^{**}, the "double dual" of V, is the set of linear functionals on V^*.

As before, the **evaluation map**, defined for each $v \in V$ by

$$E(v)(\theta) = \theta(v) \ \forall \ \theta \in V^*,$$

provides an isomorphism of V onto V^{**} and we will not normally distinguish between members of V and members of V^{**}. The notations $v(\theta)$ and $\theta(v)$ are used interchangeably for $v \in V$ and $\theta \in V^*$.

Our motivation for creating these two structures (one involving V, the other $\mathbb{R}^n$) bears some examination. Suppose we are using vector methods to help us understand a question from physics, such as surface tension on a soap bubble or deformation in a crystal lattice. In these applications a vector space V can arise which is natural for the problem. Physics is simple in V. If V itself is simple (we don't intend to rule out the possibility that $V = \mathbb{R}^n$) all is well. But often V is complicated, perhaps a quotient space in a function space or a subspace of $\mathbb{R}^k$ for larger dimension k. The representation of V in $\mathbb{R}^n$ provided by an ordered basis for V, however, may be easier to understand and provides a venue for number crunching. The isomorphisms bring the necessary geometry from V down to $\mathbb{R}^n$ where we can work undisturbed by irrelevant aspects of V. You choose where you want to work, picking the description to match what you are trying to do.

4. Change of Basis

If V possesses two ordered bases **a** given by $a_1, \dots, a_n$ and **b** given by $b_1, \dots, b_n$, we want to understand how the isomorphism $A\colon V \to \mathbb{R}^n$ and its inverse $\overline{A}$ are related to the analogous isomorphisms B and $\overline{B}$ given by

$$B(v) = \begin{pmatrix} B^1(v) \\ \vdots \\ B^n(v) \end{pmatrix} \quad \text{where} \quad v = \sum_{i=1}^n B^i(v) b_i.$$

Specifically, if x and y in $\mathbb{R}^n$ represent the same vector v in V with respect to different bases a and b respectively, how are x and y related?

Let $\mathcal{A}$ be the matrix with ith row and jth column entry $\mathcal{A}^i_j = A^i(b_j)$ and let $\mathcal{B}$ be the matrix with ith row and jth column entry $\mathcal{B}^i_j = B^i(a_j)$. The columns of $\mathcal{A}$ are the coordinates of the **b** basis vectors in terms of the **a** basis vectors. The columns of $\mathcal{B}$ are the coordinates of the **a** basis vectors in terms of the **b** basis vectors.

We call $\mathcal{A}$ the **matrix of transition from b to a** and $\mathcal{B}$ the **matrix of transition from a to b**. They will be used to aid the translation between representations of vectors in terms of the bases **a** and **b**.

Crudely (and possibly incorrectly when there is no notion of distance in V) speaking, if all the vectors of **b** are longer than those of **a**, the entries of $\mathcal{B}$ will all be smaller than 1 because the coordinates of vectors must become smaller in the new basis **b**. Conversely, if the members of **b** are all short the entries of $\mathcal{B}$ will be large.

For each i

$$\begin{aligned} b_i &= \sum_{j=1}^n A^j(b_i) a_j = \sum_{j=1}^n A^j(b_i) \sum_{k=1}^n B^k(a_j) b_k \\ &= \sum_{k=1}^n \left(\sum_{j=1}^n B^k(a_j) A^j(b_i) \right) b_k. \end{aligned}$$

By uniqueness of coefficients we have, for each k,

$$\sum_{j=1}^n B^k(a_j) A^j(b_i) = \sum_{j=1}^n \mathcal{B}^k_j \mathcal{A}^j_i = \delta^k_i \equiv \begin{cases} 0, & \text{if } i \neq k; \\ 1, & \text{if } i = k. \end{cases}$$

This means that the matrices $\mathcal{A}$ and $\mathcal{B}$ are inverse to each other.[1]

$\mathcal{B}\mathcal{A} = \mathcal{I}$, where $\mathcal{I}$ is the $n \times n$ identity matrix $\left(\delta^i_j\right)$.

So if $A(v) = x$ and $B(v) = y$ then v must be equal to

$$\sum_{j=1}^n A^j(v) a_j = \sum_{j=1}^n A^j(v) \sum_{i=1}^n B^i(a_j) b_i = \sum_{i=1}^n \left(\sum_{j=1}^n B^i(a_j) A^j(v) \right) b_i = \sum_{i=1}^n B^i(v) b_i.$$

[1] The size of the identity matrix is usually suppressed. If that might cause confusion the notation $\mathcal{I}(n) = \left(\delta^i_j(n)\right)$ can be used. The symbol δ^i_j is called the **Kronecker delta**.

So for each i,

$$y^i = B^i(v) = \sum_{j=1}^n B^i(a_j)A^j(v) = \sum_{j=1}^n B^i(a_j)x^j.$$

$$y = \mathcal{B}x \quad \text{if } A(v) = x \text{ and } B(v) = y.$$

The representation in $\mathbb{R}^n$ of a vector in V created using the ordered basis $\mathbf{b}$ is obtained from the representation using $\mathbf{a}$ by left multiplying the representation by the matrix $\mathcal{B}$.

Another way of expressing this fact is the following, for each i:

$$y^i = \sum_{j=1}^n x^j \mathcal{B}^i_j = \sum_{j=1}^n x^j B^i(a_j) \quad \text{when} \quad v = \sum_{j=1}^n x^j a_j = \sum_{j=1}^n y^j b_j \in V.$$

Thought of as a linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$, the relationship $y = \mathcal{B}x$ is differentiable and we find that the derivative is $\mathcal{B} = \left(\frac{\partial y^i}{\partial x^j}\right)$. Similarly, $\mathcal{A} = \left(\frac{\partial x^i}{\partial y^j}\right)$. These are called the **Jacobian matrices** of the linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$ given by the coordinate change.

Repeating for emphasis, a vector v has coordinates in both basis $\mathbf{a}$ and basis $\mathbf{b}$. If you change the value of *one* coefficient a little, say on basis vector a_j, you change the vector v and this new vector could have slight changes to *all* of its coefficients in basis $\mathbf{b}$.

The ijth entry of $\mathcal{B}$ is the rate of change of the ith coordinate of a vector in basis $\mathbf{b}$, the coefficient on basis vector b_i, with respect to changes in the jth coordinate of the vector in basis $\mathbf{a}$, other $\mathbf{a}$-coordinates remaining constant.

It is for this reason a symbolic notation such as

$$\frac{d\boldsymbol{b}}{d\boldsymbol{a}} = \mathcal{B} = \text{“from } \mathbf{a} \text{ to } \mathbf{b}\text{”} \quad \text{and} \quad \frac{d\boldsymbol{a}}{d\boldsymbol{b}} = \mathcal{A} = \text{“from } \mathbf{b} \text{ to } \mathbf{a}\text{”}$$

is sometimes used, making explicit reference to the bases involved. This reminder is particularly convenient when many bases are in play.

Some texts use the “derivative” notation for these matrices, and the partial derivative notation for their entries, throughout. We do not. The reason for this is purely typographical. When I change to this partial derivative notation many of the equations scattered throughout this text are harder to look at, and we are only going to be referring to two bases in these notes so a more flexible (and specific) notation is not required.

The partial derivative notation is used in applications involving differentiable manifolds, where it is appropriate and convenient.

Next, we create the ordered basis $\mathbf{b}^*$ with members $b^1, \ldots, b^n$ dual to $\mathbf{b}$ and defined for each j by

$$b^j(v) = B^j(v) \quad \text{when} \quad v = \sum_{i=1}^n B^i(v) b_i.$$

b^j "picks off" the jth coordinate of $v \in V$ when you represent v in terms of the ordered basis $\mathbf{b}$.

We create isomorphisms B^* and $\overline{B^*}$ as before and indicated in the picture below.

$$\begin{array}{ccccccccc} & & V & & & & V^* & & \\ b_i \leftrightarrow e_i & \overline{B} & \uparrow \ \downarrow & B & \text{and} & B^* & \uparrow \ \downarrow & \overline{B^*} & b^i \leftrightarrow e^i \\ & & \mathbb{R}^n & & & & \mathbb{R}^{n*} & & \end{array}$$

We will need to know how to represent each functional b^j in terms of $\mathbf{a}^*$.

Suppose $b^j = \sum_{i=1}^n C_i^j a^i$ for $j = 1, \ldots, n$. Then:

$$\begin{aligned} \delta_k^j =& b^j(b_k) = \sum_{i=1}^n C_i^j a^i(b_k) \\ =& \sum_{i=1}^n C_i^j a^i \left(\sum_{w=1}^n A^w(b_k) a_w \right) \\ =& \sum_{i=1}^n C_i^j A^i(b_k) = \sum_{i=1}^n C_i^j \mathcal{A}_k^i. \end{aligned}$$

That means the matrix C is the inverse of $\mathcal{A}$. We already know that $\mathcal{A}^{-1} = \mathcal{B}$: that is, $C = \mathcal{B}$. A symmetrical result holds for a description of the a^j in terms of $\mathbf{b}^*$.

To reiterate: for each j

$$b^j = \sum_{i=1}^n B^j(a_i) a^i = \sum_{i=1}^n \mathcal{B}_i^j a^i \quad \text{and} \quad a^j = \sum_{i=1}^n A^j(b_i) b^i = \sum_{i=1}^n \mathcal{A}_i^j b^i.$$

Now suppose we have a member $\theta = \sum_{j=1}^n f_j b^j = \sum_{j=1}^n g_j a^j$ in V^*.

$$\theta = \sum_{k=1}^n g_k a^k = \sum_{k=1}^n g_k \sum_{j=1}^n A^k(b_j) b^j = \sum_{j=1}^n \left(\sum_{k=1}^n g_k A^k(b_j) \right) b^j = \sum_{j=1}^n \left(\sum_{k=1}^n g_k \mathcal{A}_j^k \right) b^j.$$

This means that for each j

$$f_j = \sum_{i=1}^n g_i A^i(b_j) = \sum_{i=1}^n g_i \mathcal{A}_j^i \quad \text{when} \quad \theta = \sum_{j=1}^n f_j b^j = \sum_{j=1}^n g_j a^j.$$

We can look at the same calculation from a slightly different standpoint. Suppose $\sigma, \tau \in \mathbb{R}^{n*}$ represent the same linear transformation $\theta \in V^*$ with respect to bases $\mathbf{b}^*$ and $\mathbf{a}^*$ respectively.

Specifically, $\theta(v) = \sigma B(v) = \tau A(v)\ \forall\ v \in V$. So for each i

$$\begin{aligned}
\sigma_i = \sigma e_i = \sigma B(b_i) &= \theta(b_i) \\
&= \theta\left(\sum_{j=1}^{n} A^j(b_i)a_j\right) = \sum_{j=1}^{n} A^j(b_i)\theta(a_j) \ = \ \sum_{j=1}^{n} A^j(b_i)\tau A(a_j) \\
&= \sum_{j=1}^{n} A^j(b_i)\tau e_j \ = \ \sum_{j=1}^{n} \tau_j A^j(b_i) \ = \ \sum_{j=1}^{n} \tau_j \mathcal{A}^j_i.
\end{aligned}$$

$$\sigma = \tau\mathcal{A} \quad \text{if } \overline{A^*}(\theta) = \tau \text{ and } \overline{B^*}(\theta) = \sigma.$$

The representation in $\mathbb{R}^{n*}$ of a linear transformation in V^* created using the ordered basis $\mathbf{b}^*$ is obtained from the representation using $\mathbf{a}^*$ by right multiplying the representation by the matrix $\mathcal{A}$.

All these arrows are reversible.

$$\begin{tikzcd}
 & V \arrow[dl, "A"'] \arrow[dr, "B"] & \\
\mathbb{R}^n \arrow[rr, "\text{Left multiply}"', "\text{by matrix } \mathcal{B}"'] & & \mathbb{R}^n
\end{tikzcd}
\qquad
\begin{tikzcd}
 & V^* \arrow[dl, "\overline{A^*}"'] \arrow[dr, "\overline{B^*}"] & \\
\mathbb{R}^{n*} \arrow[rr, "\text{Right multiply by matrix } \mathcal{A}"'] & & \mathbb{R}^{n*}
\end{tikzcd}$$

Left multiply by matrix $\mathcal{B}$

Right multiply by matrix $\mathcal{A}$

We now raise an important point on change of basis. $\mathbb{R}^n$ and $\mathbb{R}^{n*}$ are vector spaces and we do not intend to preclude the possibility that V or V^* could be either of these. Still, when $\mathbb{R}^n$ and $\mathbb{R}^{n*}$ are considered as range of A and $\overline{A^*}$ **we will never change basis there, but always use e and e*.** Though the representative of a vector from V or V^* might change as we go along, we take the point of view that this happens **only as a result of a change of basis in V**.

5. The Einstein Summation Convention

At this point it becomes worthwhile to use a common notational contrivance: in an indexed sum of products, it often (*very often*) happens that an index symbol being summed over occurs once as subscript and once as superscript in each product in the sum. The formula $b_j = \sum_{i=1}^{n} \mathcal{A}^i_j a_i$ for $j = 1, \ldots, n$ provides an example.

Henceforth, if you see a formula in these notes involving a product within which one subscript and one superscript are indicated by the same symbol then a summation over all the values of that index is presumed. If an index is not used as an index of summation it is presumed that the formula has distinct instances, one for each potential value of that index.

This is called the **Einstein summation convention.** With this convention the n different formulae found above are represented by $b_j = \mathcal{A}^i_j a_i$, and **if anything else is intended that must be explicitly explained in situ.**

In these notes **we will never use this convention if, within a product, the same symbol is used more than once as subscript or more than once as superscript.**

So for us $a^i x^i$ is not a sum. Neither is $a^i a^i x_i$ or $(a^i + a^i)x_i$. But $(a^i)^2 x_i$ and $a^i x_i + a^i x_i$ *do* indicate summation on i, and the latter is the same as $a^i x_i + a^j x_j$.

If we intend symbols such as $a^i x^i$ to denote a sum we must use the old conventional notation $\sum_{i=1}^n a^i x^i$.

Sometimes the notation can be ambiguous, as in $\left|a_i x^i\right|$. Rather than elaborate on our convention, converting a notational convenience into a mnemonic nuisance, we will simply revert to the old summation notation whenever confusion might occur.

Be aware that many texts which focus primarily on tensor representations involving special types of vector spaces—inner product spaces—do *not* require that a subscript occur "above and below" if summation is intended, only that it be repeated twice in a term. For them, $a_i x_i$ *is* a sum. *We will never use this version of the summation convention.*

6. Tensors and the (Outer) Tensor Product

We define the product space

$$V^r_s = \underbrace{V^* \times \cdots \times V^*}_{r \text{ copies}} \times \underbrace{V \times \cdots \times V}_{s \text{ copies}}$$

where r is the number of times that V^* occurs (they are all listed first) and s is the number of times that V occurs. We will define $V^r_s(k)$ by

$$V^r_s(k) = \begin{cases} V^*, & \text{if } 1 \le k \le r; \\ V, & \text{if } r+1 \le k \le r+s. \end{cases}$$

A function $T\colon V^r_s \to \mathbb{R}$ is called **multilinear** if it is linear in each domain factor separately.

Making this precise is notationally awkward. In fact, to the beginner this whole business looks like a huge exercise in **creative** but only **marginally successful** management of notational messes. Here is an example.

T is multilinear if whenever $v(i) \in V^r_s(i)$ for each i then the function

$$T(v(1), \dots, v(j-1), x, v(j+1), \dots, v(r+s))$$

in the variable $x \in V^r_s(j)$ is linear on $V^r_s(j)$ for $j = 1, \dots, r+s$. Sometimes we say, colloquially, that T is linear in each of its **slots** separately. x in the formula above is said to be in the jth slot of T.

It is, obviously, necessary to **tweak** this condition at the beginning and end of a subscript range, where $j-1$ or $j+1$ might be "out of bounds" and it is left to the reader to do the right thing here, and in similar cases later in the text.

A **tensor**[2] **on V** is any multilinear function T as described above. $r+s$ is called the **order** or **rank** of the tensor T.

T is said to have **contravariant order r** and **covariant order s**. If $s=0$ the tensor is called **contravariant** while if $r=0$ it is called **covariant.** If neither r nor s is zero, T is called a **mixed** tensor.

A covariant tensor of order 1 is simply a linear functional, and in this context is called a **1-form**, a **covariant vector** or a **covector**. A contravariant tensor of order 1 is often called a **contravector** or, simply, a **vector**.

The sum of two tensors defined on V^r_s is a tensor defined on V^r_s, and also a constant multiple of a tensor defined on V^r_s is a tensor defined on V^r_s. So the tensors defined on V^r_s constitute a real vector space denoted $\mathcal{T}^{\mathbf{r}}_{\mathbf{s}}(\mathbf{V})$. For convenience we define $\mathcal{T}^{\mathbf{0}}_{\mathbf{0}}(\mathbf{V})$ to be $\mathbb{R}$. Members of $\mathbb{R}$ are called, variously, **real numbers, constants, invariants or scalars.**

Tensors can be multiplied to form new tensors of higher order.

If $T\colon V^r_s \to \mathbb{R}$ is a tensor and if $\widetilde{T}\colon V^{\widetilde{r}}_{\widetilde{s}} \to \mathbb{R}$ is a tensor then the **tensor product of T with $\widetilde{\mathbf{T}}$ (in that order)** is denoted $\mathbf{T}\otimes\widetilde{\mathbf{T}}$ and defined to be the multilinear function

$$T\otimes\widetilde{T}\colon V^{r+\widetilde{r}}_{s+\widetilde{s}} \to \mathbb{R}$$

given for $(v(1),\dots,v(r+\widetilde{r}+s+\widetilde{s})) \in V^{r+\widetilde{r}}_{s+\widetilde{s}}$ by the product

$$\begin{aligned}
(T\otimes\widetilde{T})&((v(1),\dots,v(r+\widetilde{r}+s+\widetilde{s}))\\
&= T(v(1),\dots,v(r),v(r+\widetilde{r}+1),\dots,v(r+\widetilde{r}+s))\\
&\quad\cdot \widetilde{T}(v(r+1),\dots,v(r+\widetilde{r}),v(r+\widetilde{r}+s+1),\dots,v(r+\widetilde{r}+s+\widetilde{s})).
\end{aligned}$$

Sometimes this product is called the **outer product** of the two tensors. $T\otimes\widetilde{T}$ is of order $r+\widetilde{r}+s+\widetilde{s}$, contravariant of order $r+\widetilde{r}$ and covariant of order $s+\widetilde{s}$.

This process of creating tensors of higher order is definitely not commutative: $T\otimes\widetilde{T} \neq \widetilde{T}\otimes T$ in general. But this process is associative so a tensor represented as $T\otimes S\otimes R$ is unambiguous.

Suppose $\mathbf{a}$ is an ordered basis of V, identified with V^{**}, and $\mathbf{a}^*$ the dual ordered basis of V^*. We use in the calculation below, for the first time, the Einstein summation convention: repeated indices which occur exactly once as superscript and exactly once as subscript in a term connote summation over the range of that index.

If $T \in \mathcal{T}^r_s(V)$ and $v=(v(1),\dots,v(r+s)) \in V^r_s$ then

$$\begin{aligned}
T(v) &= T\big(\overline{A^*}_{i_1}(v(1))a^{i_1},\dots,\overline{A^*}_{i_r}(v(r))a^{i_r},\\
&\qquad\qquad A^{i_{r+1}}(v(r+1))a_{i_{r+1}},\dots,A^{i_{r+s}}(v(r+s))a_{i_{r+s}}\big)\\
&= \overline{A^*}_{i_1}(v(1))\cdots A^{i_{r+s}}(v(r+s))T(a^{i_1},\dots,a_{i_{r+s}})\\
&= T(a^{i_1},\dots,a_{i_{r+s}})a_{i_1}\otimes\cdots\otimes a_{i_r}\otimes a^{i_{r+1}}\otimes\cdots\otimes a^{i_{r+s}}(v).
\end{aligned}$$

[2]The word **tensor** is adapted from the French "tension" meaning "strain" in English. Understanding the physics of deformation of solids was an early application.

From this exercise in multilinearity we conclude that the n^{r+s} tensors of the form $a_{i_1} \otimes \cdots \otimes a_{i_r} \otimes a^{i_{r+1}} \otimes \cdots \otimes a^{i_{r+s}}$ span $\mathcal{T}^r_s(V)$ and it is fairly easy to show **they constitute a linearly independent set and hence a basis, the standard basis for** $\mathcal{T}^r_s(V)$ for ordered basis $\mathbf{a}$.

The n^{r+s} indexed numbers $T(a^{i_1}, \dots, a_{i_{r+s}})$ determine the tensor T and we say that **this indexed collection of numbers "is" the tensor T in the ordered basis a**.

We introduce the notation

$$T^{i_1,\dots,i_r}_{i_{r+1},\dots,i_{r+s}}(\mathbf{a}) = T(a^{i_1}, \dots, a^{i_r}, a_{r_1+1}, \dots, a_{i_{r+s}}).$$

Alternatively, after letting $j_k = i_{r+k}$ for $k = 1, \dots, s$ we define

$$T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a}) = T(a^{i_1}, \dots, a^{i_r}, a_{j_1}, \dots, a_{j_s}).$$

For each ordered basis $\mathbf{a}$, $T^{\dots}_{\dots}(\mathbf{a})$ is a real valued function whose domain consists of all the ordered $r+s$-tuples of integers from the index set and whose value on any given index combination is shown above.

So as a multiple sum over $r+s$ different integer indices

$$T \;=\; T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\; a_{i_1} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}.$$

A tensor which is the tensor product of r different vectors and s different covectors is of order $t = r + s$. Tensors which can be represented this way are called **simple**. It is fairly obvious but worth noting that not all tensors are simple.

To see this note that the selection of r vectors and s covectors to form a simple tensor involves nt coefficients to represent these vectors and covectors in basis $\mathbf{a}$ when V has dimension n. On the other hand, a general tensor of order t is determined by the n^t numbers $T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})$. Solving n^t equations in nt unknowns is generally not possible for integers n and t exceeding 1, unless $n = t = 2$.

Getting back to the discussion from above, we associate T with the tensor $\widetilde{\widetilde{\mathbf{A}}}T \in \mathcal{T}^r_s(\mathbb{R}^n)$ defined by

$$\widetilde{\widetilde{\mathbf{A}}}T \;=\; T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\; e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}.$$

$\widetilde{\widetilde{\mathbf{A}}}$ is a combination of A and $\overline{A^*}$ applied, as appropriate, factor by factor.

We leave the definition of $\widetilde{\mathbf{A}}$ (see the diagram) to the imagination of the reader.

$$\begin{array}{ccc} & \mathcal{T}^r_s(V) & \\ \tilde{\mathbf{A}} & \uparrow\;\downarrow & \widetilde{\widetilde{\mathbf{A}}} \\ & \mathcal{T}^r_s(\mathbb{R}^n) & \end{array}$$

If you wish to do a calculation with a tensor T to produce a number you can use its representation $\widetilde{\widetilde{\mathbf{A}}}T$, with respect to a given ordered basis $\mathbf{a}$ in V, as a tensor on $\mathbb{R}^n$.

Specifically, we can evaluate T on a member of $(v(1), \ldots, v(r+s)) \in V_s^r$ by calculating the value of $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a}) e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$ on the member $(\overline{A^*}(v(1)), \ldots, \overline{A^*}(v(r)), A(v(r+1)), \ldots, A(v(r+s)))$ of $(\mathbb{R}^n)^r_s$.

We have arrived at several closely related definitions of, or uses for, the creatures like $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})$ seen peppering books which apply tensors.

Most simply, they are nothing more than the numbers $T(a^{i_1}, \ldots, a^{i_r}, a_{j_1}, \ldots, a_{j_s})$, the coefficients of the simple tensors $a_{i_1} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}$ in the explicit representation of the multilinear function, using the basis $\mathbf{a}$.

More abstractly, $T^{\cdots}_{\cdots}$ is a function from the set of all ordered bases of V to the set of real valued functions whose domain is the n^{r+s} possible values of our index set. The value of $T^{\cdots}_{\cdots}$ on an ordered basis $\mathbf{a}$ is the function $T^{\cdots}_{\cdots}(\mathbf{a})$.

The value of this function on an index combination is denoted $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})$, and these numbers are called the **coordinates of T in the ordered basis a**. For various index combinations but fixed ordered basis these numbers might be anything (though of course they depend on T and $\mathbf{a}$.) These values are unrelated to each other for different index combinations in the sense that they can be prescribed at will or measured and found to be any number if you are creating a tensor from scratch in a particular basis. But once they are all specified or known in one ordered basis their values are determined in all other bases, and we shall see soon how these numbers must change in a new ordered basis.

Finally, we can think of $T^{\cdots}_{\cdots}$ as defining a function from the set of all ordered bases of V to the set of tensors in $\mathcal{T}^r_s(\mathbb{R}^n)$.

For each ordered basis $\mathbf{a}$, T is assigned its own **representative tensor** $\widetilde{\widetilde{\mathbf{A}}}T \in \mathcal{T}^r_s(\mathbb{R}^n)$. Not every member of $\mathcal{T}^r_s(\mathbb{R}^n)$ can be a representative of this particular tensor using different bases. The real number coefficients must change in a coordinated way when moving from ordered basis to ordered basis, as we shall see below.

There are advantages and disadvantages to each point of view.

The list of coefficients using a single basis would suffice to describe a tensor completely, just as the constant of proportionality suffices to describe a direct variation of one real variable with another. But specifying the constant of proportionality and leaving it at that misses the *point* of direct variation, which is a functional relationship. Likewise, the coefficients which define a tensor are just numbers and thinking of them alone downplays the relationship they are intended to represent.

Guided by the impression that people think more easily about relationships taking place in some familiar environment you have the option of regarding tensors as living in $\mathbb{R}^n$ through a representative tensor there, one representative for each ordered basis of V. The utility of this viewpoint depends, of course, on familiarity with tensors in $\mathbb{R}^n$.

7. Tensor Coordinates and Change of Basis

Suppose T is the same tensor from above and $\mathbf{b}$ is a new ordered basis of V with dual ordered basis $\mathbf{b}^*$ of V^*.

Representing this tensor in terms of both the old ordered basis $\mathbf{a}$ and the new ordered basis $\mathbf{b}$ we have

$$\begin{aligned}
T \;&=\; T(b^{w_1},\ldots,b^{w_r},b_{h_1},\ldots,b_{h_s})\; b_{w_1}\otimes\cdots\otimes b_{w_r}\otimes b^{h_1}\otimes\cdots\otimes b^{h_s}\\
&=\; T(a^{i_1},\ldots,a^{i_r},a_{j_1},\ldots,a_{j_s})\; a_{i_1}\otimes\cdots\otimes a_{i_r}\otimes a^{j_1}\otimes\cdots\otimes a^{j_s}\\
&=\; T(a^{i_1},\ldots,a^{i_r},a_{j_1},\ldots,a_{j_s})\\
&\qquad\qquad\cdot\,(B^{w_1}(a_{i_1})b_{w_1})\otimes\cdots\otimes(B^{w_r}(a_{i_r})b_{w_r})\\
&\qquad\qquad\qquad\otimes\left(A^{j_1}(b_{h_1})b^{h_1}\right)\otimes\cdots\otimes\left(A^{j_s}(b_{h_s})b^{h_s}\right)\\
&=\; T(a^{i_1},\ldots,a^{i_r},a_{j_1},\ldots,a_{j_s})\\
&\qquad\qquad\cdot\; B^{w_1}(a_{i_1})\cdots B^{w_r}(a_{i_r})A^{j_1}(b_{h_1})\cdots A^{j_s}(b_{h_s})\\
&\qquad\qquad\qquad\cdot\; b_{w_1}\otimes\cdots\otimes b_{w_r}\otimes b^{h_1}\otimes\cdots\otimes b^{h_s}.
\end{aligned}$$

In a shorter form we have

$$\boxed{T^{w_1,\ldots,w_r}_{h_1,\ldots,h_s}(\mathbf{b}) = T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})\mathcal{B}^{w_1}_{i_1}\cdots\mathcal{B}^{w_r}_{i_r}\mathcal{A}^{j_1}_{h_1}\cdots\mathcal{A}^{j_s}_{h_s}.}$$

Behold the—rather messy—formula which must be satisfied by the coordinates of representations of a tensor in different bases.

We have already seen and used the coordinates of one mixed tensor, given by $\delta^i_j\, a_i\otimes a^j$. It is worth checking directly that the coordinates of this (rather unusual) tensor are the same in any ordered basis.

In deriving the formula for new tensor coordinates after change of basis we used a technique we will revisit.

In this case our assumption was simple: we assumed we were in possession of a tensor T, a multilinear real valued function on a product space formed from copies of V and V^*. No coordinates are mentioned, and play no role in the definition.

With that coordinate-free definition in hand, we *defined* coordinates in any basis and showed a relationship which must hold between coordinates in different bases.

You will most likely never see an *actual* tensor (as opposed to the "Let T be a tensor ...." type of tensor) defined in any way except through a coordinate representation in a basis. The existence of a tensor T underlying a multilinear function defined in a basis is part of the *theory of the subject you were studying when the putative tensor appeared.*

The change of basis formula is then a testable consequence for quantities you can measure of that subject-specific theory.

We will see this kind of thing in action—coordinate-free definition followed by implications for coordinates—later when we discuss topics such as contraction and raising and lowering indices.

This essential idea will be raised again in the next section.

At this point it seems appropriate to discuss the meaning of covariant and contravariant tensors and how their coordinates vary under change of basis, with the goal of getting some intuition about their differences. We illustrate this with tensors of rank 1: contravariant vectors (vectors) and and covariant vectors (covectors.)

The physical scenario we have in mind consists of two types of items in the air in front of your eyes with you seated, perhaps, at a desk.

First, we have actual physical displacements, say of various dust particles that you witness. Second, we have stacks of flat paper, numbered like pages in a book, each stack having uniform "air gap" between pages throughout its depth, though that gap might vary from stack to stack. We make no restriction about the angle any particular stack must have relative to the desk. We consider these pages as having indeterminant extent, perhaps very large pages, and the stacks to be as deep as required, though of uniform density.

The magnitude of a displacement will be indicated by the length of a line segment connecting start to finish, which we can give numerically should we decide on a standard of length. Direction of the displacement is indicated by the direction of the segment together with an "arrow head" at the finish point.

The magnitude of a stack will be indicated by the density of pages in the stack which we can denote numerically by reference to a "standard stack" if we decide on one. The direction of the stack is in the direction of increasing page number.

We now define a coordinate system on the space in front of you, measuring distances in centimeters, choosing an origin, axes and so on in some reasonable way with z axis pointing "up" from your desk.

Consider the displacement of a dust particle which moves straight up 100 centimeters from your desk, and a stack of pages laying on your desk with density 100 pages per centimeter "up" from your desk.

If you decide to measure distances in meters rather than centimeters, the vertical coordinate of displacement drops to 1, decreasing by a factor of 100. The numerical value of the density of the stack, however, increases to $10,000$.

When the "measuring stick" in the vertical direction increases in length by a factor of 100, coordinates of displacement drop by that factor and displacement is called contravariant because of this.

On the other hand, the stack density coordinate changes in the same way as the basis vector length, so we would describe the stack objects as covariant.

We haven't really shown that these stack descriptions and displacements can be regarded as vector spaces, though no doubt you have seen the geometrical procedure for scalar multiplication and vector addition of displacements. There are purely geometrical ways of combining stacks to produce a vector space structure on stacks too: two intersecting stacks create parallelogram "columns" and the sum stack has sheets that extend the diagonals of these columns.

But the important point is that if stacks and displacements are to be thought of as occurring in the same physical universe, and if a displacement is to be represented as a member u of $\mathbb{R}^3$, then a stack *must* be represented as a linear functional, a member σ of $\mathbb{R}^{3*}$.

There is a physical meaning associated with the number σu. It is the number of pages of the stack corresponding to σ crossing the shaft of the displacement corresponding to u, where this number is positive when the motion was in the direction of increasing "page number."

It is obvious on physical grounds that this number must be invariant: it *cannot* depend on the vagaries of the coordinate system used to calculate it. That is the meaning of

$$\sigma u = \sigma\left(\mathcal{A}\mathcal{B}\right)u = \left(\sigma\mathcal{A}\right)\left(\mathcal{B}u\right)$$

and why tensor coordinates of the two types change as they do.

8. Tensor Character

In applications, a real valued function is often defined by specified operations on the coordinates of several vector and/or covector variables in a convenient ordered basis **a**. It is then proposed that this function "is" a tensor.

This hypothesis has an underlying assumption, which must be remembered, and two parts, which must be verified.

The often unstated (or at least underemphasized and forgotten by beginners) assumption is that there actually is a vector space V which "is" reality, or closer to it anyway, whose coordinates in a basis generate a *representative* tensor in $\mathcal{T}^r_s(\mathbb{R}^n)$. You will likely be working with nothing *but* representative tensors, but the "real thing" is a member of $\mathcal{T}^r_s(V)$.

The first part of the hypothesis to be verified is that the coordinate process is linear in each vector or covector (each "slot") separately.

Secondly, it must be shown or simply assumed (possibly for physical reasons) that the process by which you produced the coordinates has **tensor character**: that is, had you applied the same process by which you calculated the tensor coordinates $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})$ in the *first* coordinate system **a** to a *second* coordinate system **b**, the tensor coordinates $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{b})$ thereby produced would be the properly transformed tensor coordinates from the first basis.

In order to provide examples and make the connection with classical treatments we first do some spadework. We specialize the example to consider a tensor of rank 3, selecting that as the "sweet spot" between too simple and too complicated.

The symbols x and y are intended to be the representatives in ordered basis $\mathbf{a}^*$ of generic covectors, while z is the representative in ordered basis **a** of a generic vector. In our examples below we assume the dimension of these underlying spaces to be at least three (so x, y and z have at least three coordinates each) though that is not really relevant to the issue.

A **coordinate polynomial** for x, y and z is any polynomial involving the coordinates of x, y and z and, after simplification perhaps, no other indeterminates.

If, after simplification, each nonzero term in this polynomial is of third degree and involves one coordinate from each of x, y and z then the polynomial is said to be **multilinear in x, y and z**.

So the polynomial $x_1\, y_2\, z^2 - 7x_2\, y_2\, z^3$ is multilinear in $x,\ y$ and z but $x_1\, x_3\, z^2 - 7x_2\, y_2\, z^1$ is not and $x_1\, y_2\, z^1 - 7x_2\, y_2$ is not. Though $8x_3\, y_2 - x_2\, y_1$ is not multilinear in $x,\ y$ and z, it *is* multilinear in x and y.

Multilinearity of a coordinate polynomial is necessary if it is to represent the effect of a member T of $\mathcal{T}^2_1(V)$ on generic vectors and covectors in its domain: it encapsulates multilinearity of T.

More generally, sometimes the coordinates of a proposed tensor T of order k are given by formulae involving an explicit ordered basis $\mathbf{a}$. Coordinates of the appropriate combination of k different generic vectors or covectors in this ordered basis $\mathbf{a}$ are named, as was x, y and z above, and the calculated tensor coordinates affixed as coefficient for each coordinate monomial, creating a coordinate polynomial.

Multilinearity means that the polynomial must (possibly after simplification and collection of like terms) be the sum of polynomial terms each of which is degree k and involves one coordinate from each of the k generic vectors or covectors. The (simplified) polynomial is called the **coordinate polynomial for T in ordered basis a**.

Going back to our example of order 3, if the polynomial $x_1\, y_2\, z^3 - 7x_2\, y_2\, z^1$ is the coordinate polynomial for a tensor $T \in \mathcal{T}^2_1$ in ordered basis $\mathbf{a}$ where V has dimension 3, the tensor coordinates are $T^{1,2}_3(\mathbf{a}) = 1,\ T^{2,2}_1(\mathbf{a}) = -7$ and in this ordered basis the other twenty-five tensor coefficients are 0.

After satisfying ourselves that a proposed coordinate polynomial is multilinear, which is often obvious, we still have some work to do. The numbers 1, -7 and 0, the coordinates in ordered basis $\mathbf{a}$, came from somewhere. Usually it is an explicit formula involving the basis $\mathbf{a}$. There is only one tensor T *could* be:

$$T = a_1 \otimes a_2 \otimes a^3 \ - \ 7\, a_2 \otimes a_2 \otimes a^1.$$

The question is, do the formulae that produced 1, -7 and 0 for basis $\mathbf{a}$ have tensor character: will they, if applied in a second basis $\mathbf{b}$ yield the necessary coefficients in that new basis?

Let's get specific. If basis $\mathbf{b}$ happens to be $b_1 = 2a_1$ and $b_2 = a_1 + a_2$ and $b_3 = a_3$ then

$$\mathcal{A} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \mathcal{B} = \begin{pmatrix} \frac{1}{2} & \frac{-1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So $T^{2,2}_1(\mathbf{b})$ must equal

$$\begin{aligned} T^{i_1,i_2}_{j_1}(\mathbf{a})\, \mathcal{B}^2_{i_1}\, \mathcal{B}^2_{i_2}\, \mathcal{A}^{j_1}_1 &= T^{1,2}_3(\mathbf{a})\, \mathcal{B}^2_1\, \mathcal{B}^2_2\, \mathcal{A}^3_1 \ + \ T^{2,2}_1(\mathbf{a})\, \mathcal{B}^2_2\, \mathcal{B}^2_2\, \mathcal{A}^1_1 \\ &= 1 \cdot 0 \cdot 1 \cdot 0 + (-7) \cdot 1 \cdot 1 \cdot 2 = -14. \end{aligned}$$

In the new basis the coordinate polynomial will have a term

$$T^{2,2}_1(\mathbf{b})\, \overline{x}_2\ \overline{y}_2\ \overline{z}^1 = -14\ \overline{x}_2\ \overline{y}_2\ \overline{z}^1$$

where $\overline{x},\ \overline{y}$ and $\overline{z}$ are the representatives in basis $\mathbf{b}$ of the *same* two covectors and the *same* vector, in the same order, that x, y and z represent in basis $\mathbf{a}$.

In the notation we have used up to this section, T will have a term $-14\, b_2 \otimes b_2 \otimes b^1$ when represented in basis $\mathbf{b}$.

There are 27 coordinates to be checked in a *generic* second basis if you need to demonstrate tensor character when the tensor coordinates (or coordinate polynomial coefficients) are supplied as formulae involving the basis.

To show that formulae producing coefficients in a basis of a member of $\mathcal{T}^r_s(V)$ have tensor character involves n^{r+s} potentially painful calculation involving coordinates in a *generic* second basis and the associated matrices of transition, the boxed equations in Section 7.

It is fortunate that, in applications, it is rare to find tensors for which either of n or $r+s$ exceed 4: that would still, at worst, require 256 calculations, though symmetries among the coordinates can cut this down considerably.

Taking an alternative point of view, if it is a *theory* that the procedure which produces coordinates in a basis *should have tensor character*, then the value of each coordinate in each basis is a prediction: it provides an *opportunity for experimental verification of the theory.*

Here's a simple example of a procedure that *fails* to produce coefficients with tensor character, using the same bases **a** and **b**.

Let us suppose that we define a multilinear (actually, this one is linear) function to be the sum of the first three coordinates of a covector in a basis. So using basis **a** the corresponding coordinate polynomial is $x_1 + x_2 + x_3$ which corresponds to tensor $a_1+a_2+a_3$. But using basis **b** the procedure produces coordinate polynomial $\overline{x}_1 + \overline{x}_2 + \overline{x}_3$ which corresponds to tensor $b_1 + b_2 + b_3$. These tensors are not the same. Therefore this process, though linear, does not define a tensor.

On a cautionary note, we mention a practice common in older tensor treatments in which tensors are defined using coordinate polynomials only, without explicit elaboration of the structure we have built from V and V^*.

You might see symbols $T^{i\,j}$ and $S^{k\,l}$ named as contravariant tensors. In our language, these are the coefficients of tensors T and S in a basis, and when covectors x, y, u and w are represented in terms of that basis

$$T(x,y) = T^{i\,j} x_i\, y_j \quad \text{and} \quad S(u,w) = S^{k\,l} u_k\, w_l.$$

Note that all the numbers on right side of both equations above are dependent on the explicit basis, but tensor character is invoked and we conclude that the sum of these products is invariant. In a new "barred basis"

$$T(x,y) = \overline{T}^{\,i\,j} \overline{x}_i\, \overline{y}_j \quad \text{and} \quad S(u,w) = \overline{S}^{\,k\,l} \overline{u}_k\, \overline{w}_l.$$

Usually, T and $T(x,y)$ and S and $S(x,y)$ are not mentioned explicitly, nor are the bases named. You are supposed to simply understand they are there.

Tensor operations are defined on these coordinate polynomials alone. Sometimes it is understood that the index i is associated with x, j with y, k with u and l with w. When an index symbol is associated with a variable symbol, statements can arise that seem odd to us.

For instance, the tensor products $T^{i\,j} S^{k\,l}$ and $S^{k\,l} T^{i\,j}$ refer to the coordinate polynomials

$$\left(T^{i\,j} x_i\, y_j\right) \left(S^{k\,l} u_k\, w_l\right) \quad \text{and} \quad \left(S^{k\,l} u_k\, w_l\right) \left(T^{i\,j} x_i\, y_j\right)$$

which are the same by the ordinary commutative and distributive properties of real numbers. Tensor product commutes using these conventions! Of course, what is happening behind the scenes is that you have decided on an order for your four covectors (perhaps alphabetical) and are simply sticking to it.

Many people find the silent mental gymnastics needed to keep the older notation straight to be unsettling, and tend to lose track in really complex settings. The gyrations we have chosen to replace them are, at least, explicit.

9. Relative Tensors

Certain types of processes that define a linear or multilinear function for each basis don't seem to be that useful or reflect much other than the features of the particular basis. For instance if $v = v^i(\mathbf{a})\, a_i \in V$ the functions

$$f_{\mathbf{a}},\, g_{\mathbf{a}} \colon V \to \mathbf{R} \quad \text{given by} \quad f_{\mathbf{a}}(v) = \sum_{i=1}^{n} v^i(\mathbf{a}) \quad \text{and} \quad g_{\mathbf{a}}(v) = v^1(\mathbf{a})$$

(same formula for any basis) are both linear but are *not* covectors.

They correspond to linear coordinate polynomials

$$x^1 + \cdots + x^n \qquad \text{and} \qquad x^1 \qquad \text{in any basis.}$$

They certainly might be useful for *something* (giving us useful information about each basis perhaps) but whatever that something might be it will probably *not*, by itself, describe "physical law."

But among those processes that define a linear or multilinear function for each basis, it is not only tensors which have physical importance or which have consistent, basis-independent properties. We will discuss some of these entities (relative tensors) after a few paragraphs of preliminary material.

The set of matrices of transition with matrix multiplication forms a group called the **general linear group** over $\mathbb{R}$. We denote this set with this operation by $\mathbf{GL}(\mathbf{n}, \mathbb{R})$.

A real valued function $f \colon GL(n, \mathbb{R}) \to \mathbb{R}$ is called **multiplicative** if $f(\mathcal{C}\mathcal{D}) = f(\mathcal{C})f(\mathcal{D})$ for every pair of members $\mathcal{C}$ and $\mathcal{D}$ of $GL(n, \mathbb{R})$.

Any nontrivial (i.e. nonzero) multiplicative function must take the identity matrix to 1, and also a matrix and its inverse matrix must be sent to reciprocal real numbers. Any such function must satisfy $f\left(\mathcal{A}^{-1}\,\mathcal{B}\,\mathcal{A}\right) = f\left(\mathcal{B}\right)$, for instance.

Given two nontrivial multiplicative functions, both their product and their ratio are multiplicative. If a nontrivial multiplicative function is always positive, any real power of that function is multiplicative.

In this section we will be interested in several kinds of multiplicative functions.

First, let **det** denote the **determinant** function on square matrices. Its properties are explored in any linear algebra text. We find in [8] p.154 that $\det(MN) = \det(M)\det(N)$ for compatible square matrices M and N. Since $\det(\mathcal{I}) = 1$ this yields $\det\left(M^{-1}\right) = (\det(M))^{-1}$ when M is invertible. We then find that

$$\det\left(M^{-1} N M\right) = \det(N).$$

Determinant is an invariant under similarity transformations.

The determinant is also multiplicative on $GL(n, \mathbb{R})$.

So is the absolute value of the determinant function, $|\det|$, and also the ratio **sign** $= \frac{|\det|}{\det}$. This last function is 1 when applied to matrices with positive determinant and -1 otherwise.

If w is any real number, the functions f and g defined by

$$f(\mathcal{C}) = \operatorname{sign}(\mathcal{C})\,|\det(\mathcal{C})|^{w} \quad \text{and} \quad g(\mathcal{C}) = |\det(\mathcal{C})|^{w}$$

are multiplicative and, in fact, these functions exhaust the possibilities among continuous multiplicative functions[3] on $GL(n, \mathbb{R})$.

$f(\mathcal{C})$ is negative exactly when $\det(\mathcal{C})$ is negative, and is called an **odd weighting function with weight w** with reference to this fact. g is always positive, and is called an **even weighting function with weight w**.

A **relative tensor** is defined using ordinary tensors but with a different transformation law when going to a new coordinate system. These transformation laws take into account how volumes could be measured in V, or the "handedness" of a coordinate basis there, and relative tensors are vital when theories of physical processes refer to these properties in combination with multilinearity.

Specifically, one measures or calculates or otherwise determines numbers called coordinates of a proposed relative tensor T in one basis $\mathbf{a}$, the numbers $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})$, and in this basis the tensor density T is the multilinear function

$$T = T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})\ a_{i_1} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}.$$

The difference is that when switching to a new coordinate system $\mathbf{b}$ the coordinates of a relative tensor change according to one of the two formulae

$$\begin{aligned} T^{k_1,\ldots,k_r}_{h_1,\ldots,h_s}(\mathbf{b}) &= \operatorname{sign}(\mathcal{B})\,|\det(\mathcal{B})|^{w}\ T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})\mathcal{B}^{k_1}_{i_1}\cdots\mathcal{B}^{k_r}_{i_r}\mathcal{A}^{j_1}_{h_1}\cdots\mathcal{A}^{j_s}_{h_s} \\ &\text{or} \\ T^{k_1,\ldots,k_r}_{h_1,\ldots,h_s}(\mathbf{b}) &= |\det(\mathcal{B})|^{w}\ T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})\mathcal{B}^{k_1}_{i_1}\cdots\mathcal{B}^{k_r}_{i_r}\mathcal{A}^{j_1}_{h_1}\cdots\mathcal{A}^{j_s}_{h_s}. \end{aligned}$$

Relative tensors of the first kind are called **odd relative tensors of weight w** while tensors of the second kind are called **even relative tensors of weight w**.

If $\mathcal{C}$ is the matrix of transition from ordered basis $\mathbf{b}$ to ordered basis $\mathbf{c}$ then $\mathcal{C}\,\mathcal{B}$ is the matrix of transition from ordered basis $\mathbf{a}$ directly to ordered basis $\mathbf{c}$.

Because of this, it is not hard to show that the coordinates $T^{k_1,\ldots,k_r}_{h_1,\ldots,h_s}(\mathbf{c})$ for a relative tensor in ordered basis $\mathbf{c}$ can be derived from the coordinates in ordered basis $\mathbf{a}$ directly, or through the intermediary coordinates in ordered basis $\mathbf{b}$ using two steps: that is, the multiplicative nature of the weight function is exactly what we need for the coordinate change requirement to be consistent.

An even relative tensor of weight 0 is just a tensor, so the concept of relative tensor generalizes that of tensor. If it needs to be emphasized, ordinary tensors are called **absolute tensors** to distinguish them from relative tensors with more interesting weighting functions.

[3]See [2] page 349.

This subject has a long history and many applications, so a lot of vocabulary is to be expected.

Odd relative tensors are described variously (and synonymously) as

- **axial tensors**
- **pseudotensors**
- **twisted tensors**
- **oriented tensors.**

The "axial" adjective seems more common when the relative tensor is of rank 0 or 1: **axial scalars** or **axial vectors or covectors,** but sometimes you will see these relative tensors referred to as **pseudoscalars** or **pseudovectors or pseudocovectors.**

An even relative tensor (for example an ordinary tensor) is sometimes called a **polar** tensor.

If the weight w is positive (usually 1) the phrase **tensor density** is likely to be used for the relative tensor, while if the weight is negative (often -1) we take a tip from [19] and call these **tensor capacities,** though many sources refer to any relative tensor of nonzero weight as a tensor density.

These adjectives can be combined: e.g., axial vector capacity or pseudoscalar density.

The columns of the change of basis matrix $\mathcal{B}$ represents the old basis vectors $\mathbf{a}$ in terms of the new basis vectors $\mathbf{b}$. The number $|\det(\mathcal{B})|$ is often interpreted as measuring the "relative volume" of the parallelepiped formed with edges along basis $\mathbf{a}$ as compared to the parallelepiped formed using $\mathbf{b}$ as edges. So, for example, if the new basis vectors $\mathbf{b}$ are all just twice as long as basis vectors $\mathbf{a}$, then $|\det(\mathcal{B})| = 2^{-n}$.

With that situation in mind, imagine V to be filled with uniformly distributed dust. Since the unit of volume carved out by basis $\mathbf{b}$ is so much larger, the measured density of dust using basis $\mathbf{b}$ would be greater by a factor of 2^n than that measured using basis $\mathbf{a}$. On the other hand, the numerical measure assigned to the "volume" or "capacity" of a given geometrical shape in V would be smaller by just this factor when measured using basis $\mathbf{b}$ as compared to the number obtained using basis $\mathbf{a}$ as a standard volume.

The vocabulary "tensor density" for positive weights and "tensor capacity" for negative weights acknowledges and names these (quite different) behaviors.

One can form the product of relative tensors (in a specified order) to produce another relative tensor. The weight of the product is the sum of the weights of the factors. The product relative tensor will be odd exactly when an odd number of the factor relative tensors are odd.

Many of the constructions involving tensors apply also to relative tensors with little or no change. These new objects will be formed by applying the construction to the "pure tensor" part of the relative tensor and applying the ordinary distributive property (or other arithmetic) to any weighting function factors, as we did in the last paragraph.

10. Push Forward and Pull Back

Up to this point we have used only two vector spaces: generic V and particular $\mathbb{R}^n$. We have a way of associating tensors on V with tensors on $\mathbb{R}^n$.

It is useful in applications to generalize this procedure.

Suppose $H\colon V \to W$ is a linear function between two finite dimensional vector spaces.

Recall our definition of product space

$$V^r_s = \underbrace{V^* \times \cdots \times V^*}_{r \text{ copies}} \times \underbrace{V \times \cdots \times V}_{s \text{ copies}}$$

A member of $\mathfrak{T}^r_s(V)$ is a multilinear function $T\colon V^r_s \to \mathbb{R}$ and any such can be represented in basis $\mathbf{a}$ of V as

$$T \;=\; T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\; a_{i_1} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}.$$

Define

$$\boldsymbol{H^*(a_{i_1} \otimes \cdots \otimes a_{i_r}) = H(a_{i_1}) \otimes \cdots \otimes H(a_{i_r}) \in \mathfrak{T}^r_0(W)}$$

and extend H^* to a map $\boldsymbol{H^*\colon \mathfrak{T}^r_0(V) \to \mathfrak{T}^r_0(W)}$ by linearity.

The linearity of H implies that this vector space homomorphism does not depend on the basis $\mathbf{a}$ of V used to define it.

If H is one-to-one so is H^*. If H is onto W then H^* is onto $\mathfrak{T}^r_0(W)$. These facts can be most easily seen by considering the case $r = 1$.

Now suppose $\mathbf{c}$ is a basis of vector space W and define

$$\boldsymbol{H_*(c^{i_1} \otimes \cdots \otimes c^{i_r}) = c^{i_1} \circ H \otimes \cdots \otimes c^{i_r} \circ H \in \mathfrak{T}^0_s(V)}$$

and extend H_* to a map $\boldsymbol{H_*\colon \mathfrak{T}^0_s(W) \to \mathfrak{T}^0_s(V)}$ by linearity.

This vector space homomorphism does not depend on the basis $\mathbf{c}^*$ of W^* used to define it.

If H is one-to-one then H^* is onto $\mathfrak{T}^0_s(V)$. If H is onto W then H^* is one-to-one.

In case H is an isomorphism both H^* and H_* are isomorphisms too. And in this case we can define isomorphisms

$$\boldsymbol{H^*\colon \mathfrak{T}^r_s(V) \to \mathfrak{T}^r_s(W) \quad \text{and} \quad H_*\colon \mathfrak{T}^r_s(W) \to \mathfrak{T}^r_s(V)}$$

by applying either H or H^{-1} at each factor, as appropriate.

$\boldsymbol{H^*}$ is called the **pushforward map of H** and $\boldsymbol{H_*}$ is called the **pullback map of H** and one uses the transitive verb form “to pull back” a tensor or “to push forward” a tensor using H.

11. An Identification of $\mathcal{H}om(V,V)$ and $\mathcal{T}^1_1(V)$

There is an identification of the vector space $\boldsymbol{\mathcal{H}om(V,V)}$ of all linear functions $F\colon V \to V$ and the tensor space $\mathcal{T}^1_1$ which we describe now.

Whenever you see a function with vector output called a tensor, and they are common, it is this association to which they refer.

Define $\Psi(F)\colon V^* \times V \to \mathbb{R}$ to be the function $\Psi(F)(\omega, v) = \omega(F(v))$.

$\Psi(F)$ is obviously multilinear and $\Psi(F + c\,G) = \Psi(F) + c\,\Psi(G)$ for any other linear map $G\colon V \to V$ and any scalar c. So we have created a member of $\mathcal{T}^1_1$ and a homomorphism from $\mathcal{H}om(V,V)$ into $\mathcal{T}^1_1$, and each space has dimension n^2.

If $\Psi(F)$ is the zero tensor then $\omega(F(v)) = 0$ for every ω and v so F is the zero transformation. That means the kernel of Ψ is trivial and we have created an isomorphism $\Psi\colon \mathcal{H}om(V,V) \to \mathcal{T}^1_1$. You will note that no reference was made to a basis in the definition of Ψ.

Consider the map $\Theta\colon \mathcal{T}^1_1 \to \mathcal{H}om(V,V)$ given, for $T = T^i_j(\mathbf{a})\, a_i \otimes a^j \in \mathcal{T}^1_1$, by

$$\Theta(T)(v) = T^i_j(\mathbf{a})\, v^j\, a_i = T^i_j(\mathbf{a})\, a^j(v)\, a_i \qquad \text{whenever } v = v^j a_j.$$

If $v = v^j a_j$ and $\omega = \omega_j a^j$

$$\begin{aligned}(\Psi \circ \Theta)(T)(\omega, v) &= \omega(\,\Theta(T)(v)\,) = \omega_k a^k T^i_j(\mathbf{a})\, a^j(v)\, a_i = \omega_i T^i_j(\mathbf{a})\, v^j \\ &= T^i_j(\mathbf{a})\, a_i \otimes a^j\,(\omega, v) = T(\omega, v).\end{aligned}$$

It follows that $\Theta = \Psi^{-1}$.

It may not be obvious that the formula for Θ is independent of basis, but it must be since the definition of Ψ is independent of basis, and the formula for Θ produces Ψ^{-1} in any basis.

12. Contraction and Trace

At this point we are going to define a process called **contraction** which allows us to reduce the order of a relative tensor by 2: by one covariant and one contravariant order value. We will initially define this new tensor by a process which uses a basis.

Specifically, we consider $T \in \mathcal{T}^r_s(V)$ where both r and s are at least 1. Select one contravariant index position α and one covariant index position β. We will define the tensor $\mathbf{C}^\alpha_\beta \mathbf{T}$, called the **contraction of T with respect to the αth contravariant and βth covariant index.**

Define in a particular ordered basis $\mathbf{a}$ the numbers

$$\boxed{H^{i_1,\ldots,i_{r-1}}_{j_1,\ldots,j_{s-1}}(\mathbf{a}) = T^{i_1,\ldots,i_{\alpha-1},k,i_\alpha,\ldots,i_{r-1}}_{j_1,\ldots,j_{\beta-1},k,j_\beta,\ldots,j_{s-1}}(\mathbf{a}).}$$

This serves to define real numbers for all the required index values in $\mathcal{T}^{r-1}_{s-1}(V)$ and so these real numbers $H^{i_1,\ldots,i_{r-1}}_{j_1,\ldots,j_{s-1}}(\mathbf{a})$ do define *some* tensor in $\mathcal{T}^{r-1}_{s-1}(V)$.

Any process, including this one, that creates a multilinear functional in one ordered basis defines a tensor: namely, the tensor obtained in other bases by transforming the coordinates according to the proper pattern for that order of tensor.

It is worth belaboring the point raised in Section 8: that is not what people usually mean when they say that a process creates or "is" a tensor. **What they usually mean is that the procedure outlined, though it might use a basis inter alia, would produce the same tensor whichever basis had been initially picked.** We illustrate this idea by demonstrating the tensor character of contraction.

Contracting T in the αth contravariant against the βth covariant coordinate in the ordered basis $\mathbf{b}$ gives

$$\begin{aligned} H^{w_1,\dots,w_{r-1}}_{h_1,\dots,h_{s-1}}(\mathbf{b}) &= T^{w_1,\dots,w_{\alpha-1},k,w_\alpha,\dots,w_{r-1}}_{h_1,\dots,h_{\beta-1},k,h_\beta,\dots,h_{s-1}}(\mathbf{b}) \\ &= T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\mathcal{B}^{w_1}_{i_1}\cdots\mathcal{B}^{w_{\alpha-1}}_{i_{\alpha-1}}\mathcal{B}^{k}_{i_\alpha}\mathcal{B}^{w_\alpha}_{i_{\alpha+1}}\cdots\mathcal{B}^{w_{r-1}}_{i_r}\mathcal{A}^{j_1}_{h_1}\cdots\mathcal{A}^{j_{\beta-1}}_{h_{\beta-1}}\mathcal{A}^{j_\beta}_{k}\mathcal{A}^{j_{\beta+1}}_{h_\beta}\cdots\mathcal{A}^{j_s}_{h_{s-1}}. \end{aligned}$$

We now note that $\mathcal{B}^k_{i_\alpha}\mathcal{A}^{j_\beta}_k = \delta^{j_\beta}_{i_\alpha}$. So we make four changes to the last line:

- Factor out the sum $\mathcal{B}^k_{i_\alpha}\mathcal{A}^{j_\beta}_k$ and replace it with $\delta^{j_\beta}_{i_\alpha}$.
- Set $q = j_\beta = i_\alpha$ to eliminate $\delta^{j_\beta}_{i_\alpha}$ from the formula.
- Define $\bar{i}_m = \begin{cases} i_m, & \text{if } m = 1,\dots \alpha-1; \\ i_{m+1}, & \text{if } m = \alpha,\dots,r-1. \end{cases}$
- Define $\bar{j}_m = \begin{cases} j_m, & \text{if } m = 1,\dots,\beta-1; \\ j_{m+1}, & \text{if } m = \beta,\dots,s-1. \end{cases}$

The last line in the equation from above then becomes

$$\begin{aligned} &T^{\bar{i}_1,\dots,\bar{i}_{\alpha-1},q,\bar{i}_\alpha,\dots,\bar{i}_{r-1}}_{\bar{j}_1,\dots,\bar{j}_{\beta-1},q,\bar{j}_\beta,\dots,\bar{j}_{s-1}}(\mathbf{a})\mathcal{B}^{w_1}_{\bar{i}_1}\cdots\mathcal{B}^{w_{\alpha-1}}_{\bar{i}_{\alpha-1}}\mathcal{B}^{w_\alpha}_{\bar{i}_\alpha}\cdots\mathcal{B}^{w_{r-1}}_{\bar{i}_{r-1}}\mathcal{A}^{\bar{j}_1}_{h_1}\cdots\mathcal{A}^{\bar{j}_{\beta-1}}_{h_{\beta-1}}\mathcal{A}^{\bar{j}_\beta}_{h_\beta}\cdots\mathcal{A}^{\bar{j}_{s-1}}_{h_{s-1}} \\ &\qquad = H^{\bar{i}_1,\dots,\bar{i}_{r-1}}_{\bar{j}_1,\dots,\bar{j}_{s-1}}(\mathbf{a})\mathcal{B}^{w_1}_{\bar{i}_1}\cdots\mathcal{B}^{w_{r-1}}_{\bar{i}_{r-1}}\mathcal{A}^{\bar{j}_1}_{h_1}\cdots\mathcal{A}^{\bar{j}_{s-1}}_{h_{s-1}} \end{aligned}$$

This is exactly how the coordinates of H must change in going from ordered basis $\mathbf{a}$ to $\mathbf{b}$ if the contraction process is to generate the same tensor when applied to the coordinates of T in these two bases. Contraction can be applied to the representation of a tensor in *any* ordered basis and will yield the same tensor.

So we define the tensor $\mathbf{C}^\alpha_\beta\mathbf{T}$ by

$$\boxed{(C^\alpha_\beta T)^{\cdots}_{\cdots} = H^{\cdots}_{\cdots}.}$$

Whenever $T \in \mathcal{T}^1_1$ (for instance, a tensor such as $\Psi(F)$ from Section 11 above) we have an interesting case. Then the number $(C^1_1T)(\mathbf{a}) = T^k_k(\mathbf{a}) = T(a^k, a_k)$ is called the **trace of T** (or the trace of $F\colon V \to V$ when $T = \Psi(F)$ as formed in the previous section.)

We have just defined a number $\mathbf{trace(T)} = C^1_1T \in \mathcal{T}^0_0 = \mathbb{R}$ for each member T of $\mathcal{T}^\mathbf{1}_\mathbf{1}(\mathbf{V})$. No basis is mentioned here: the trace does not depend on the basis. It is an invariant.

This usage of the word "trace" is related to the trace of a matrix as follows.

Suppose $\theta \in V^*$ and $v \in V$. In the ordered basis $\mathbf{a}$ the tensor T is $T^i_j(\mathbf{a})\ a_i \otimes a^j$ and θ and v have representations as $\theta = \theta_i a^i$ and $v = v^i a_i$. So if $\big(T^i_j(\mathbf{a})\big)$ is the matrix with ijth entry $T^i_j(\mathbf{a})$

$$\begin{aligned}
T(\theta, v) =& T^i_j(\mathbf{a})\ a_i \otimes a^j(\theta, v)\\
=& T^i_j(\mathbf{a})\ \theta_i v^j\\
=& (\theta_1, \dots, \theta_n) \left(T^i_j(\mathbf{a})\right) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}\\
=& \overline{A^*}(\theta) \left(T^i_j(\mathbf{a})\right) A(v).
\end{aligned}$$

In other words, down in $\mathbb{R}^n$ the tensor $T^i_j(\mathbf{a})\, e_i \otimes e^j$ is, essentially, "matrix multiply in the middle" with the representative row matrix $\overline{A^*}(\theta)$ of θ on the left and the column $A(v)$ on the right. $trace(T)$ is the ordinary trace of this "middle matrix," the sum of its diagonal entries.

Two matrices P and Q are said to be **similar** if there is invertible matrix M with $P = M^{-1}QM$, and the process of converting Q into P in this way is called a **similarity transformation**.

If we change basis to $\mathbf{b}$ the "matrix in the middle" from above becomes

$$\left(T^i_j(\mathbf{b})\right) = \left(\mathcal{B}^w_i\ T^i_j(\mathbf{a})\ \mathcal{A}^j_h\right) = \mathcal{B}\left(T^i_j(\mathbf{a})\right)\mathcal{A} = \mathcal{A}^{-1}\left(T^i_j(\mathbf{a})\right)\mathcal{A}.$$

Any invertible matrix $\mathcal{A}$ (this requires a **tiny proof**) can be used to generate a new ordered basis $\mathbf{b}$ of V, and we recover by direct calculation the invariance of trace under similarity transformations.

It is interesting to see how this relates to the linear transformation $F\colon V \to V$ and associated $\Psi(F)$ and the matrix of F under change of basis, and we leave that comparison to the reader.

If $T \in \mathcal{T}^r_s(V)$ is any tensor with $0 < \alpha \le r$ and $0 < \beta \le s$, the contraction $C^\alpha_\beta T$ is related to the trace of a certain matrix, though it is a little harder to think about with all the indices floating around. First, we fix all but the αth contravariant and βth covariant index, thinking of all the rest as constant for now. We define the matrix $M(\mathbf{a})$ to be the matrix with ijth entry

$$M^i_j(\mathbf{a}) = T^{i_1,\dots,i_{\alpha-1},i,i_{\alpha+1},\dots,i_r}_{j_1,\dots,j_{\beta-1},j,j_{\beta+1},\dots,j_s}(\mathbf{a}).$$

So for each index combination with entries fixed except in these two spots, we have a member of $\mathcal{T}^1_1(V)$ and $M(\mathbf{a})$ is the matrix of its representative, using ordered basis $\mathbf{a}$, in $\mathcal{T}^1_1(\mathbb{R}^n)$. **The contraction is the trace of this matrix: the trace of a different matrix for each index combination.**

Contraction of relative tensors proceeds in an identical fashion. Contraction does not change the weight of a relative tensor, and the properties "even" or "odd" persist, unchanged by contraction.

13. Evaluation

Suppose $T \in \mathcal{T}^{\mathbf{r}}_{\mathbf{s}}(\mathbf{V})$. By definition, T is a multilinear function defined on $V^r_s = V^* \times \cdots \times V^* \times V \times \cdots \times V$, where there are r copies of V^* followed by s copies of V in the domain.

If $\theta \in V^*$ and $1 \leq k \leq r$ we can evaluate T at θ in its kth slot yielding the function

$$E(\theta, k)(T) = T(\cdots, \theta, \cdots)$$

where the dots indicate remaining "unfilled" slots of T, if any.

It is obvious that $E(\theta, k)(T)$ is also a tensor, and a member of $\mathcal{T}^{\mathbf{r-1}}_{\mathbf{s}}(\mathbf{V})$.

Suppose $\mathbf{a}$ is an ordered basis of V and $\theta = \theta_i(\mathbf{a})\, a^i$. Then

$$T = T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\ a_{i_1} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}$$

so

$$\begin{aligned}
&E(\theta, k)(T)(\mathbf{a})\\
&\quad = T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\ a_{i_1} \otimes \cdots \otimes a_{i_{k-1}} \otimes a_{i_k}(\theta) \otimes a_{i_{k+1}} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}\\
&\quad = T^{i_1,\dots,i_r}_{j_1,\dots,j_s}(\mathbf{a})\ \theta_{i_k}(\mathbf{a})\, a_{i_1} \otimes \cdots \otimes a_{i_{k-1}} \otimes a_{i_{k+1}} \otimes \cdots \otimes a_{i_r} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}\\
&\quad = G^{i_1,\dots,i_{r-1}}_{j_1,\dots,j_s}(\mathbf{a})\ a_{i_1} \otimes \cdots \otimes a_{i_{r-1}} \otimes a^{j_1} \otimes \cdots \otimes a^{j_s}
\end{aligned}$$

where

$$G^{i_1,\dots,i_{r-1}}_{j_1,\dots,j_s}(\mathbf{a}) = T^{i_1,\dots,i_{k-1},w,i_k,\dots,i_{r-1}}_{j_1,\dots,j_s}(\mathbf{a})\ \theta_w(\mathbf{a}).$$

$G = E(\theta, k)(T)$ is the tensor you would obtain by taking the tensor product of T with θ and then contracting in the appropriate indices:

$$G = E(\theta, k)(T) = C^k_{s+1}(T \otimes \theta).$$

By a calculation identical to the above, evaluation at a member $v \in V$ in slot $r + k$, one of the last s slots of T, is a tensor in $\mathcal{T}^{\mathbf{r}}_{\mathbf{s-1}}(\mathbf{V})$, and that tensor is

$$H = E(v, k)(T) = C^{r+1}_k(T \otimes v).$$

In ordered basis $\mathbf{a}$, if $v = v^i(\mathbf{a})\, a_i$ then H has coordinates

$$H^{i_1,\dots,i_r}_{j_1,\dots,j_{s-1}}(\mathbf{a}) = T^{i_1,\dots,i_r}_{j_1,\dots,j_{k-1},w,j_k,\dots,j_{s-1}}(\mathbf{a})\ v^w(\mathbf{a})\,.$$

By combining the two operations, we see that evaluating T in any number of its slots at specific choices of members of V or V^* (as appropriate for the slots involved) will produce a tensor. If k slots are to be filled, that tensor can be formed by the k-fold tensor product of T with the specified members of V or V^*, followed by k contractions.

We make an obvious remark: if we evaluate a tensor at *all* of its slots, we get a tensor of order 0, an **invariant** or (synonymous) **scalar**. If we evaluate the representation of that same tensor at the representations of those same vectors **in any basis** we get the same numerical output.

This remark has a converse.

Recall that the coordinates in a basis of a member of V^r_s is a member of $(\mathbb{R}^n)^r_s$. Suppose you determine a multilinear function for each basis of V which acts on

$(\mathbb{R}^n)^r_s$. Suppose further that, for each $v \in V^r_s$, if you evaluate the multilinear function corresponding to a basis at the coordinates corresponding to v in that basis you always get the same (that is, basis independent) numerical output. Then the process by which you created the multilinear function in each basis has tensor character: you have defined a tensor in $\mathcal{T}^{\mathbf{r}}_{\mathbf{s}}(\mathbf{V})$.

Some sources find this observation important enough to name, usually calling it (or something equivalent to it) **The Quotient Theorem.**[4]

Combining our previous remarks about products and contraction of relative tensors, we see that evaluation of a relative tensor on vectors or covectors can be performed at any of its slots, and the process does not change the weight of a relative tensor. The properties "even" or "odd" persist unchanged by evaluation.

Evaluation of a relative tensor on *relative* vectors or *relative* covectors *could* change both weight and whether the resulting relative tensor is even or odd, and we leave it to the reader to deduce the outcome in this case.

14. Symmetry and Antisymmetry

It is obvious that, for a tensor of order 2, $T(x, y)$ need not bear any special relation to $T(y, x)$. For one thing, $T(x, y)$ and $T(y, x)$ won't both be defined unless the domain is either V^2_0 or V^0_2. But even then they need not be related in any specific way.

A covariant or contravariant relative tensor T of order L is called **symmetric** when switching exactly two of its arguments never changes its value. Specifically, for each pair of distinct positive integers j and k not exceeding L and all $v \in V^0_L$ or V^L_0, whichever is the domain of T,

$$\begin{aligned} &T(v(1), \ldots, v(L)) \\ &\quad = T(v(1), \ldots, v(j-1), v(k), v(j+1), \ldots, v(k-1), v(j), v(k+1), \ldots, v(L)). \end{aligned}$$

A covariant or contravariant relative tensor is called (any or all of) **antisymmetric**, **alternating** or **skew symmetric** if switching any two of its arguments introduces a minus sign but otherwise does not change its value. Specifically:

$$\begin{aligned} &T(v(1), \ldots, v(L)) \\ &\quad = -\ T(v(1), \ldots, v(j-1), v(k), v(j+1), \ldots, v(k-1), v(j), v(k+1), \ldots, v(L)). \end{aligned}$$

The following facts about permutations are discussed in many sources, including [9] p. 46. A **permutation of the set $\{\mathbf{1}, \ldots, \mathbf{L}\}$**, the set consisting of the first L positive integers, is a one-to-one and onto function from this set to itself. Let $\mathcal{P}_{\mathbf{L}}$ denote the set of all permutations on this set. The composition of any two permutations is also a permutation. There are $L!$ distinct permutations of a set containing L elements. Any permutation can be built by switching two elements

[4]Our statement has a generalization. Suppose that we identify a fixed collection of slots of the multilinear function defined as above using a basis. Suppose that whenever you fill those particular slots with coordinates of vectors/covectors in a basis the resulting contracted multilinear function is independent of the basis: had you used coordinates in a different basis the new coefficients would be the same as the old ones, properly transformed so as to have tensor character. Then the means by which you formed the original coefficients has tensor character.

of the set at a time gradually building up the final permutation by composition of a sequence of such switches. A "switch," involving as it does exactly two distinct elements, is also called a **2-cycle**. You might be able to build P by this process in many ways. If you can build a permutation as the composition of an even number of 2-cycles it is called an **even permutation** while if you can build a permutation as the composition of an odd number of 2-cycles it is called an **odd permutation**. It is a fact that a permutation is even or odd but cannot be both. There are exactly $L!/2$ even permutations and $L!/2$ odd ones. If P is even, define $\mathbf{sgn(P) = 1}$ and if P is odd define $\mathbf{sgn(P) = -1}$. For any permutations P and Q we have $sgn(P \circ Q) = sgn(P)\, sgn(Q)$.

If P is a permutation of $\{1, \dots, L\}$ and T is a covariant or contravariant relative tensor of order L let T_P denote the tensor whose values are given by

$$T_P(v(1), \dots, v(L)) = T(v(P(1)), \dots, v(P(L))).$$

In other words, T_P is the relative tensor obtained from T by reordering the arguments using the permutation P. The function that sends T to T_P is sometimes called a **braiding map.** Braiding maps are isomorphisms.

Using this notation, T is symmetric precisely if $T_P = T$ whenever P is a 2-cycle. T is alternating precisely when $T_P = -T$ whenever P is a 2-cycle.

Since every permutation is a composition of 2-cycles, this means that T is symmetric precisely when $T_P = T$ for all permutations P, while T is alternating precisely if $T_P = sgn(P)T$ for any permutation P.

A covariant or contravariant relative tensor T of order L can undergo a process called **symmetrization**. It is defined in a coordinate-independent way by

$$\mathbf{Sym(T)} = \frac{1}{L!} \sum_{P \in \mathcal{P}_L} T_P$$

where the sum is over all $L!$ members of the set of permutations $\mathcal{P}_L$ of $\{1, \dots, L\}$.

Antisymmetrization is accomplished by the coordinate-independent formula

$$\mathbf{Alt(T)} = \frac{1}{L!} \sum_{P \in \mathcal{P}_L} sgn(P)\, T_P.$$

It is sometimes said that $Sym(T)$ is the **symmetric part** of T, while $Alt(T)$ is said to be the **skew part** of T.

For instance if a^1, a^2 are 1-forms then

$$Sym(a^1 \otimes a^2) = \frac{1}{2}\,(\, a^1 \otimes a^2 + a^2 \otimes a^1 \,) \text{ and } Alt(a^1 \otimes a^2) = \frac{1}{2}\,(\, a^1 \otimes a^2 - a^2 \otimes a^1 \,)$$

and if a^3 is a third 1-form then

$$\begin{aligned} Alt(a^1 \otimes a^2 \otimes a^3) = &\frac{1}{6}\,(\, a^1 \otimes a^2 \otimes a^3 - a^2 \otimes a^1 \otimes a^3 + a^2 \otimes a^3 \otimes a^1 \\ &- a^3 \otimes a^2 \otimes a^1 + a^3 \otimes a^1 \otimes a^2 - a^1 \otimes a^3 \otimes a^2 \,). \end{aligned}$$

Both of these operations are linear when viewed as functions on the space of tensors of a given order. Applied to a representation of T as $T = T_{i_1,\dots,i_L}(\mathbf{a})\, a^{i_1} \otimes$

$\cdots \otimes a^{i_L}$ in a basis, this gives $Sym(T) = T_{i_1,\ldots,i_L}(\mathbf{a})\, Sym(a^{i_1} \otimes \cdots \otimes a^{i_L})$ and $Alt(T) = T_{i_1,\ldots,i_L}(\mathbf{a})\, Alt(a^{i_1} \otimes \cdots \otimes a^{i_L})$.

In some sources, the Alt operation is indicated on coordinates of a tensor by using **square brackets** on the indices, such as $T_{[i_1,\ldots,i_L]}$. So

$$T_{[i_1,\ldots,i_L]}(\mathbf{a}) = Alt(T)_{i_1,\ldots,i_L}(\mathbf{a}).$$

In these sources **round brackets** may be used to indicate symmetrization:

$$T_{(i_1,\ldots,i_L)}(\mathbf{a}) = Sym(T)_{i_1,\ldots,i_L}(\mathbf{a}).$$

One also sees these notations spread across tensor products as in the expression $S_{[i_1,\ldots,i_m}T_{j_1,\ldots,j_L]}$, used to indicate the coordinates of $Alt(S \otimes T)$.

It is not necessary to involve all indices in these operations. Sometimes you may see brackets enclosing only some of the indices, as

$$T_{[i_1,\ldots,i_k],i_{k+1},\ldots,i_L}(\mathbf{a}) \quad \text{or} \quad T_{(i_1,\ldots,i_k),i_{k+1},\ldots,i_L}(\mathbf{a}).$$

The intent here is that these tensors are calculated on a selection of L vectors by permuting the vectors in the first k "slots" of T in all possible ways, leaving the last $L-k$ arguments fixed. After summing the results of this (with factor "-1" where appropriate) divide by $k!$.

In the calculation to follow, we let $Q = P^{-1}$ for a generic permutation P, and note that $sgn(P) = sgn(Q)$. For a particular basis tensor $a^{i_1} \otimes \cdots \otimes a^{i_L}$,

$$\begin{aligned}
&Alt(a^{i_1} \otimes \cdots \otimes a^{i_L})(v(1),\ldots,v(L)) \\
&\quad= \frac{1}{L!} \sum_{P \in \mathcal{P}_L} sgn(P)(a^{i_1} \otimes \cdots \otimes a^{i_L})(v(P(1)),\ldots,v(P(L))) \\
&\quad= \frac{1}{L!} \sum_{P \in \mathcal{P}_L} sgn(P)\, v(P(1))^{i_1} \cdots v(P(L))^{i_L} \\
&\quad= \frac{1}{L!} \sum_{Q \in \mathcal{P}_L} sgn(Q)\, v(1)^{i_{Q(1)}} \cdots v(L)^{i_{Q(L)}} \\
&\quad= \frac{1}{L!} \sum_{Q \in \mathcal{P}_L} sgn(Q)(a^{i_{Q(1)}} \otimes \cdots \otimes a^{i_{Q(L)}})(v(1),\ldots,v(L)).
\end{aligned}$$

In a basis, if $T = T_{i_1,\ldots,i_L}(\mathbf{a})\, a^{i_1} \otimes \cdots \otimes a^{i_L}$ then

$$\begin{aligned}
Alt(T) &= \frac{1}{L!} \sum_{Q \in \mathcal{P}_L} sgn(Q)\, T_{i_1,\ldots,i_L}(\mathbf{a})\, a^{i_{Q(1)}} \otimes \cdots \otimes a^{i_{Q(L)}} \\
&= \frac{1}{L!} \sum_{Q \in \mathcal{P}_L} sgn(Q)\, T_{i_{Q(1)},\ldots,i_{Q(L)}}(\mathbf{a})\, a^{i_1} \otimes \cdots \otimes a^{i_L}.
\end{aligned}$$

(Sum on an index pair i_m and $i_{Q(q)}$ whenever $m = Q(q)$.)

So the permutations can be applied to the domain factors or to the basis superscripts in a basis representation or to the coefficients in a basis to calculate Alt(T). Obviously, the same fact pertains for Sym(T).

It is worthwhile to check that $Sym(T)$ is symmetric, while $Alt(T)$ is alternating. It is also true that if T is symmetric then $Sym(T) = T$ and $Alt(T) = 0$. If T is alternating then $Sym(T) = 0$ and $Alt(T) = T$.

Our choice of scaling constant $\frac{1}{L!}$ in the definitions of Sym and Alt makes these linear functions, when restricted to tensors of a specified rank, into projections onto the spaces of symmetric and alternating tensors of that rank, respectively.

Choice of constant here and in the wedge product below are not universally adopted, even within the work of a single author, and in each new source you must note the convention used there. Formulas differ by various "factorial" factors from text to text as a regrettable consequence.

Note that the coefficients of any covariant 2-tensor T can be written as

$$T_{i,j} = \frac{T_{i,j} + T_{j,i}}{2} + \frac{T_{i,j} - T_{j,i}}{2}$$

so any such tensor is the sum of a symmetric and an alternating tensor.

But if the rank exceeds 2 it is *not true* that every covariant or contravariant relative tensor is the sum of an alternating with a symmetric tensor, and at this point it might be a useful exercise to verify the following facts concerning $T = a^1 \otimes a^2 \otimes a^3 - a^2 \otimes a^1 \otimes a^3$ and $S = T - Alt(T)$.

- T is neither symmetric nor alternating yet $Sym(T) = 0$.
- S is not symmetric and not alternating yet $Alt(S) = Sym(S) = 0$.
- Conclude that S is not the sum of a symmetric with an alternating tensor.

If $\mathbf{S}$ and $\mathbf{T}$ are both covariant relative tensors, of orders s and t respectively, we define the **wedge product of S with T** (in that order) by

$$\boxed{S \wedge T = \frac{(s+t)!}{s!\ t!}\, Alt(S \otimes T).}$$

The wedge product is in some sources called the **exterior product,** not to be confused with the ordinary tensor (outer) product.

It is not too hard to show that if $\mathbf{S}$ and $\mathbf{R}$ are both covariant relative tensors of order s with the same weighting function and if $\mathbf{T}$ is a covariant relative tensor of order t and if k is a constant then

$$\boxed{S \wedge T = (-1)^{st}\, T \wedge S \qquad \text{and} \qquad (kS + R) \wedge T = kS \wedge T + R \wedge T.}$$

If R is a covariant relative tensor of order r we would like to know that $(R \wedge S) \wedge T = R \wedge (S \wedge T)$: that is, wedge product is associative. This is a bit harder to show, but follows from a straightforward, if messy, calculation.

In preparation we state the following facts about permutations. Suppose $P \in \mathcal{P}_k$ and $0 < k \leq L$. Define $\overline{P} \in \mathcal{P}_L$ to be the permutation

$$\overline{P}(i) = \begin{cases} P(i), & \text{if } i \leq k; \\ i, & \text{if } \ L \geq i > k. \end{cases}$$

First, any member Q of $\mathcal{P}_L$ is $Q \circ \overline{P}^{\,-1} \circ \overline{P}$ and so can be written in $k!$ different ways as $R \circ \overline{P}$ where $R = Q \circ \overline{P}^{\,-1}$ for various $P \in \mathcal{P}_k$.

Second, $sgn\left(\overline{P}\right) = sgn(P) = sgn\left(\overline{P}^{\,-1}\right)$ and $sgn\left(R\circ\overline{P}\right) = sgn(R)\, sgn(P)$.

$$\begin{aligned}
&Alt\left(Alt\left(a^{i_1}\otimes\cdots\otimes a^{i_k}\right)\otimes a^{i_{k+1}}\otimes\cdots\otimes a^{i_L}\right)\\
&=Alt\left(\frac{1}{k!}\sum_{P\in\mathcal{P}_k} sgn(P)\, a^{i_{P(1)}}\otimes\cdots\otimes a^{i_{P(k)}}\otimes a^{i_{k+1}}\otimes\cdots\otimes a^{i_L}\right)\\
&=\frac{1}{k!\;L!}\sum_{R\in\mathcal{P}_L,\;P\in\mathcal{P}_k} sgn(R)\, sgn\left(\overline{P}\right)\; a^{i_{R\circ P(1)}}\otimes\cdots\otimes a^{i_{R\circ P(k)}}\otimes a^{i_{R(k+1)}}\otimes\cdots\otimes a^{i_{R(L)}}\\
&=\frac{1}{k!\;L!}\sum_{R\in\mathcal{P}_L,\;P\in\mathcal{P}_k} sgn\left(R\circ\overline{P}\right)\; a^{i_{R\circ\overline{P}(1)}}\otimes\cdots\otimes a^{i_{R\circ\overline{P}(L)}}.
\end{aligned}$$

If we focus on a fixed permutation $i_{Q(1)},\ldots,i_{Q(L)}$ of the superscripts in the last double sum above we see that each is repeated $k!$ times, once for permutation $R = Q\circ\overline{P}^{\,-1}$ for each permutation P. So the last double sum above is

$$\begin{aligned}
&\frac{1}{k!\;L!}\sum_{Q\in\mathcal{P}_L,\;P\in\mathcal{P}_k} sgn(Q)\, a^{i_{Q(1)}}\otimes\cdots\otimes\cdots\otimes a^{i_{Q(L)}}\\
&\qquad=\frac{1}{k!\;L!}\;k!\sum_{Q\in\mathcal{P}_L} sgn(Q)\, a^{i_{Q(1)}}\otimes\cdots\otimes\cdots\otimes a^{i_{Q(L)}} = Alt\left(a^{i_1}\otimes\cdots\otimes a^{i_L}\right).
\end{aligned}$$

The obvious modification shows also that

$$Alt\left(a^{i_1}\otimes\cdots\otimes a^{i_k}\otimes Alt\left(a^{i_{k+1}}\otimes\cdots\otimes a^{i_L}\right)\right) = Alt\left(a^{i_1}\otimes\cdots\otimes a^{i_L}\right).$$

Suppose $R = R_{i_1,\ldots,i_r}(\mathbf{a})\, a^{i_1}\otimes\cdots\otimes a^{i_r}$ and $S = S_{j_1,\ldots,j_s}(\mathbf{a})\, a^{j_1}\otimes\cdots\otimes a^{j_s}$ and $T = T_{k_1,\ldots,k_r}(\mathbf{a})\, a^{k_t}\otimes\cdots\otimes a^{k_t}$ are any covariant tensors.

By linearity of Alt and the result above we find that

$$\boxed{Alt(Alt(R\otimes S)\otimes T) = Alt(R\otimes S\otimes T) = Alt(R\otimes Alt(S\otimes T)).}$$

We remark in passing that the argument above is easily adapted to show that Sym satisfies:

$$\boxed{Sym(Sym(R\otimes S)\otimes T) = Sym(R\otimes S\otimes T) = Sym(R\otimes Sym(S\otimes T)).}$$

From these facts about Alt, it follows immediately that

$$\boxed{R\wedge(S\wedge T) = \frac{(r+s+t)!}{r!\;s!\;t!}\;Alt(R\otimes S\otimes T) = (R\wedge S)\wedge T.}$$

Since the operation $\wedge$ is associative, we have found that an expression involving three or more wedge factors, such as $R\wedge S\wedge T$, is unambiguous. We note also that this wedge product is an alternating tensor, and its order is the sum of the orders of its factors.

As a special but common case, if $\beta^1,\ldots,\beta^L\in V^*$ an induction argument gives

$$\beta^1\wedge\cdots\wedge\beta^L = \sum_{Q\in\mathcal{P}_L} sgn(Q)\,\beta^{Q(1)}\otimes\cdots\otimes\beta^{Q(L)}.$$

So this product of 1-forms can be evaluated on $v_1,\ldots,v_L\in V$ by

$$\left(\beta^1 \wedge \cdots \wedge \beta^L\right)(v_1, \dots, v_L) = \sum_{Q \in \mathcal{P}_L} sgn(Q) \left(\beta^{Q(1)} \otimes \cdots \otimes \beta^{Q(L)}\right))(v_1, \dots, v_L)$$
$$= \sum_{Q \in \mathcal{P}_L} sgn(Q)\, \beta^{Q(1)}(v_1) \cdots \beta^{Q(L)}(v_L) = \det\left(\beta^i(v_j)\right)$$

where **det** stands for the **determinant** function[5], in this case applied to the $L \times L$ matrix with ith row, jth column entry $\beta^i(v_j)$.

And it follows that if $Q \in \mathcal{P}_L$ then the wedge product of 1-forms satisfies

$$\beta^1 \wedge \cdots \wedge \beta^L = sgn(Q)\, \beta^{Q(1)} \wedge \cdots \wedge \beta^{Q(L)}.$$

Every part of the discussion in this section works with minor modifications if all tensors involved are contravariant rather than covariant tensors, and these do occur in applications.

As with tensor products, the weight of a wedge product of relative covariant or contravariant tensors is the sum of the weights of the factors and the wedge product of relative tensors will be odd exactly when an odd number of the factor relative tensors are odd.

15. A Basis for $\Lambda_r(V)$

Define $\mathbf{\Lambda_s(V)}$ to be the subset of $\mathcal{T}^0_s(V)$ consisting of tensors T with $Alt(T) = T$: that is, $\Lambda_s(V)$ is the set of alternating members of $\mathcal{T}^0_s(V)$, a condition trivially satisfied when $s = 0$ or 1.

Members of $\Lambda_s(V)$ are called **s-forms**, specifically the s-forms on V.

The alternating members of $\mathcal{T}^s_0(V)$, denoted $\Lambda^s(V)$. They are called **s-vectors**. Note that 1-forms on V^* are identified with V, so $\Lambda_1(V^*) = \Lambda^1(V) = V$ and, generally $\Lambda^s(V) = \Lambda_s(V^*)$.

We confine our work, mostly, to $\Lambda_s(V)$ in this section but invite the reader to make the obvious notational adaptations needed to state and prove the same results in $\Lambda^s(V)$.

If $S \in \Lambda_s(V)$ and $T \in \Lambda_t(V)$ there is no reason to suppose that $S \otimes T$, which is in $\mathcal{T}^0_{s+t}(V)$, will be alternating. However $S \wedge T$ *will* be alternating, so wedge product defines an operation on forms that produces a form as output.

Suppose $R \in \Lambda_r(V)$. In an ordered basis $\mathbf{a}$, R is written

$$R = R_{i_1,\dots,i_r}(\mathbf{a})\, a^{i_1} \otimes \cdots \otimes a^{i_r}$$

where, because R is alternating, the coordinates satisfy the antisymmetry condition

$$R_{i_1,\dots,i_r}(\mathbf{a}) = sgn(Q)\, R_{i_{Q(1)},\dots,i_{Q(r)}}(\mathbf{a}) \quad \text{for } Q \in \mathcal{P}_r.$$

Because of antisymmetry, if there is any repeat of an index value within one term the coefficient of that term must be 0. Nonzero coefficients must be among

[5] The sum in the last line is usually taken to be the definition of the determinant of an $L \times L$ matrix with these numerical entries.

those in which the list of index values consists of r distinct integers. This means that any alternating tensor of order exceeding n must be the zero tensor. So let's focus on a particular list $i_1, \dots, i_r$ and for specificity we choose them so that $1 \le i_1 < \dots < i_r \le n$. (This implies, of course, that $r \le n$.) A list of indices like this will be called **increasing**.

The part of the sum given for R in the $\mathbf{a}$ coordinates involving any way of rearranging this particular list of integers is

$$\begin{aligned}
&\sum_{Q \in \mathcal{P}_r} R_{i_{Q(1)},\dots,i_{Q(r)}}(\mathbf{a})\ a^{i_{Q(1)}} \otimes \dots \otimes a^{i_{Q(r)}} && \text{(Sum on } Q \text{ only.)}\\
&\quad = R_{i_1,\dots,i_r}(\mathbf{a}) \sum_{Q \in \mathcal{P}_r} sgn(Q)\ a^{i_{Q(1)}} \otimes \dots \otimes a^{i_{Q(r)}} && \text{(Sum on } Q \text{ only.)}\\
&\quad = R_{i_1,\dots,i_r}(\mathbf{a})\ a^{i_1} \wedge \dots \wedge a^{i_r}. && \text{(This is not a sum.)}
\end{aligned}$$

We learn two important things from this.

First, **$\mathbf{\Lambda_r(V)}$ is spanned by wedge products of the form**

$$a^{i_1} \wedge \dots \wedge a^{i_r}. \quad \text{(These are increasing indices.)}$$

Second, when we gather together all the terms as above, **the coefficient on $\mathbf{a^{i_1}} \wedge \dots \wedge \mathbf{a^{i_r}}$ (increasing indices) is the same as the coefficient on $\mathbf{a^{i_1}} \otimes \dots \otimes \mathbf{a^{i_r}}$ in the representation of alternating R as a member of $\mathcal{T}^0_r(V)$ using ordered basis a**.

It remains to show that these increasing-index wedge products are linearly independent, and so form a basis of $\Lambda_r(V)$.

Suppose we have a linear combination of these tensors

$$T_{i_1,\dots,i_r}(\mathbf{a})\ a^{i_1} \wedge \dots \wedge a^{i_r} \quad \text{(Sum here on increasing indices only.)}$$

Evaluating this tensor at a particular $(a_{j_1}, \dots, a_{j_r})$ with $1 \le j_1 < \dots < j_r \le n$ we find that all terms are 0 except possibly

$$T_{j_1,\dots,j_r}(\mathbf{a})\ a^{j_1} \wedge \dots \wedge a^{j_r}(a_{j_1}, \dots, a_{j_r}) \quad \text{(This is not a sum.)}$$

Expanding this wedge product in terms of Alt and evaluating, we find that this is $T_{j_1,\dots,j_r}(\mathbf{a})$. So if this linear combination is the zero tensor, every coefficient is 0: that is, these increasing wedge products form a linearly independent set.

So tensors of the form $a^{i_1} \wedge \dots \wedge a^{i_r}$ where $1 \le i_1 < \dots < i_r \le n$ constitute a basis of $\Lambda_r(V)$, the **standard basis of $\mathbf{\Lambda_r(V)}$ for ordered basis a.**

$$\boxed{\begin{aligned}
&R_{i_1,\dots,i_r}(\mathbf{a})\ a^{i_1} \otimes \dots \otimes a^{i_r} && (R \text{ is alternating, sum here on all indices.})\\
&\qquad = R_{i_1,\dots,i_r}(\mathbf{a})\ a^{i_1} \wedge \dots \wedge a^{i_r}. && \text{(Sum here on increasing indices only.)}
\end{aligned}}$$

There are $\frac{n!}{(n-r)!}$ ways to form r-tuples of distinct numbers between 1 and n.

There are $r!$ distinct ways of rearranging any particular list of this kind so there are $\frac{n!}{(n-r)!\ r!}$ ways of selecting r increasing indices from among n index values.

So **the dimension of $\Lambda_r(V)$ is**

$$\frac{n!}{(n-r)!\ r!} \ = \ \binom{n}{r}.$$

Note that if $r = n$ or $r = 0$ the dimension is 1. If $r > n$ the dimension of $\Lambda_r(V)$ is 0. $\Lambda_1(V) = \mathcal{T}^0_1(V) = V^*$ so if $r = 1$ this dimension is n, and this is the case when $r = n - 1$ as well.

Though not used as often in applications, we might as well consider the space of ***symmetric tensors of covariant order*** **r** ***and find it's dimension*** too.

Any symmetric covariant order r tensor can be written as a linear combination of tensors of the form

$$Sym(a^{i_1} \otimes \cdots \otimes a^{i_r})$$

for arbitrary choice of $i_1, \ldots, i_r$ between 1 and n.

But any permutation of a particular list choice will yield the same symmetrized tensor here so we examine the collection of all $Sym(a^{i_1} \otimes \cdots \otimes a^{i_r})$ where $i_1, \ldots, i_r$ is a *nondecreasing list of indices.*

This collection constitutes a linearly independent set by the similar argument carried out for alternating tensors above, and hence constitutes a basis for the space of symmetric tensors of order r. We want to count how many lists $i_1, \ldots, i_r$ satisfy

$$1 \leq i_1 \leq i_2 \leq i_3 \leq \cdots \leq i_r \leq n.$$

At first blush, counting how many ways this can be done seems to be a much harder proposition than counting the number of *strictly increasing* sequences as we did for alternating tensors, but it can be converted into that calculation by examining the equivalent sequence

$$1 \leq i_1 < i_2 + 1 < i_3 + 2 < \cdots < i_r + r - 1 \leq n + r - 1.$$

Thus, **the dimension of the space of symmetric tensors of order r is**

$$\frac{(n+r-1)!}{(n-1)!\ r!} \ = \ \binom{n+r-1}{r}.$$

Returning from this symmetric-tensor interlude, **there is a simpler formula (or, at least, one with fewer terms) for calculating the wedge product of two tensors known, themselves, to be alternating. This involves "shuffle permutations."**

Suppose s, t and L are positive integers and $L = s + t$. An **s-shuffle** in $\mathcal{P}_L$ is a permutation P in $\mathcal{P}_L$ with $P(1) < P(2) < \cdots < P(s)$ and also $P(s+1) < P(s+2) < \cdots < P(L)$.

It is clear that for each permutation Q there are $s!\,t!$ different permutations P for which the sets $\{Q(1), Q(2), \ldots, Q(s)\}$ and $\{P(1), P(2), \ldots, P(s)\}$ are equal, and only *one* of these is a shuffle permutation.

We let the symbol $\mathcal{S}_{s,L}$ denote the set of these shuffle permutations.

If $\mathbf{a}$ is an ordered basis of V and S and T are alternating

$$S \wedge T = \frac{L!}{s!\,t!}\, Alt(S \otimes T) = \frac{1}{s!\,t!} \sum_{P \in \mathcal{P}_L} sgn(P)(S \otimes T)_P$$

$$= \frac{1}{s!\,t!} \sum_{P \in \mathcal{P}_L} sgn(P) S_{i_1,\ldots,i_s} T_{i_{s+1},\ldots,i_L} \left(\left(a^{i_1} \wedge \cdots \wedge a^{i_s}\right) \otimes \left(a^{i_{s+1}} \wedge \cdots \wedge a^{i_L}\right) \right)_P .$$

This is a huge sum involving indices for which $i_1, \ldots, i_s$ is in increasing order and also $i_{s+1}, \ldots, i_L$ is in increasing order, as well as the sum over permutations.

Suppose $i_1, \ldots, i_s, i_{s+1}, \ldots, i_L$ is a particular choice of such indices.

If any index value among the $i_1, \ldots, i_s$ is repeated in $i_{s+1}, \ldots, i_L$ the result after summation on the permutations will be 0 so we may discard such terms and restrict our attention to terms with entirely distinct indices.

If P and Q are two permutations in $\mathcal{P}_L$ and $\{Q(1), Q(2), \ldots, Q(s)\}$ is the same set as $\{P(1), P(2), \ldots, P(s)\}$ then P can be transformed into Q by composition with a sequence of 2-cycles that switch members of the common set $\{Q(1), Q(2), \ldots, Q(s)\}$, followed by a sequence of 2-cycles that switch members of the set $\{Q(s+1), Q(s+2), \ldots, Q(L)\}$. Each of these switches, if applied to a term in the last sum above, would introduce a minus sign factor in one of the two wedge product factors, and also introduce a minus sign factor into $sgn(P)$. The net result is no change, so each permutation can be transformed into an s-shuffle permutation without altering the sum above. Each shuffle permutation would be repeated $s!\,t!$ times.

We conclude:

$$S \wedge T = \sum_{P \in \mathcal{S}_{s,L}} sgn(P)(S \otimes T)_P$$

when S, T are alternating, and where $\mathcal{S}_{s,L}$ are the s-shuffle permutations.

For instance, if we have two single-term-wedge-product alternating tensors and $r = 3$ and $s = 2$ this sum is over 10 permutations rather than 120. If $r = 1$ and $s = 3$ we are summing over 4 permutations rather than 24.

16. Determinants and Orientation

The single tensor $a^1 \wedge \cdots \wedge a^n$ spans $\Lambda_n(V)$ when V has dimension n. If σ is any member of $\Lambda_n(V)$ then, in ordered basis $\mathbf{a}$, σ has only one coordinate $\sigma_{1,\ldots,n}(\mathbf{a})$. It is common to skip the subscript entirely in this case and write $\sigma = \sigma(\mathbf{a})\, a^1 \wedge \cdots \wedge a^n$.

We can be explicit about the effect of σ on vectors in V. Suppose $H_j = A^i(H_j) a_i$ for $j = 1, \ldots, n$ are generic members of V. Form matrix $H(\mathbf{a})$ with ijth entry $H(\mathbf{a})^i_j = A^i(H_j)$. $H(\mathbf{a})$ is the $n \times n$ matrix formed from the columns $A(H_j)$.

Suppose that $\sigma \in \Lambda_n(V)$. Then for some constant k,

$$\sigma(H_1, \ldots, H_n) = \; k \left(a^1 \wedge \cdots \wedge a^n\right)(H_1, \ldots, H_n) = k \det(H(\mathbf{a})).$$

You may recall various properties we cited earlier from [8] p.154: $\det(MN) = \det(M)\det(N)$ for compatible square matrices M and N, $\det(\mathcal{I}) = 1$ and the fact that $\det(N) = \det(M^{-1}NM)$ when M and N are compatible and M invertible.

A couple of other facts are also useful when dealing with determinants.

First, by examining the effect of replacing permutations P by $Q = P^{-1}$ we see that

$$\det(M) = \sum_{P\in\mathcal{P}_n} sgn(P)\; M_1^{P(1)} \cdots M_n^{P(n)} = \sum_{Q\in\mathcal{P}_n} sgn(Q)\; M^1_{Q(1)} \cdots M^n_{Q(n)}$$

which implies that the determinant of a matrix and its transpose are equal.

Second, define $\mathbf{M^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,j-1,j+1,\dots,n}}$ to be the $(n-1)\times(n-1)$ matrix obtained from the square matrix M by deleting the ith row and jth column. In [8] p.146 (and also in this work on page 42) it is shown that

$$\det(M) = \sum_{i=1}^{n} (-1)^{i+j}\; M^i_j\; \det\left(M^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,j-1,j+1,\dots,n}\right) \quad \text{for each } j = 1,\dots,n.$$

Calculating the determinant this way, in terms of smaller determinants for a particular j, is called a **Laplace expansion** or, more descriptively, "expanding the determinant around the jth column." Coupled with the remark concerning the transpose found above we can also "expand the determinant around any row."

Returning now from this determinant theory interlude, if $H_j = a_j$ for each j then $H(\mathbf{a})$ is the identity matrix and so $\det(H(\mathbf{a})) = 1$ and we identify the constant k as $\sigma(a_1, \cdots, a_n) = \sigma(\mathbf{a})$. So for any choice of these vectors H_j

$$\sigma(H_1,\dots,H_n) = \sigma(a_1,\cdots,a_n)\; \det(H(\mathbf{a})) = \sigma(\mathbf{a})\; \det(H(\mathbf{a})).$$

In addition to its utility as a tool to calculate $\sigma(H_1,\dots,H_n)$, this provides a way of classifying bases of V into two groups, called **orientations**. Given any particular nonzero σ, if $\sigma(a_1,\dots,a_n)$ is positive the ordered basis $\mathbf{a}$ belongs to one orientation, called the **orientation determined by** σ. If $\sigma(a_1,\dots,a_n)$ is negative the ordered basis $\mathbf{a}$ is said to belong to the **opposite orientation**. This division of the bases into two groups does not depend on σ (though the signs associated with each group could switch.)

17. Change of Basis for Alternating Tensors and Laplace Expansion

If we change coordinates to another ordered basis $\mathbf{b}$ the coordinates of the r-form R change as

$$R_{j_1,\dots,j_r}(\mathbf{b}) = R_{h_1,\dots,h_r}(\mathbf{a})\, \mathcal{A}^{h_1}_{j_1} \cdots \mathcal{A}^{h_r}_{j_r}.$$

We can choose $j_1,\dots,j_r$ to be increasing indices if we are thinking of R as a member of $\Lambda_r(V)$ but the sum over $h_1,\dots,h_r$ is over all combinations of indices, not just increasing combinations. We will find a more convenient expression.

Let's focus on a particular increasing combination of indices $i_1,\dots,i_r$. Then each term in the sum on the right involving these particular indices in any specified order is of the form

$$\begin{aligned} &R_{i_{Q(1)},\dots,i_{Q(r)}}(\mathbf{a})\, \mathcal{A}^{i_{Q(1)}}_{j_1} \cdots \mathcal{A}^{i_{Q(r)}}_{j_r} && \text{(This is not a sum.)}\\ &\qquad = R_{i_1,\dots,i_r}(\mathbf{a})\, sgn(Q)\, \mathcal{A}^{i_{Q(1)}}_{j_1} \cdots \mathcal{A}^{i_{Q(r)}}_{j_r} && \text{(This is not a sum.)} \end{aligned}$$

for some permutation $Q \in \mathcal{P}_r$. You will note that the sum over all permutations yields $R_{i_1,\dots,i_r}(\mathbf{a})$ times the determinant of the $r \times r$ matrix obtained from $\mathcal{A}$ by selecting rows $i_1, \dots, i_r$ and columns $j_1, \dots, j_r$. If you denote this matrix $\mathcal{A}^{\mathbf{i_1,\dots,i_r}}_{\mathbf{j_1,\dots,j_r}}$ we obtain the change of basis formula in $\Lambda_r(V)$:

$$R_{j_1,\dots,j_r}(\mathbf{b}) = R_{i_1,\dots,i_r}(\mathbf{a}) \det\left(\mathcal{A}^{i_1,\dots,i_r}_{j_1,\dots,j_r}\right) \quad \text{(Sum on increasing indices only.)}$$

It is worth noting here that if T is an r-vector (an alternating member of $\mathcal{T}^r_0(V)$) then by a calculation identical to the one above we have

$$T^{j_1,\dots,j_r}(\mathbf{b}) = T^{i_1,\dots,i_r}(\mathbf{a}) \det\left(\mathcal{B}^{j_1,\dots,j_r}_{i_1,\dots,i_r}\right) \quad \text{(Sum on increasing indices only.)}$$

As examples, consider the n-form $\sigma = \sigma(\mathbf{a})\ a^1 \wedge \cdots \wedge a^n$ and the n-vector $x = x(\mathbf{a})\ a_1 \wedge \cdots \wedge a_n$. The change of basis formula[6] gives

$$\sigma(\mathbf{b}) = \sigma(\mathbf{a})\ \det(\mathcal{A}) \quad \text{and} \quad x(\mathbf{b}) = x(\mathbf{a})\ \det(\mathcal{B}).$$

This implies immediately

$$b^1 \wedge \cdots \wedge b^n = \det(\mathcal{B})\, a^1 \wedge \cdots \wedge a^n \quad \text{and} \quad b_1 \wedge \cdots \wedge b_n = \det(\mathcal{A})\, a_1 \wedge \cdots \wedge a_n$$

with interesting consequences when $\det(\mathcal{B}) = \det(\mathcal{A}) = 1$.

Let's see how this meshes with the calculation shown above for σ evaluated on the H_j.

Form matrix $H(\mathbf{b})$ from columns $B(H_j)$. Recall that $B(H_j) = \mathcal{B}A(H_j)$. This implies that $H(\mathbf{b}) = \mathcal{B}H(\mathbf{a})$ and so $\det(H(\mathbf{b})) = \det(\mathcal{B})\det(H(\mathbf{a}))$.

From above,

$$\begin{aligned}\sigma(H_1, \dots, H_n) &= \sigma(\mathbf{b})\ \det(H(\mathbf{b})) = \sigma(\mathbf{a})\ \det(\mathcal{A})\ \det(\mathcal{B}H(\mathbf{a}))\\ &= \sigma(\mathbf{a})\ \det(\mathcal{A})\ \det(\mathcal{B})\ \det(H(\mathbf{a})) = \sigma(\mathbf{a})\ \det(H(\mathbf{a})).\end{aligned}$$

We make an important observation: Suppose $h^1, \dots, h^r$ are members of V^*. If there is a linear relation among them then one of them, say h^1, can be written in terms of the others: $h^1 = c_2 h^2 + \cdots + c_r h^r$. Using standard summation notation

$$h^1 \wedge \cdots \wedge h^r = \sum_{i=2}^{r} c_i\, h^i \wedge h^2 \wedge \cdots \wedge h^r = 0$$

because each tensor in the sum is alternating and there is a repeated factor in each wedge product.

On the other hand, suppose $b^1, \dots, b^r$ is a linearly independent list of members of V^*. Then this list can be extended to an ordered basis of V^* dual to an ordered basis $\mathbf{b}$ of V. The n-form $b^1 \wedge \cdots \wedge b^n$ is nonzero so

$$0 \neq b^1 \wedge \cdots \wedge b^n = \left(b^1 \wedge \cdots \wedge b^r\right) \wedge \left(b^{r+1} \wedge \cdots \wedge b^n\right)$$

[6] Note the similarity to the change of basis formula for a tensor capacity in the case of n-forms, and tensor density in case of n-vectors. The difference, of course, is that *here* the determinant incorporates the combined effect of the matrices of transition, and is not an extra factor.

which implies that $b^1 \wedge \cdots \wedge b^r \neq 0$.

We conclude that

$$h^1 \wedge \cdots \wedge h^r = 0 \text{ if and only if } h^1, \dots, h^r \text{ is a dependent list in } V^*.$$

A similar result holds for **s**-vectors in $\Lambda_s(V^*)$.

We can use the change of basis formula and the associativity of wedge product to prove a generalization of the Laplace expansion for determinants. The shuffle permutations come in handy here.

Suppose $\mathcal{A}$ is an invertible $n \times n$ matrix and $\mathbf{b}$ is an ordered basis of V. So $\mathcal{A}$ can be regarded as a matrix of transition to a new basis $\mathbf{a}$.

From above, we see that $b_1 \wedge \cdots \wedge b_n = \det(\mathcal{A})\, a_1 \wedge \cdots \wedge a_n$.

But if $0 < r < n$, we also have $b_1 \wedge \cdots \wedge b_n = (b_1 \wedge \cdots \wedge b_r) \wedge (b_{r+1} \wedge \cdots \wedge b_n)$. Applying the change of base formula to the factors on the right, we see that

$$\begin{aligned}
&(b_1 \wedge \cdots \wedge b_r) \wedge (b_{r+1} \wedge \cdots \wedge b_n) \qquad \text{(Sum on increasing indices only below.)}\\
&= \left[\det\left(\mathcal{A}^{j_1,\dots,j_r}_{1,\dots,r}\right) a_{j_1} \wedge \cdots \wedge a_{i_r}\right] \wedge \left[\det\left(\mathcal{A}^{k_1,\dots,k_{n-r}}_{r+1,\dots,n}\right) a_{k_1} \wedge \cdots \wedge a_{k_{n-r}}\right].
\end{aligned}$$

The nonzero terms in the expanded wedge product consist of *all and only* those pairs of increasing sequences for which $j_1, \dots, j_r, k_1, \dots, k_{n-r}$ constitutes an r-shuffle permutation of $1, \dots, n$. So

$$\begin{aligned}
&\left[\det\left(\mathcal{A}^{j_1,\dots,j_r}_{1,\dots,r}\right) a_{j_1} \wedge \cdots \wedge a_{i_r}\right] \wedge \left[\det\left(\mathcal{A}^{k_1,\dots,k_{n-r}}_{r+1,\dots,n}\right) a_{k_1} \wedge \cdots \wedge a_{k_{n-r}}\right]\\
&\quad \text{(Sum on increasing indices only in the line above.)}\\
&\quad = \sum_{Q \in \mathcal{S}_{r,n}} \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{1,\dots,r}\right) \det\left(\mathcal{A}^{Q(r+1),\dots,Q(n)}_{r+1,\dots,n}\right) a_{Q(1)} \wedge \cdots \wedge a_{Q(n)}\\
&\quad \text{(Sum on shuffle permutations only in the line above and the line below.)}\\
&\quad = \sum_{Q \in \mathcal{S}_{r,n}} sgn(Q) \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{1,\dots,r}\right) \det\left(\mathcal{A}^{Q(r+1),\dots,Q(n)}_{r+1,\dots,n}\right) a_1 \wedge \cdots \wedge a_n.
\end{aligned}$$

We conclude that

$$\det(\mathcal{A}) = \sum_{Q \in \mathcal{S}_{r,n}} sgn(Q) \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{1,\dots,r}\right) \det\left(\mathcal{A}^{Q(r+1),\dots,Q(n)}_{r+1,\dots,n}\right).$$

Note that the Laplace expansion around the first column appears when $r = 1$.

More generally, if P is *any* permutation of $1, \dots, n$ then

$$\det(\mathcal{A}) = \sum_{Q \in \mathcal{S}_{r,n}} sgn(P)\, sgn(Q) \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{P(1),\dots,P(r)}\right) \det\left(\mathcal{A}^{Q(r+1),\dots,Q(n)}_{P(r+1),\dots,P(n)}\right)$$

which recovers, among other things, the Laplace expansion around any column.

18. Evaluation in $\Lambda_s(V)$ and $\Lambda^s(V)$: the Interior Product

One often sees a contraction or evaluation operator applied to s-forms or s-vectors, which we define now.

It is called the **interior product**, indicated here using the "angle" symbol, $\lrcorner$. It involves one vector and one s-form, or one covector and one s-vector.

If v is a vector and $c \in \Lambda_0(V) = \mathbb{R}$ we define the **"angle" operation** on c, called the **interior product of v with c**, by $v \lrcorner c = 0$.

More generally, for $\theta \in \Lambda_r(V)$ we define $v \lrcorner \theta$, the interior product of v with θ, to be the contraction of the tensor product of θ against v in the first index of θ.

This is the tensor we called $E(v,1)(\theta)$ in Section 13, and considered as an evaluation process it is obvious that $v \lrcorner \theta \in \Lambda_{r-1}(V)$ for each $r > 0$.

We note that interior product is linear, both in v and θ.

We examine a few cases and extract properties of this operation. We suppose coordinates below are in a fixed basis **a** of V.

If $\theta \in \Lambda_1(V)$ then $v \lrcorner \theta = \theta(v) = \theta_i \, v^i$.

If $\theta = \theta_{1,2} \, v^1 \wedge v^2 \in \Lambda_2(V)$ we have

$$\begin{aligned} v \lrcorner \theta \; = \; \theta(v, \cdot) \; &= \theta_{1,2} \, v^1 \, a^2 + \theta_{2,1} \, v^2 \, a^1 \\ &= \theta_{1,2} \left(v^1 \, a^2 - v^2 \, a^1 \right). \end{aligned}$$

If w is another vector,

$$w \lrcorner (v \lrcorner \theta) \;=\; \theta_{1,2} \, v^1 \, w^2 \, - \theta_{1,2} \, v^2 \, w^1 \;=\; -v \lrcorner (w \lrcorner \theta).$$

This readily generalizes to $\theta \in \Lambda_r(V)$ for any $r \geq 1$

$$\boxed{v \lrcorner \, (\, w \lrcorner \theta \,) = - \, w \lrcorner \, (\, v \lrcorner \theta \,)}$$

as can be seen by examining the pair of evaluations $\theta(v, w, \dots)$ and $\theta(w, v, \dots)$.

This implies that for any θ

$$v \lrcorner \, (\, v \lrcorner \theta \,) = 0.$$

You will note that

$$v \lrcorner \left(a^1 \wedge a^2 \right) = \left(v \lrcorner a^1 \right) \wedge a^2 + (-1)^1 \, a^1 \wedge \left(v \lrcorner a^2 \right).$$

We will prove a similar formula for θ of higher order.

If $\theta \in \Lambda_r(V)$ and $\tau \in \Lambda_s(V)$ for r and s at least 1 then

$$\boxed{v \lrcorner \, (\theta \wedge \tau) = (\, v \lrcorner \theta \,) \wedge \tau + (-1)^r \, \theta \wedge (\, v \lrcorner \tau \,)}$$

We prove the result when $v = a_1$. Since $\mathbf{a}$ is an arbitrary basis and interior product is linear in v, this will suffice for the general result.

We further presume that $\theta = a^{i_1} \wedge \cdots \wedge a^{i_r}$ and $\tau = a^{j_1} \wedge \cdots \wedge a^{j_s}$ for increasing indices. If we can prove the result in this case then by linearity of wedge product in each factor and linearity of the evaluation map we will have the result we seek.

We consider four cases. In each case, line (A) is $v \lrcorner\, (\theta \wedge \tau)$, line (B) is $(\, v \lrcorner\, \theta\,) \wedge \tau$, while line (C) is $(-1)^r\, \theta \wedge (\, v \lrcorner\, \tau\,)$. In each case, the first is the sum of the last two.

Case One: $i_1 = 1$ and $j_1 \neq 1$.

$$(A) \quad a_1 \lrcorner \left(a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge \cdots \wedge a^{j_s}\right) \;=\; a^{i_2} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge \cdots \wedge a^{j_s}.$$

$$(B) \quad \left(a_1 \lrcorner \left(a^{i_1} \wedge \cdots \wedge a^{i_r}\,\right)\right) \wedge a^{j_1} \wedge \cdots \wedge a^{j_s} \;=\; a^{i_2} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge \cdots \wedge a^{j_s}.$$

$$(C) \quad (-1)^r\, a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge \left(a_1 \lrcorner \left(a^{j_1} \wedge \cdots \wedge a^{j_s}\,\right)\right) \;=\; 0.$$

Case Two: $i_1 \neq 1$ and $j_1 = 1$.

$$\begin{aligned}
(A) \quad & a_1 \lrcorner \left(a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge \cdots \wedge a^{j_s}\right) \\
&= a_1 \lrcorner \left((-1)^r\, a^{j_1} \wedge a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_2} \wedge \cdots \wedge a^{j_s}\right) \\
&= (-1)^r\, a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_2} \wedge \cdots \wedge a^{j_s}.
\end{aligned}$$

$$(B) \quad \left(a_1 \lrcorner \left(a^{i_1} \wedge \cdots \wedge a^{i_r}\,\right)\right) \wedge a^{j_1} \wedge \cdots \wedge a^{j_s} \;=\; 0 \wedge a^{j_1} \wedge \cdots \wedge a^{j_s} \;=\; 0.$$

$$\begin{aligned}
(C) \quad & (-1)^r\, a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge \left(a_1 \lrcorner \left(a^{j_1} \wedge \cdots \wedge a^{j_s}\,\right)\right) \\
&= (-1)^r\, a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_2} \wedge \cdots \wedge a^{j_s}.
\end{aligned}$$

Case Three: $i_1 \neq 1$ and $j_1 \neq 1$.

$$(A) = (B) = (C) = 0.$$

Case Four: $i_1 = 1$ and $j_1 = 1$.

$$(A) \quad a_1 \lrcorner \left(a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge \cdots \wedge a^{j_s}\right) \;=\; a_1 \lrcorner\, 0 \;=\; 0.$$

$$(B) \quad \left(a_1 \lrcorner \left(a^{i_1} \wedge \cdots \wedge a^{i_r}\,\right)\right) \wedge a^{j_1} \wedge \cdots \wedge a^{j_s} \;=\; a^{i_2} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge \cdots \wedge a^{j_s}.$$

$$\begin{aligned}
(C) \quad & (-1)^r\, a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge \left(a_1 \lrcorner \left(a^{j_1} \wedge \cdots \wedge a^{j_s}\,\right)\right) \\
&= (-1)^r\, a^{i_1} \wedge \cdots \wedge a^{i_r} \wedge a^{j_2} \wedge \cdots \wedge a^{j_s} \\
&= (-1)^r\, (-1)^{r-1} a^{i_2} \wedge \cdots \wedge a^{i_r} \wedge a^{i_1} \wedge a^{j_2} \wedge \cdots \wedge a^{j_s} \\
&= (-1)\, a^{i_2} \wedge \cdots \wedge a^{i_r} \wedge a^{j_1} \wedge a^{j_2} \wedge \cdots \wedge a^{j_s}.
\end{aligned}$$

The interior product is defined for a covector-and-s-vector pair by identical means, yielding the analogous properties.

Applied to relative tensors, the weight of an interior product of relative tensors is, as with wedge and tensor products, the sum of the weights of the two factors. An interior product of relative tensors will be odd when one but not both of the two factors is odd.

19. Bilinear Forms

Since V^* and V are both n dimensional, there are many isomorphisms between them. The obvious isomorphism sending basis to dual basis is basis-dependent: it is a feature of the basis, not the space and will not in general reflect geometry, which usually comes from a nondegenerate bilinear form on V. This structure encapsulates physical properties and physical invariants of great interest to practical folk such as physicists and engineers. Sometimes it will correspond to the analogue of ordinary Euclidean distance in V. Sometimes it is created by a Lorentzian metric on V, coming perhaps from relativistic space-time considerations.

A **bilinear form on V** is simply a member of $\mathcal{T}_2^0(V)$.

Suppose $\mathbf{a}$ is an ordered basis for V. The coordinates of a bilinear form G in the ordered basis $\mathbf{a}$ are $G_{i,j}(\mathbf{a}) = G(a_i, a_j)$. Define the **matrix representation** of G with respect to ordered basis $\mathbf{a}$ to be the matrix $\mathbf{G_a}$ with ijth entry ${\mathbf{G_a}}^i_j = G_{i,j}(\mathbf{a})$.

We wish to distinguish between the tensor G with coordinates $G_{i,j}(\mathbf{a})$ in the ordered basis $\mathbf{a}$ and the matrix $\mathbf{G_a}$ with ijth entry $G_{i,j}(\mathbf{a})$. They refer to distinct (obviously related) ideas.

G is the tensor

$$G = G_{i,j}(\mathbf{a})\ a^i \otimes a^j.$$

If $v = v^i(\mathbf{a})\, a_i$ and $w = w^i(\mathbf{a})\, a_i$

$$G(v,w) = G\left(v^i(\mathbf{a})\, a_i, w^j(\mathbf{a})\, a_j\right) = G_{i,j}(\mathbf{a})\ v^i(\mathbf{a})\ w^j(\mathbf{a}) = A(v)^t \mathbf{G_a}\, A(w).$$

Evaluating $G(v,w)$ in $\mathbb{R}^n$ as $A(v)^t \mathbf{G_a}\, A(w)$ takes advantage of our common knowledge of how to multiply matrices, but *we are careful (and it is important) to note that the matrix* $\mathbf{G_a}$ *is not the representative tensor of* G *in ordered basis* $\mathbf{a}$. That representative tensor is

$$G_{i,j}(\mathbf{a})\ e^i \otimes e^j.$$

Any bilinear form can be written as

$$\begin{aligned} G &= G_{i,j}(\mathbf{a})\ a^i \otimes a^j \\ &= \frac{1}{2}\left(G_{i,j}(\mathbf{a}) a^i \otimes a^j + G_{i,j}(\mathbf{a}) a^j \otimes a^i\right) + \frac{1}{2}\left(G_{i,j}(\mathbf{a}) a^i \otimes a^j - G_{i,j}(\mathbf{a}) a^j \otimes a^i\right). \end{aligned}$$

The first term is symmetric while the second is alternating.

A bilinear form is called **degenerate** if there is $v \in V$ with

$$v \neq 0 \quad \text{and} \quad G(v,w) = 0 \quad \forall\ w \in V.$$

Otherwise the bilinear form is called **nondegenerate**.

The two most common types of bilinear forms are symplectic forms and inner products.

A **symplectic form** on V is an alternating nondegenerate bilinear form on V. Symplectic forms are used many places but one common application is in the **Hamiltonian** formulation of mechanics.

The matrix of a symplectic form (and in fact any 2-form) satisfies $\mathbf{G_a}^t = -\,\mathbf{G_a}$: that is, the matrix is **skew symmetric**. The fact that a symplectic form is nondegenerate implies that this matrix is invertible. The inverse matrix is also skew symmetric.

An **inner product on V** is a symmetric nondegenerate bilinear form on V. Sometimes an inner product is called a **metric tensor.**

A symmetric bilinear form is called **non-negative** if

$$G(v,v) \geq 0 \quad \forall \text{ nonzero } v \in V$$

and **positive definite** if

$$G(v,v) > 0 \quad \forall \text{ nonzero } v \in V.$$

Negative definite and **non-positive** symmetric bilinear forms are defined by making the obvious adaptations of these definitions.

We **specifically do not require** our inner product to be positive definite.

But if an inner product *is* positive definite it can be used to define lengths of vectors and angles between vectors by calculations which are identical to the case of the standard inner product in $\mathbb{R}^n$. Some of the geometrical ideas spawned by **positive definite** symmetric G are listed below.

The **length of a vector** v is defined to be $\|\mathbf{v}\|_{\mathbf{G}} = \sqrt{\mathbf{G(v,v)}}$. When the inner product G is understood by all we drop the subscript and write $\|\mathbf{v}\|$. We define the **angle between nonzero vectors v and w** to be

$$\text{The angle between nonzero } v \text{ and } w\text{:} \quad \theta_{v,w} = \arccos\left(\frac{G(v,w)}{\|v\|\,\|w\|}\right).$$

It is a bit of work to show that this definition of length scales with linear multiples as we would expect[7] and satisfies the Polarization Identity[8], the Parallelogram Law[9] and the Triangle[10] and Cauchy-Schwarz[11] Inequalities, and that an angle is always defined between pairs of nonzero vectors and never violates our understanding of how angles should behave.[12] We forgo the proofs of these facts.

A vector space endowed with this notion of distance and angle (i.e. one that comes from, or could have come from, a positive definite inner product) is called **Euclidean.**

The matrix of an inner product (and in fact any symmetric covariant tensor of rank 2) is symmetric: $\mathbf{G_a} = \mathbf{G_a}^t$. The inverse of the matrix of an inner product is symmetric too.

---

[7]Scaling of Length: $\|cv\| = |c|\,\|v\|$ for any $c \in \mathbb{R}$ and $v \in V$.

[8]Polarization Identity: $G(v,w) = \frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2$.

[9]Parallelogram Law: $\|v+w\|^2 + \|v-w\|^2 = 2\,\|v\|^2 + 2\,\|w\|^2$.

[10]Triangle Inequality: $\|v+w\| \leq \|v\| + \|w\|$.

[11]Cauchy-Schwarz Inequality: $|G(v,w)| \leq \|v\|\,\|w\|$

[12]For instance, if v, s and w are nonzero vectors we would hope that $\theta_{v,w} \leq \theta_{v,s} + \theta_{s,w}$ with equality when $s = cv + dw$ for positive real numbers c and d.

If G is nondegenerate then $\mathbf{G_a}$ must be invertible. Let $\overline{\mathbf{G_a}}$ be the inverse of $\mathbf{G_a}$. We will denote the ijth entry of $\overline{\mathbf{G_a}}$ by $\overline{\mathbf{G_a}}^{i}_{\ j} = G^{i,j}(\mathbf{a})$.

$\mathbf{G_a}$ and $\overline{\mathbf{G_a}}$ are inverse matrices, so

$$G^{i,k}(\mathbf{a})G_{k,j}(\mathbf{a}) \;=\; G_{i,k}(\mathbf{a})G^{k,j}(\mathbf{a}) \;=\; \delta^i_j$$

for each i and j.

Note bene: Matrices have entries M^i_j in the ith row and jth column. For bilinear forms **we associate these matrix entries with tensor coefficients whose indices are both high or both low.** With this convention we can utilize the Einstein summation notation in certain calculations we need to perform later.

20. An Isomorphism Induced by a Nondegenerate Bilinear Form

Suppose we are given a symplectic form or inner product $G = \; G_{i,j}(\mathbf{a})\; a^i \otimes a^j$ with matrix $\mathbf{G_a}$ in ordered basis $\mathbf{a}$ as indicated above.

The ijth entry of $\mathbf{G_a}$ is $G_{i,j}(\mathbf{a})$ and the ijth entry of the inverse matrix $\overline{\mathbf{G_a}}$ is a number we have denoted $G^{i,j}(\mathbf{a})$. This collection of numbers looks like the coefficients of a tensor, and we will see that it is.

In this section we will be working with a single basis $\mathbf{a}$ for V, and will temporarily suspend explicit reference to that basis in most coordinates.

We define $\flat\colon V \to V^*$, using the musical "flat" symbol, by

$$\flat(u)(w) = G(u,w) = A(u)^t\; \mathbf{G_a}\; A(w) \quad \forall\, w \in V.$$

If (and, generally, only if) $\mathbf{G_a}$ is the identity matrix, left multiplication by $A(u)^t\mathbf{G_a}$ corresponds to simple "dot product" involving the coordinates of u against the coordinates of v.

If $\flat(u)$ is the zero transformation then u must be 0 since G is nondegenerate. So $\flat$ is one-to-one and since the vector spaces V and V^* have the same dimension, $\flat$ is an isomorphism onto V^*.

$$\text{If } u = u^i\; a_i \in V \quad \text{then} \quad \flat(u) = G_{i,j}\; u^i\; a^j \in V^*.$$

Consider the following calculation, and its implication in the box below.

$$\flat(\;\sigma_m\; G^{m,w}\; a_w\;) = G_{w,j}\;\sigma_m\; G^{m,w}\; a^j = \sigma_m\;\delta^m_j\; a^j = \sigma_j\; a^j = \sigma.$$

Define the function $\sharp\colon V^* \to V$, using the musical "sharp" symbol, by

$$\sharp(\sigma) \;=\; G^{i,j}\;\sigma_i\; a_j \text{ for } \sigma = \sigma_i\; a^i \in V^*.$$

We see that $\quad \sharp \;=\; \flat^{-1}$.

So $\sharp(\sigma)$ evaluated on τ will, generally, look like the sum of simple products $\sigma_i \tau_i$ only when $\overline{\mathbf{G_a}}$ (and so $\mathbf{G_a}$) is the identity matrix.

$\flat$ was defined without reference to a basis, and so both it and its inverse $\sharp$ have meaning independent of basis. The formulas we create involving their effect on vectors and covectors in a basis are valid in *any* basis.

We use the musical "flat" symbol $\flat$ and the "sharp" symbol $\sharp$ as mnemonic aids. $\flat$ lowers or "flattens" the index on a coefficient turning a vector, usually associated with an arrow, into a covector, which is usually associated with a family of parallel planes (or hyperplanes.) The $\sharp$ symbol raises the index, turning a "flat" object into a "sharp" one.

We can use $\flat$ to create $\mathbf{G}^\star$, a tensor of order 2, on V^*. $G^\star$ will be a member of $\mathcal{T}_0^2(V)$ and is defined by:

$$G^\star(\sigma, \tau) = G\left(\sharp(\tau), \sharp(\sigma)\right) \quad \forall\ \sigma, \tau \in V^*.$$

The reversal of order you see in σ and τ above is not relevant when G is symmetric, as it usually will be. In the symplectic case, this reversal avoids a minus sign in the formula.

If $\sigma = \sigma_i\ a^i$ and $\tau = \tau_i\ a^i$ then

$$\begin{aligned}(\sigma, \tau) =& G\left(\tau_m\ G^{m,w}\ a_w\ ,\ \sigma_L\ G^{L,s}\ a_s\right)\\ =& \tau_m\ \sigma_L\ G^{m,w}\ G^{L,s}\ G_{w,s} = \tau_m\ \sigma_L\ \delta_s^m\ G^{L,s}\\ =& \sigma_L\ \tau_m\ G^{L,m}.\end{aligned}$$

So $G^\star$ is the tensor

$$G^\star = G^{L,m}\ a_L \otimes a_m.$$

The tensor $G^\star$ is said to be the inner product **conjugate** to G.

We can evaluate $G^\star$ down in $\mathbb{R}^{n*}$ as:

$$G^\star(\sigma, \tau) = \overline{A^*}(\sigma)\ \overline{\mathbf{G_a}}\ \overline{A^*}(\tau)^t.$$

We call $\overline{\mathbf{G_a}}$ the matrix of $G^\star$.

The bilinear form G and its conjugate $G^\star$ induce tensors on $\mathbb{R}^n$ and $\mathbb{R}^{n*}$ by transporting vectors to V and V^* using $\overline{A}$ and A^*.

$$\tilde{\tilde{\mathbf{A}}}G(x, y)\ =\ G\left(\ \overline{A}(x), \overline{A}(y)\ \right)$$

and

$$\tilde{\tilde{\mathbf{A}}}G^\star(\sigma, \tau)\ =\ G^\star\left(\ A^*(\sigma), A^*(\tau)\ \right)$$

We have produced the nondegenerate tensor $G^\star$ defined on $V^* \times V^*$ and an isomorphism $\flat$ from V to V^*. We can carry this process forward, mimicking our construction of $G^\star$ and $\flat$, to produce a nondegenerate bilinear form on $V^{**} = V$ and an isomorphism from V^* to V.

You will find that this induced bilinear form is G (i.e. the tensor conjugate to $G^\star$ is G) and the isomorphism is $\sharp$.

21. Raising and Lowering Indices

Suppose, once again, we are given a nondegenerate 2-form G with representation $G_{i,j}(\mathbf{a})\ a^i \otimes a^j$ in ordered basis $\mathbf{a}$. Though the initial calculations make no symmetry assumptions, in practice G will be an inner product or, possibly, a symplectic form.

In this section as in the last we will be working with a single basis $\mathbf{a}$ for V and will temporarily suspend reference to that basis in most coordinates.

The functions $\flat \colon V \to V^*$ and $\sharp = \flat^{-1}$ and the conjugate tensor $G^\star$ from the last section were defined using G alone and so are not dependent on coordinate systems, though we have found representations for their effect in a basis.

$\flat$ and $\sharp$ can be used to modify tensors in a procedure called "**raising or lowering indices**."

This process is entirely dependent on a fixed choice of G and extends the procedure by which we created $G^\star$ from G.

We illustrate the process on vectors and covectors before introducing a notational change to help us with higher orders.

Suppose $\tau = \tau_i\ a^i$ is in V^*.

$\sharp$ is an isomorphism from V^* to V, and τ is linear on V so the composition $\tau \circ \sharp$ is a linear real valued function on V^*: that is, a vector.

Define the "raised index version of τ" to be this vector, and denote its coordinates in basis $\mathbf{a}$ by τ^k so $\tau \circ \sharp = \tau^k\, a_k$.

If $\sigma = \sigma_i\ a^i$ is in V^*,

$$\tau \circ \sharp(\sigma) = \tau_j\, a^j(\sharp(\sigma)) = \tau_j\, a^j \left(\ G^{k,L}\, \sigma_k\ a_L\ \right) = \tau_j\, G^{k,j}\, \sigma_k = \tau_j\, G^{k,j}\ a_k(\sigma).$$

So $\tau^k = \tau_j\, G^{k,j}$ for each k. From this you can see that $\tau \circ \sharp$ could be obtained by the tensor product of τ with $G^\star$ followed by a contraction against the second index of $G^\star$.

Suppose $x = x^i\ a_i$ is in V.

The linear function $\flat$ has domain V and range V^* and x has domain V^* and range $\mathbb{R}$ so the composition $x \circ \flat$ is a covector.

If $v = v^k a_k \in V$,

$$x \circ \flat(v) = x^j\, a_j(\flat(v)) = x^j\, a_j \left(\ G_{k,L}\, v^k\ a^L\ \right) = x^j\, G_{k,j}\, v^k = x^j\, G_{k,j}\ a^k(v).$$

Define the "lowered index version of x" to be this covector, and denote its coordinates in basis $\mathbf{a}$ by x_k so $x \circ \flat = x_k\, a^k$.

As you can see, the covector $x \circ \flat$ could have been obtained by tensor product of x with G followed by contraction against the second index of G.

We have, in compact form, for vectors and covectors:

$$\boxed{x_k = x^j\, G_{k,j} \quad \text{and} \quad \tau^k = \tau_j\, G^{k,j} \quad \text{(independent of basis)}}$$

Since $\flat = \sharp^{-1}$ we find that $x \circ \flat \circ \sharp = x$ and $\tau \circ \sharp \circ \flat = \tau$ so raising a lowered index and lowering a raised index gets you back where you started.

There is a common ambiguity which can cause problems and must be mentioned, centered around using the same base symbol x for coordinates x^i and also coordinates x_i. That invites the practice of using the same symbol x to denote both the vector and the covector obtained by lowering its index, and a similar practice for a covector τ and its raised index version. The version which is intended must be deduced from context or made explicit somehow.

We will now introduce the notational change mentioned above.

The domain of our tensors in $\mathcal{T}^r_s(V)$ was $V^* \times \cdots \times V^* \times V \times \cdots \times V$ where all the copies of V are listed last. But nothing we did required that they be last. It was a notational convenience only that prompted this. It was the number of factors of V and V^* which mattered, not where they were on the list. It is more convenient for the remainder of this section for us to allow V and V^* to appear anywhere as factors in the domain of a tensor and we create a notation to accommodate this.

A tensor formerly indicated as $\mathbf{T^{i_1,\dots,i_r}_{j_1,\dots,j_s}}$ in $\mathcal{T}^r_s(V)$ will now look like $\mathbf{T^{i_1\dots i_r}{}_{j_1\dots j_s}}$ where we give a separate column in the block of coordinates to the right of T for each factor in the domain. Factors of V^* in the domain correspond to raised indices while lowered ones correspond to domain factors V. The "column" notation does not use the comma to separate indices.

So a tensor with coordinates $\mathbf{T^{a}{}_{b}{}^{c}{}_{de}{}^{fg}}$ has domain $V^* \times V \times V^* \times V \times V \times V^* \times V^*$ and has contravariant order 4 and covariant order 3. To recover our previous situation merely slide all factors of V^* to the left and pack the the indices to the left in order, inserting commas where they are required.

Suppose T is any tensor of order k, contravariant of order r and covariant of order s, with coordinates given in this column notation. The domain of T is a set W which is a product of copies of V and V^* in some order: we do not require all the copies of V^* to be on the left. We presume that the ith index of T is contravariant: that is, the index in column i is "high" and that V^* is the ith factor of W.

We let W_i be the factor space obtained by substituting V in place of V^* at the ith factor.

Define the function $\flat_i \colon W_i \to W$ by

$$\flat_i \left(\, (v(1), \dots, v(k)) \, \right) \; = \; \left(\, v(1), \dots, \flat(v(i)), \dots, v(k) \, \right).$$

In other words, $\flat_i$ leaves a member of W unchanged except in its ith factor, where the vector $v(i)$ is replaced by the covector $\flat(v(i))$.

We now consider the tensor $T \circ \flat_i \colon W_i \to \mathbb{R}$. This tensor is still of rank k but now its contravariant rank is $r-1$ and its covariant rank is $s+1$. It is said to be a **version of T whose ith index is low**, and the process of forming $T \circ \flat_i$ from T is called **lowering the ith index on T**.

Similarly, if the jth index of T is covariant then the index in column j is "low" and V is the jth factor of W.

We let W^j be the factor space obtained by substituting V^* in place of V as the jth factor.

Define the function $\sharp^j : W^j \to W$ by

$$\sharp^j \; (\; (v(1), \dots, v(k)) \;) \; = \; (\; v(1), \dots, \sharp(v(j)), \dots, v(k) \;).$$

In other words, $\sharp^j$ leaves a member of W unchanged except in its jth factor, where the covector $v(j)$ is replaced by the vector $\sharp(v(j))$.

We now consider the tensor $T \circ \sharp^j : W^j \to \mathbb{R}$. This tensor has contravariant rank $r+1$ and covariant rank $s-1$. It is said to be a **version of T whose jth index is high**, and the process of forming $T \circ \sharp^j$ from T is called **raising the jth index on T**.

Generally, for a fixed tensor of order k there will be 2^k different index position choices. Confusingly, all of these different tensors can be, and in some treatments routinely are, denoted by the same symbol T.

Presumably, you will start out with one specific known version of T as above. It will have, perhaps, coordinates $T^i{}_j$. The other three versions in this rank-2 case will be

- $T \circ \flat_1$ with coordinates denoted $T_{i\,j}$
- $T \circ \sharp_2$ with coordinates denoted $T^{i\,j}$
- $T \circ \flat_1 \circ \sharp_2$ with coordinates denoted $T_i{}^j$.

When dealing with coordinates, there can be no ambiguity here. T has a column for each index and a position (high or low) for an index in each column. If the index is not where it should be it must have been raised by $\sharp$ or lowered by $\flat$ at that domain factor. The problems arise when coordinate-free statements are preferred and in that case using the same symbol for all versions of T is **bad notation**. The practice is, however, ubiquitous.

By means identical to the rank 1 case considered above, we find the relationship between coordinates of the various versions of rank-2 T to be related by

$$T_{i\,j} \;= T_i{}^L \; G_{j,L} \;=\; T^L{}_j \; G_{i,L} \;=\; T^{L\,M} \; G_{i,L} \; G_{j,M}$$

and

$$T^{i\,j} \;= T^i{}_L \; G^{j,L} \;=\; T_L{}^j \; G^{i,L} \;=\; T_{L\,M} \; G^{i,L} \; G^{j,M}.$$

Once again, to raise an index on a version of T you take the tensor product of that version with $\mathbf{G}^\star$ and contract against the second index of $\mathbf{G}^\star$. To lower an index on a version of T you take the tensor product of that version with G and contract at the second index of G.

If G (and hence $G^\star$) is symmetric the choice of index of contraction on G or $G^\star$ is irrelevant. If G is *not* symmetric, it does matter.

The statement in bold above applies to tensors of any rank.

We can apply this process to the tensor G itself. Raising both indices gives coordinates

$$G_{Lk} \; G^{j,k} \; G^{i,L} = \delta^i_k \; G^{j,k} = G^{j,i} = G^{i,j}.$$

It is in anticipation of this result that we used the raised-index notation for the coordinates of $G^\star$. The tensor obtained from G by raising both indices using G

actually is the conjugate metric tensor, so we have avoided a clash of notations. We also note that $G^k{}_i = \delta^k_i = G_i{}^k$.

It is possible to **contract a tensor by two contravariant or two covariant index positions**, by first raising or lowering one of the two and then contracting this new pair as before. **We restrict our attention here to symmetric G**.

We illustrate the process with an example. Let us suppose given a tensor with coordinates $T^a{}_b{}^c{}_{de}{}^{fg}$, and that we wish to contract by the two contravariant indices in the $6th$ and $7th$ columns. We first lower the 7th index to form

$$T^a{}_b{}^c{}_{de}{}^f{}_g = T^a{}_b{}^c{}_{de}{}^{fh}\, G_{gh}.$$

We then contract to form $C^{67}T$ as

$$(C^{67}T)^a{}_b{}^c{}_{de} = T^a{}_b{}^c{}_{de}{}^k{}_k = T^a{}_b{}^c{}_{de}{}^{kh}\, G_{kh}.$$

Done in the other order we have

$$T^a{}_b{}^c{}_{dek}{}^k = T^a{}_b{}^c{}_{de}{}^{hk}\, G_{kh}$$

which is the same by symmetry of G.

Contracting in the 4th and 5th index positions, which are covariant, is defined by first raising one of these indices and then contracting at the new pair.

$$(C_{45}T)^a{}_b{}^{cfg} = T^a{}_b{}^c{}_k{}^{kfg} = T^a{}_b{}^c{}_{hk}{}^{fg}\, G^{kh}.$$

In this case too, symmetry of G guarantees that the result does not depend on which of these two indices was raised.

To put some of this in perspective, **we consider a very special case of an ordered basis a and an inner product G with matrix $\mathbf{G_a} = \mathcal{I} = \overline{\mathbf{G_a}}$, the identity matrix.** This means that the inner product in $\mathbb{R}^n$ created by G using this ordered basis is the ordinary **Euclidean inner product**, or **dot product**. If T is a member of V (that is, a contravariant tensor of order 1) or a member of V^* (a covariant tensor of order 1) the act of raising or lowering the index corresponds, in $\mathbb{R}^n$, to taking the transpose of the representative there. The numerical values of coordinates of any tensor do not change when an index is raised or lowered in this very special but rather common case.

Once again **restricting attention to symmetric G**, indices can be raised or lowered on alternating tensors just as with any tensor, though the result could not be alternating unless *all* indices are raised or lowered *together.* In that case, though, the transformed tensor *will* be alternating as we now show.

If R is an r-form, it can be represented in a basis as $R_{j_1 \dots j_r}\, a^{j_1} \otimes \cdots \otimes a^{j_r}$ where the sum is over all indices, or as $R_{j_1 \dots j_r}\, a^{j_1} \wedge \cdots \wedge a^{j_r}$ where the sum is over increasing indices.

In the first representation, the fact that R is alternating is encapsulated in the equation

$$R_{j_1 \dots j_r} = sgn(Q)\, R_{j_{Q(1)} \dots j_{Q(r)}} \quad \text{for any permutation } Q \in \mathcal{P}_r.$$

Raising all indices on R yields

$$R^{\, j_1\, j_2 \dots j_r} = R_{i_1\, i_2 \dots i_r}\, G^{i_1\, j_1} G^{i_2\, j_2} \cdots G^{i_r\, j_r}$$

where the sums on the right are over all indices.

Raising indices with j_1 and j_2 permuted gives

$$\begin{aligned} R^{\,j_2\, j_1 \dots j_r} &= R_{i_1\, i_2 \dots i_r}\, G^{i_1\, j_2} G^{i_2\, j_1} \cdots G^{i_r\, j_r} = -R_{i_2\, i_1 \dots i_r}\, G^{i_1\, j_2} G^{i_2\, j_1} \cdots G^{i_r\, j_r} \\ &= -\, R_{i_2\, i_1 \dots i_r}\, G^{i_2\, j_1} G^{i_1\, j_2} \cdots G^{i_r\, j_r} = -R^{\,j_1\, j_2 \dots j_r} \end{aligned}$$

so the raised index version of R is alternating as stated.

We recall a notation we used earlier and let $\overline{\mathbf{G_a}}\,{}^{i_1,\dots,i_r}_{j_1,\dots,j_r}$ denote the matrix obtained from $\overline{\mathbf{G_a}}$ by selecting rows $i_1, \dots, i_r$ and columns $j_1, \dots, j_r$.

By means identical to that used for the change of basis formula for wedge products found on page 41 we conclude that for alternating R

$$R^{\,j_1 \dots j_r} = R_{i_1 \dots i_r} \det\left(\overline{\mathbf{G_a}}\,{}^{i_1,\dots,i_r}_{j_1,\dots,j_r}\right) \qquad \text{(Sum here on increasing indices only.)}$$

We observe also that if T is an r-vector (an alternating member of $\mathcal{T}^r_0(V)$) then we can lower all indices using matrix G rather than $\overline{G}$ to obtain for alternating T

$$T_{j_1 \dots j_r} = T^{\,i_1 \dots i_r} \det\left(\mathbf{G_a}\,{}^{j_1,\dots,j_r}_{i_1,\dots,i_r}\right) \qquad \text{(Sum here on increasing indices only.)}$$

Finally, the various processes we have described can be adapted with the obvious modifications to raise or lower the index on a relative tensor. The procedure does not alter the weighting function.

22. Four Facts About Tensors of Order 2

In this section we will consider how to calculate all four types of tensors of order 2 on V and V^* using matrices acting in specific ways on the representatives of vectors and covectors in $\mathbb{R}^n$ and $\mathbb{R}^{n*}$. In this text, these matrices are called the **matrix representations of the relevant tensors.**

We will then show how the matrix representations change as the basis changes from basis $\mathbf{a}$ to $\mathbf{b}$ in V.

Then we will make three different kinds of observations about these matrices.

We will examine how these matrices are related in a fixed ordered basis if the tensors involved are all "raised or lowered versions" of the same tensor with respect to an inner product.

Then we will comment on how these matrices change when the matrices of transition are of a special type—orthogonal matrices.

Finally, we will examine the determinants of these matrices.

For the following calculations v, w are generic members of V, while θ, τ are to be generic members of V^*.

Suppose P is a tensor with domain $V \times V$. We saw that $P_{ij}(\mathbf{b}) = P_{kL}(\mathbf{a})\, \mathcal{A}^k_i\, \mathcal{A}^L_j$. We will determine this fact directly to show exactly how this carries over to a calculation with matrices in $\mathbb{R}^n$.

Let matrix $M_I(\mathbf{a})$ be the matrix with ijth entry $P_{ij}(\mathbf{a})$ and $M_I(\mathbf{b})$ be the matrix with ijth entry $P_{ij}(\mathbf{b})$.

$$\begin{aligned} P(v,w) =& P_{ij}(\mathbf{a})\ A^i(v)\ A^j(w) \\ =& A(v)^t\ M_I(\mathbf{a})\ A(w) \\ =& \left(\mathcal{A}\ B(v)\right)^t\ M_I(\mathbf{a})\ \mathcal{A}\ B(w) \\ =& B(v)^t\ \mathcal{A}^t\ M_I(\mathbf{a})\ \mathcal{A}\ B(w). \end{aligned}$$

We conclude that

$$\boxed{\text{Case I} \qquad M_I(\mathbf{b}) = \mathcal{A}^t\ M_I(\mathbf{a})\ \mathcal{A} \qquad \text{where } {M_I}\,{}^i_j = P_{ij}.}$$

Suppose Q is a tensor with domain $V \times V^*$. Then $Q_i{}^j(\mathbf{b}) = Q_k{}^L(\mathbf{a})\ \mathcal{A}^k_i\ \mathcal{B}^j_L$.

Let matrix $M_{II}(\mathbf{a})$ be the matrix with ijth entry $Q_i{}^j(\mathbf{a})$ and $M_{II}(\mathbf{b})$ be the matrix with ijth entry $Q_i{}^j(\mathbf{b})$.

$$\begin{aligned} Q(v,\theta) =& Q_i{}^j(\mathbf{a})\ A^i(v)\ \overline{A^*}_j(\theta) \\ =& A(v)^t\ M_{II}(\mathbf{a})\ \overline{A^*}(\theta)^t \\ =& \left(\mathcal{A}\ B(v)\right)^t\ M_{II}(\mathbf{a})\ \left(\overline{B^*}(\theta)\ \mathcal{B}\right)^t \\ =& B(v)^t\ \mathcal{A}^t\ M_{II}(\mathbf{a})\ \mathcal{B}^t\ \overline{B^*}(\theta)^t. \end{aligned}$$

We conclude that

$$\boxed{\text{Case II} \qquad M_{II}(\mathbf{b}) = \mathcal{A}^t\ M_{II}(\mathbf{a})\ \mathcal{B}^t \qquad \text{where } {M_{II}}\,{}^i_j = Q_i{}^j.}$$

Suppose R is a tensor with domain $V^* \times V$. Then $R^i{}_j(\mathbf{b}) = R^k{}_L(\mathbf{a})\ \mathcal{B}^i_k\ \mathcal{A}^L_j$.

Let matrix $M_{III}(\mathbf{a})$ be the matrix with ijth entry $R^i{}_j(\mathbf{a})$ and $M_{III}(\mathbf{b})$ be the matrix with ijth entry $R^i{}_j(\mathbf{b})$.

$$\begin{aligned} R(\theta,v) =& R^i{}_j(\mathbf{a})\ \overline{A^*}_i(\theta)\ A^j(v) \\ =& \overline{A^*}(\theta)\ M_{III}(\mathbf{a})\ A(v) \\ =& \overline{B^*}(\theta)\ \mathcal{B}\ M_{III}(\mathbf{a})\ \mathcal{A}\ B(v). \end{aligned}$$

We conclude that

$$\boxed{\text{Case III} \qquad M_{III}(\mathbf{b}) = \mathcal{B}\ M_{III}(\mathbf{a})\ \mathcal{A} \qquad \text{where } {M_{III}}\,{}^i_j = R^i{}_j.}$$

Suppose S is a tensor with domain $V^* \times V^*$. Then $S^{ij}(\mathbf{b}) = S^{kL}(\mathbf{a})\ \mathcal{B}^i_k\ \mathcal{B}^j_L$.

Let matrix $M_{IV}(\mathbf{a})$ be the matrix with ijth entry $S^{ij}(\mathbf{a})$ and $M_{IV}(\mathbf{b})$ be the matrix with ijth entry $S^{ij}(\mathbf{b})$.

$$\begin{aligned}
S(\theta,\tau) =& S_i{}^j(\mathbf{a})\ \overline{A^*}_i(\theta)\ \overline{A^*}_j(\tau)\\
=& \overline{A^*}(\theta)\ M_{IV}(\mathbf{a})\ \overline{A^*}(\tau)^t\\
=& \overline{B^*}(\theta)\ \mathcal{B}\ M_{IV}(\mathbf{a})\ (\overline{B^*}(\tau)\ \mathcal{B})^t\\
=& \overline{B^*}(\theta)\ \mathcal{B}\ M_{IV}(\mathbf{a})\ \mathcal{B}^t\ \overline{B^*}(\tau)^t.
\end{aligned}$$

We conclude that

Case IV $\quad M_{IV}(\mathbf{b}) = \mathcal{B}\ M_{IV}(\mathbf{a})\ \mathcal{B}^t \quad$ where $M_{IV}{}^{\,i}_{\,j} = S^{ij}$.

We will now consider the situation **where these four tensors are obtained by raising or lowering indices of (any) one of them with the services of an inner product G**.

We will let $M_I = (P_{ij})$ as above and relate the other matrices from above to this one. So

$$\begin{aligned}
P_i{}^j(\mathbf{a}) &= P_{ik}(\mathbf{a})\,G^{kj}(\mathbf{a}) & M_{II}(\mathbf{a}) &= M_I(\mathbf{a})\,\overline{\mathbf{G_a}}\\
P^i{}_j(\mathbf{a}) &= G^{ik}(\mathbf{a})\,P_{kj}(\mathbf{a}) & M_{III}(\mathbf{a}) &= \overline{\mathbf{G_a}}\,M_I(\mathbf{a})\\
P^{ij}(\mathbf{a}) &= G^{ik}(\mathbf{a})\,P_{kL}(\mathbf{a})\,G^{Lj}(\mathbf{a}) & M_{IV}(\mathbf{a}) &= \overline{\mathbf{G_a}}\,M_I(\mathbf{a})\,\overline{\mathbf{G_a}}
\end{aligned}$$

Since $\mathbf{G_a}^{-1} = \overline{\mathbf{G_a}}$ you can modify the above to write any of them in terms of the others and the matrices $\mathbf{G_a}$ and $\overline{\mathbf{G_a}}$.

We make explicit note of this calculation in an important and common case:

When these four tensors are all "raised or lowered" versions of each other, and if $\mathbf{G_a}$ and $\overline{\mathbf{G_a}}$ are the identity matrices, then the matrix of each of these tensors is the same.

A basis for which matrix $\mathbf{G_a}$ (and so $\overline{\mathbf{G_a}}$ too) is the identity matrix is called **orthonormal with respect to G**. We will have more to say about such bases later.

We now consider "special" matrices of transition.

Sometimes the change of basis matrix $\mathcal{A}$ satisfies

$$\mathcal{A}^t = \mathcal{A}^{-1} = \mathcal{B} \quad \text{and so} \quad \mathcal{B}^t = \mathcal{B}^{-1} = \mathcal{A}.$$

This happens exactly when $\mathcal{A}\mathcal{A}^t = \mathcal{I}$ where $\mathcal{I}$ is the $n \times n$ identity matrix, or

$$\sum_{k=1}^{n} \mathcal{A}^k_i \mathcal{A}^k_j = \delta^i_j.$$

The columns of $\mathcal{A}$ form an **orthonormal ordered basis of $\mathbb{R}^{\mathbf{n}}$ with respect to the usual inner product** there. Similarly, so do the rows. Matrices like this are called **orthogonal**.

The matrix representations of the four types of tensors of order 2, the original cases I through IV, all transform in the same way under coordinate changes with orthogonal matrices of transition.

Finally, we take determinants of the equations in cases I through IV.

The determinants of matrices $\mathcal{A}$ and $\mathcal{A}^t$ are the same. So are the determinants of matrices $\mathcal{B}$ and $\mathcal{B}^t$. The determinants of matrices $\mathcal{A}$ and $\mathcal{B}$ are reciprocals of each other.

So for the matrices of the mixed tensors, Cases II and III, we find that

$$\det\left(M_{II}(\mathbf{b})\right) = \det\left(M_{II}(\mathbf{a})\right) \quad \text{and} \quad \det\left(M_{III}(\mathbf{b})\right) = \det\left(M_{III}(\mathbf{a})\right).$$

The determinants of these matrices are invariant.

On the other hand, in Cases I and IV, the purely contravariant and covariant cases, respectively, these determinants do depend on basis. We find that

$$\det\left(M_{I}(\mathbf{b})\right) = \left(\det(\mathcal{A})\right)^2 \det\left(M_{I}(\mathbf{a})\right)$$

and

$$\det\left(M_{IV}(\mathbf{b})\right) = \left(\det(\mathcal{B})\right)^2 \det\left(M_{IV}(\mathbf{a})\right).$$

Of particular interest is the case where the underlying tensors are versions of an inner product G. So

$$M_I(\mathbf{a}) = \mathbf{G_a} \quad \text{and} \quad M_I(\mathbf{b}) = \mathbf{G_b}$$

while

$$M_{IV}(\mathbf{a}) = \overline{\mathbf{G_a}} \quad \text{and} \quad M_{IV}(\mathbf{b}) = \overline{\mathbf{G_b}}.$$

and all four matrices are invertible.

In this case we have

$$\boxed{\det(\mathcal{A}) = sign(\mathcal{A})\sqrt{\frac{\det\left(\mathbf{G_b}\right)}{\det\left(\mathbf{G_a}\right)}}}$$

and

$$\boxed{\det(\mathcal{B}) = sign(\mathcal{B})\sqrt{\frac{\det\left(\overline{\mathbf{G_b}}\right)}{\det\left(\overline{\mathbf{G_a}}\right)}}.}$$

23. Matrices of Symmetric Bilinear Forms and 2-Forms

We discuss here various details regarding matrix representation of symmetric bilinear forms and 2-forms which can be useful in calculations. Both types of rank 2 tensors may be present in an application, each contributing to different aspects of the geometry.

We accumulate several facts first.

The tensor G we will work with is given, in basis $\mathbf{a}$, by $G_{i,j}(\mathbf{a})a^i \otimes a^j$.

Define the set $$\begin{aligned} W &= \{\, v \in V \mid G(v,w) = 0 \text{ for all } w \in V \,\} \\ &= \{\, v \in V \mid G(w,v) = 0 \text{ for all } w \in V \,\}. \end{aligned}$$

If $W = \{0\}$ then G is either an inner product or a symplectic form.

If $W \neq \{0\}$ then an ordered basis $w_1, \ldots, w_k$ of W can be extended to an ordered basis $y_1, \ldots, y_{n-k}, w_1, \ldots, w_k$ of V.

If Y is the span of $y_1, \ldots, y_{n-k}$ then G restricted to Y is either an inner product or a symplectic form, and the last k columns and k rows of the matrix of G with respect to the basis $y_1, \ldots, y_{n-k}, w_1, \ldots, w_k$ of V contain only zero entries.

Recall from section 22 that the matrix $\mathbf{G_b}$ for bilinear form G in ordered basis $\mathbf{b}$ can be obtained from the matrix $\mathbf{G_a}$ in ordered basis $\mathbf{a}$ as $\mathbf{G_b} = \mathcal{A}^t \mathbf{G_a} \mathcal{A}$.

If $\mathcal{A}$ is any orthogonal matrix and $\mathcal{A}^t M \mathcal{A} = N$ the matrices N and M are called **orthogonally equivalent**.

We will start with symmetric bilinear form G.

The matrix of any symmetric bilinear form is symmetric. In [8] p.369 we find that **any symmetric matrix is orthogonally equivalent to a diagonal matrix** with r negative diagonal entries listed first, then s positive diagonal entries and zero diagonal entries listed last. We find there that the numbers r and s of negative and positive entries on the diagonal of $\mathbf{G_a}$ is independent of basis when $\mathbf{G_a}$ is diagonal, a result referred to as **Sylvester's Law of Inertia.**

s is the dimension of any subspace of V of maximal dimension upon which G is positive definite. r is the dimension of any subspace of V of maximal dimension upon which G is negative definite. $r+s$ is the dimension of any subspace of V of maximal dimension upon which G is nondegenerate. $r+s$ is the rank of any matrix representation of G.

The ordered pair (s, r) is called the **signature of G** and the number $s - r$ is sometimes called the **index of G**. Be warned: this vocabulary is not used in a consistent way in the literature.

Writing this out in more detail, suppose that $\mathbf{G_a}$ is the matrix of symmetric G with respect to $\mathbf{a}$. Then there is an orthogonal matrix $\mathcal{B}$, which can be used as a matrix of transition from ordered basis $\mathbf{a}$ to a new ordered basis $\mathbf{b}$, and for which (by orthogonality) and $\mathcal{B}^t = \mathcal{B}^{-1} = \mathcal{A}$ and, finally, for which the matrix $\mathbf{G_b}$ is diagonal.

$$\mathbf{G_b} = \mathcal{A}^t \mathbf{G_a} \mathcal{A} = \mathcal{A}^{-1} \mathbf{G_a} \mathcal{A}$$

$$= \begin{pmatrix} G_{1,1}(\mathbf{b}) & & & & & & & & \\ & \ddots & & & & & & & \\ & & G_{r,r}(\mathbf{b}) & & & & & & \\ & & & G_{r+1,r+1}(\mathbf{b}) & & & & & \\ & & & & \ddots & & & & \\ & & & & & G_{r+s,r+s}(\mathbf{b}) & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

The orthogonal matrix $\mathcal{B}$ can be chosen so that

$$G(b_i, b_j) = e_i^t \ \mathbf{G_b} \ e_j = G_{i,j}(\mathbf{b}) = 0 \text{ unless } i = j$$

and also

$$G = \sum_{i=1}^{n} G_{i,i}(\mathbf{b}) \ b^i \otimes b^i \text{ where } G_{i,i}(\mathbf{b}) \ \text{ is } \begin{cases} \text{negative}, & \text{if } 1 \leq i \leq r; \\ \text{positive} \ , & \text{if } r+1 \leq i \leq r+s; \\ \ \ 0, & \text{if } r+s+1 \leq i \leq n. \end{cases}$$

Now suppose G is an inner product on V: that is, $r + s = n$.

Taking the ordered basis $\mathbf{b}$ as above create a new ordered basis $\mathbf{c}$ using diagonal matrix of transition

$$\mathcal{K} = \begin{pmatrix} \frac{1}{\sqrt{-G_{1,1}(\mathbf{b})}} & & & & & \\ & \ddots & & & & \\ & & \frac{1}{\sqrt{-G_{r,r}(\mathbf{b})}} & & & \\ & & & \frac{1}{\sqrt{G_{r+1,r+1}(\mathbf{b})}} & & \\ & & & & \ddots & \\ & & & & & \frac{1}{\sqrt{G_{n,n}(\mathbf{b})}} \end{pmatrix}$$

In other words, $c_j = \left(\frac{1}{\sqrt{\pm \ G_{j,j}(\mathbf{b})}} \right) b_j$ for each j, where the sign is chosen to make the root real. Note: $\mathcal{K}^t \neq \mathcal{K}^{-1}$, in general, so **we lose orthogonal equivalence at this point—and similarity too**.

The new matrix of G, obtained as $\mathbf{G_c} = \mathcal{K}^t \mathbf{G_b} \mathcal{K}$, is still diagonal and

$$\boxed{G = \sum_{i=1}^{n} G_{i,i}(\mathbf{c}) \ c^i \otimes c^i \ , \qquad G_{i,i}(\mathbf{c}) = \begin{cases} -1, & \text{if } 1 \leq i \leq r; \\ \ \ 1, & \text{if } r+1 \leq i \leq r+s = n. \end{cases}}$$

Whenever the matrix of inner product G with respect to an ordered basis is diagonal with only ± 1 on the diagonal **and with the negative entries first** the ordered basis is called **orthonormal with respect to this inner product**. It should be pointed out that some texts reverse this order.

There is an orthonormal ordered basis for any inner product.

Suppose **c** and **w** are any two orthonormal ordered bases and $\mathcal{H} = \left(h^i_j\right)$ is the matrix of transition to **w** from **c**. So

$$\begin{aligned} G(w_K, w_L) \;&=\; \begin{cases} -1, & \text{if } L = K \le r; \\ \;\;1, & \text{if } r < L = K \le n; \\ \;\;0, & \text{if } L \ne K; \end{cases} \\ &= G\left(h^i_K c_i, h^j_L c_j\right) \;=\; h^i_K h^j_L G\left(c_i, c_j\right) \\ &=\; -\sum_{i=1}^{r} h^i_K h^i_L + \sum_{i=r+1}^{n} h^i_K h^i_L. \end{aligned}$$

So the transpose of matrix $\mathcal{H}$ can be tweaked to produce $\mathcal{H}^{-1}$. Specifically, divide the transpose of $\mathcal{H}$ into four blocks at row and column r. Multiply the off-diagonal blocks by -1 and you have $\mathcal{H}^{-1}$.

The multiplication can be accomplished by multiplying the first r rows and then the first r columns of $\mathcal{H}^t$ by -1, and each such multiplication will affect the determinant by a factor of -1. The combined effect is to leave the determinant unchanged, and since $\det\left(\mathcal{H}^t\right) = \det\left(\mathcal{H}\right)$ we conclude that $\det\left(\mathcal{H}^{-1}\right) = \det\left(\mathcal{H}\right) = \pm 1$.

A matrix of transition $\mathcal{H}$ between two bases which are orthonormal with respect to any inner product has determinant ± 1.

Also, if G is positive definite (i.e. $r = 0$) then $\mathcal{H}^t \;=\; \mathcal{H}^{-1}$.

A matrix of transition $\mathcal{H}$ between two bases which are orthonormal with respect to any positive definite inner product is an orthogonal matrix: that is, it satisfies $\mathcal{H}^t = \mathcal{H}^{-1}$.

A diagonalization algorithm for any symmetric bilinear form by a change of basis from any specified basis is found using a modification of the **Gram-Schmidt process**, which we briefly outline below. Though the matrices of transition we create have determinant ± 1, they will not in general be orthogonal even in the positive definite case.

Suppose **a** is any ordered basis of n dimensional V and G is a symmetric bilinear form on V.

Define the set of basis vectors $S \;=\; \{\, a^i \in \mathbf{a} \mid G(a^i, a^j) = 0 \text{ for all } j = 1, \ldots, n \,\}$.

Note that it is quite possible for *all* the diagonal entries $G(a^i, a^i)$ of the matrix for G in this basis to be 0. The basis vectors in S (if there are any) are those a^i for which the *entire ith row* of that matrix is filled with zeroes, and by symmetry the entire ith column is filled with zeroes too.

We will now *re-order* **a** so that all these vectors occur *last*. If S contains k vectors the matrix of G in basis **a** has nothing but zeroes in the bottom $k \times n$ block, and nothing but zeroes in the rightmost $n \times k$ block. So the rank of this matrix cannot exceed $n - k$. Any reordering of members of an ordered basis involves a matrix of transition with determinant ± 1.

With this preliminary tweaking out of the way, we proceed with the algorithm. The first step involves three possible cases and the last two both lead to Step 2.

Step 1a: If the matrix of G is diagonal (in particular, it is diagonal if it is the zero matrix) we are done.

Step 1b: If the matrix of G is not diagonal and $G(a_i, a_i) \neq 0$ for some smallest i reorder the members of the ordered basis **a** if necessary so that $G(a_1, a_1) \neq 0$. Again, this reordering of members of the ordered basis involves a matrix of transition with determinant ± 1.

Step 1c: If the matrix of G is not diagonal and $G(a_i, a_i) = 0$ for all i then $G(a_1, a_i) \neq 0$ for some smallest i. Replace a_1 by $a_1 + a_i$. The span of the new list of basis vectors is still V and now $G(a_1, a_1) \neq 0$. This change of ordered basis involves a matrix of transition with determinant 1.

Step 2: Let

$$b_1 = a_1 \text{ and } b_i = a_i - \frac{G(a_i, a_1)}{G(a_1, a_1)}\, a_1 \text{ for } i = 2, \ldots, n.$$

The matrix of transition between these bases has determinant 1. Note that if v is a linear combination of the last $n - 1$ basis vectors $b_2, \ldots, b_n$ then $G(b_1, v) = 0$.

Step 3: The matrix of G in the ordered basis $b_1, \ldots, b_n$ has all zeros in the first row and in the first column except possibly for the first diagonal entry, which is $G(b_1, b_1)$. It also has the bottom k rows and the last k columns filled with zeroes.

Step 4: If this matrix is actually diagonal, we are done. If the matrix is not yet diagonal, reorder the $b_1, \ldots, b_{n-1}$ if necessary (leaving b_1 alone) so that $G(b_2, b_2) \neq 0$ or, if that is impossible, replace b_2 with $b_2 + b_i$ for some smallest i so that the inequality holds. Let

$$c_1 = b_1 = a_1 \text{ and } c_2 = b_2 \text{ and } c_i = b_i - \frac{G(b_i, b_2)}{G(b_2, b_2)}\, b_2 \text{ for } i = 3, \ldots, n.$$

Once again, the matrix of transition has determinant 1. This time the matrix of G in the ordered basis $c_1, \ldots, c_n$ has all zeros in the first *two* rows and the first *two* columns except for the first *two* diagonal entries, which are $G(c_1, c_1)$ and $G(c_2, c_2)$. Of course the last k rows and the right-most k columns are still filled with zeroes.

We carry on in the pattern of Step 4 for (at most) $n - 3 - k$ more steps yielding an ordered basis for V for which the matrix of G is diagonal. And we can do a final re-ordering, if we wish, so that the negative eigenvalues are listed first followed by the positive eigenvalues on the diagonal.

The matrix of transition which produces this diagonal form from a specific basis **a**, the product of all the intermediary matrices of transition, will have determinant ± 1 but it will *not*, in general, be orthogonal.

We conclude this section with a similar discussion for a 2-form S.

In this case any matrix of S is skew symmetric. We find in [1] p.163 that there is a unique (dependent on S of course) natural number r and there is an ordered basis $\mathbf{b}$ for which the matrix $\mathbf{S_b}$ has the form

$$\mathbf{S_b} = \begin{pmatrix} 0 & \mathcal{I} & 0 \\ -\mathcal{I} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{each } \mathcal{I} \text{ is an } r \times r \text{ identity matrix.}$$

If you start with a matrix $\mathbf{S_a}$ for any 2-form S in ordered basis $\mathbf{a}$ the new matrix is of the form $\mathbf{S_b} = \mathcal{A}^t \mathbf{S_a} \mathcal{A}$, just as above. Once again $\mathcal{A}^t \neq \mathcal{A}^{-1}$ (at least, not necessarily) so $\mathbf{S_b}$ **will not in general be either orthogonally equivalent or similar to $\mathbf{S_a}$**.

The rank of any matrix like $\mathbf{S_b}$ must be an even number, so if **S is a symplectic form on V the dimension of V must be even.**

$$\text{The entries of } \mathbf{S_b} \text{ are zero except for } \begin{cases} -1, & \text{row } r+i \text{ column } i \text{ for } 1 \le i \le r; \\ 1, & \text{row } i \text{ column } r+i \text{ for } 1 \le i \le r. \end{cases}$$

So with this basis all $S(b_i, b_j)$ are 0 except

$$S(b_{r+i}, b_i) = e_{r+i}^t \; G_{\mathbf{b}} \; e_i = -1 \quad \text{for } 1 \le i \le r$$

and

$$S(b_i, b_{r+i}) = e_i^t \; G_{\mathbf{b}} \; e_{r+i} = 1 \quad \text{for } 1 \le i \le r$$

So

$$S = \sum_{i=1}^{r} \left(b^i \otimes b^{r+i} - b^{r+i} \otimes b^i \right) = \sum_{i=1}^{r} \; b^i \wedge b^{r+i}.$$

There is an algorithm similar to the Gram-Schmidt process which will produce an ordered basis of the kind we are discussing for any 2-form.

Suppose $\mathbf{a}$ is any ordered basis of n dimensional V and S is a 2-form.

For the purpose of this construction only, we will call a matrix "skew-diagonal" if its entries are all zero except for 2×2 blocks along the diagonal of the form

$$\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} \text{ where } x \text{ is a real number.}$$

Our goal is to create from the ordered basis $\mathbf{a}$ a new ordered basis for which the matrix of S is skew-diagonal using matrices of transition whose determinants are 1 (not ± 1.) It is then an easy process to create a basis in the form specified above, though the last matrix of transition will not have determinant 1, necessarily.

Step 1: If the matrix of S is skew-diagonal (in particular, it is skew-diagonal if it is the zero matrix) we are done. Otherwise, S is not the zero tensor. Since $S(v, v) = 0$ for all $v \in V$, if S is not the zero tensor we must be able to reorder the members of the ordered basis $\mathbf{a}$ (if necessary) so that $S(a_1, a_2) \neq 0$. Because we are reordering basis members two at a time, this reordering can be done with a matrix of transition with determinant 1.

Step 2: Let

$$b_1 = a_1 \text{ and } b_2 = a_2 \text{ and } b_i = a_i - \frac{S(a_i, a_2)}{S(a_1, a_2)}\ a_1 + \frac{S(a_i, a_1)}{S(a_1, a_2)}\ a_2 \text{ for } i = 3, \dots, n.$$

The matrix of transition between these bases has determinant 1. If v is any linear combination of the last $n-2$ basis vectors $b_3, \dots, b_n$ then $S(b_1, v) = S(b_2, v) = 0$.

Step 3: The matrix of S in the ordered basis $b_1, \dots, b_n$ has all zeros in the first two rows and in the first two columns except for a 2×2 skew-diagonal block on the far upper left, whose entries are

$$\begin{pmatrix} 0 & S(b_1, b_2) \\ -S(b_1, b_2) & 0 \end{pmatrix}.$$

If the matrix of S in this new ordered basis is skew-diagonal we are done.

Step 4: If the matrix is not yet skew-diagonal, reorder the $b_1, \dots, b_n$ if necessary (leaving b_1 and b_2 alone) so that $S(b_3, b_4) \neq 0$. This reordering can be done with a matrix of transition with determinant 1. Let

$$c_i = b_i \text{ for } i = 1, \dots, 4 \text{ and } c_i = b_i - \frac{S(b_i, b_4)}{S(b_3, b_4)}\ b_3 + \frac{S(b_i, b_3)}{S(b_3, b_4)}\ b_4 \text{ for } i = 5, \dots, n.$$

Once again, the matrix of transition has determinant 1. This time the matrix of S in the ordered basis $c_1, \dots, c_n$ has all zeros in the first *four* rows and the first *four* columns except for the first *two* 2×2 diagonal blocks, whose entries are those of the matrices

$$\begin{pmatrix} 0 & S(c_1, c_2) \\ -S(c_1, c_2) & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & S(c_3, c_4) \\ -S(c_3, c_4) & 0 \end{pmatrix}.$$

We carry on in this fashion yielding an ordered basis for V for which the matrix of G is skew-diagonal. The product of all these matrices of transition has determinant 1. If $\mathbf{p}$ is the end-product ordered basis obtained after r iterations of this process

$$S = \sum_{i=1}^{r} S(p_{2i-1}, p_{2i})\ p^{2i-1} \wedge p^{2i}.$$

An interesting and immediate consequence of this representation result is that a 2-form S on a space of dimension $2n$ is nondegenerate (and therefore a symplectic form) if and only if $\underbrace{S \wedge \cdots \wedge S}_{n \text{ factors}} \neq 0$.

An n-fold wedge product, as seen in the last line, is sometimes denoted $\mathbf{S}^{\wedge \mathbf{n}}$. This particular $2n$-form on the $2n$-dimensional V, if nonzero, is called the **volume element generated by the symplectic form S**.

24. Cartesian Tensors

We presume in this section that our vector space V is endowed with a preferred and agreed-upon positive definite inner product, yielding preferred bases—the orthonormal bases—and the orthogonal matrices of transition which connect them. The matrices of transition between these orthonormal bases have determinant ± 1, so these bases form two groups called **orientations.** Two bases have the same orientation if the matrices of transition between them have determinant 1.

We suppose, first, that a multilinear function is defined or proposed whose coordinates are given for orthonormal bases. If these coordinates have tensor character *when transforming from one of these orthonormal bases to another* the multilinear function defined in each orthonormal basis is said to have **Cartesian character** and to define a **Cartesian tensor.**

If the coordinates are only intended to be calculated in, and transformed among, orthonormal coordinate systems of a pre-determined orientation, tensor character restricted to these bases is called **direct Cartesian character** and we are said to have defined a **direct Cartesian tensor.**

Of course, ordinary tensors are both Cartesian and direct Cartesian tensors. Cartesian tensors are direct Cartesian tensors.

Roughly speaking, there are *more* coordinate formulae in a basis that yield Cartesian tensors and direct Cartesian tensors than coordinate formulae that have (unrestricted) tensor character. Our meaning here is clarified in the next paragraph.

The coordinates of Cartesian or direct Cartesian tensors given by coordinate formulae can certainly be consistently transformed by the ordinary rules for change of basis to bases other than those intended for that type of tensor. Often the original coordinate formulae make sense and can be calculated in these more general bases too. But there is no guarantee that these transformed coordinates will match those calculated directly, and in common examples they do not.

In general, when the vocabulary "Cartesian tensor" is used and coordinates represented, every basis you see will be presumed to be orthonormal unless explicit mention is made to the contrary. If direct Cartesian tensors are discussed, all these orthonormal bases will be presumed to have the given orientation.

With that assumption, there are important simplifications in some of the standard calculations involving tensors. These simplifications can be so significant that some folks will never use anything *but* Cartesian tensors if a natural inner product can be found. However these simplifications can be misleading, concealing interesting phenomenon behind the symmetry of these bases.

For instance, raising or lowering an index using the inner product has no effect on the coordinates of a Cartesian tensor, so coordinates $T^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(\mathbf{a})$ in some treatments are all written as $T_{j_1,\ldots,j_{r+s}}(\mathbf{a})$ and only the total order rather than covariant or contravariant order is significant. It is then natural to modify the Einstein summation notation to indicate summation on *any* doubly-repeated index.

This conflation of 2^{r+s} different tensors is handy but rather confusing if you ever decide to leave the orthonormal realm.

As we remarked above, sometimes formulae for coordinates in a basis with Cartesian or direct Cartesian character are intended to be taken literally in any basis, and in this wider setting might not have (unrestricted) tensor character.

Even or odd relative tensors with nontrivial weighting function provide common examples of this. Relative tensors on an orientation class of orthonormal bases are direct Cartesian tensors. Even relative tensors of any weight are Cartesian tensors in orthonormal bases. Relative tensors and the "background" tensor they agree with in these orthonormal bases are very different objects conceptually and the fact that their coordinates are numerically identical in these special bases might be convenient but cannot be counted as illuminating.

We give two examples of Cartesian tensors.

Define for basis $\mathbf{a}$ the function $T\colon V\times V \to \mathbf{R}$ given by $T(v,w) = \sum_{i=1}^{n} v^i(\mathbf{a})w^i(\mathbf{a})$ when $v = v^i(\mathbf{a})\, a_i$ and $w = w^i(\mathbf{a})\, a_i$. In the paragraph below we check that this formula (obviously multilinear) is invariant when the matrix of transition $\mathcal{B}$ to a new basis $\mathbf{b}$ is orthogonal, using the fact that the inverse of $\mathcal{B}$ is its transpose.

We have $v = v^i(\mathbf{a})\, a_i = v^i(\mathbf{a})\, \mathcal{B}^k_i\, b_k$ and $w = w^i(\mathbf{a})\, a_i = w^i(\mathbf{a})\, \mathcal{B}^k_i\, b_k$ so the formula for T applied to the coordinates of these vectors in basis $\mathbf{b}$ is

$$\sum_{k=1}^{n} v^i(\mathbf{a})\mathcal{B}^k_i\, w^j(\mathbf{a})\mathcal{B}^k_j = \sum_{k=1}^{n} v^i(\mathbf{a})w^j(\mathbf{a})\mathcal{B}^k_i\mathcal{B}^k_j = \sum_{i=1}^{n} v^i(\mathbf{a})\, w^j(\mathbf{a})\, \delta^i_j = \sum_{i=1}^{n} v^i(\mathbf{a})\, w^i(\mathbf{a}).$$

Considering the same problem from another (nicer) point of view, the numbers

$$\delta^i{}_j = \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j \end{cases}$$

define the coordinates of a mixed tensor and this tensor has the same coordinates in any basis, orthonormal or not.

Given any inner product G, we saw earlier that $\delta_{i\,j}$ are the coordinates of G and $\delta^{i\,j}$ are the coordinates of $G^\star$.

But since G is positive definite, $\delta^i{}_j$, $\delta^{i\,j}$, $\delta_{i\,j}$ and $\delta_i{}^j$ are all the same in any orthonormal basis, and so they are (all of them) coordinates of Cartesian tensors, and usually all represented using common coordinates $\delta_{i\,j}$. The "dot product" Cartesian tensor we worked with above is $\delta_{i\,j}\, a^i \otimes a^j$.

As a second example, consider the function

$$\epsilon_{i_1\,\ldots\,i_n} = \begin{cases} -1, & \text{if the ordered list } \; i_1,\ldots,i_n \; \text{ is an odd permutation} \\ & \text{of the integers } 1,\ldots,n; \\ 1, & \text{if the ordered list } \; i_1,\ldots,i_n \; \text{ is an even permutation} \\ & \text{of the integers } 1,\ldots,n; \\ 0, & \text{otherwise.} \end{cases}$$

This is called the **permutation symbol, or Levi-Civita symbol, of order n** .

These are the coordinates of a *direct* Cartesian tensor, another example of a tensor with constant coordinates, though in this case to retain "tensor character" we need to restrict the bases even more than for the Kronecker delta. Here we must look only at orthonormal bases of a fixed orientation.

If $v^1,\ldots,v^n$ are any members of V^*, a quick examination shows that

$$v^1 \wedge \cdots \wedge v^n \;=\; \epsilon_{i_1\,\ldots\,i_n}\, v^{i_1} \otimes \cdots \otimes v^{i_n}.$$

If $\mathbf{a}$ and $\mathbf{b}$ are any bases, we saw that if σ is any n-form, the coefficients $\sigma(\mathbf{a})$ and $\sigma(\mathbf{b})$ defined by

$$\sigma \;=\; \sigma(\mathbf{a})\; a^1 \wedge \cdots \wedge a^n \;=\; \sigma(\mathbf{b})\; b^1 \wedge \cdots \wedge b^n$$

are related by

$$\sigma(\mathbf{b}) = (\det(\mathcal{B}))^{-1}\,\sigma(\mathbf{a}).$$

That means three things. First, if $\mathbf{a}$ and $\mathbf{b}$ are both orthonormal bases of the predetermined orientation, then $\det(\mathcal{B}) = 1$ so the fixed numbers $\epsilon_{i_1 \ldots i_n}$ satisfy

$$\epsilon_{i_1 \ldots i_n}\, a^{i_1} \otimes \cdots \otimes a^{i_n} \;=\; \epsilon_{i_1 \ldots i_n}\, b^{i_1} \otimes \cdots \otimes b^{i_n}$$

as suggested above: the numbers $\epsilon_{i_1 \ldots i_n}$ are the coordinates of a covariant direct cartesian tensor.

Second, we see that these numbers are the coordinates of a covariant *odd relative Cartesian tensor with weighting function sign*$(\mathcal{B})$: that is, if you switch to an orthonormal basis of the "other" orientation you multiply by the sign of the determinant of the matrix of transition to the new basis.

Finally, we see that when transferring to an arbitrary basis these constant coefficients require the weighting function $(\det(\mathcal{B}))^{-1}$ in the new basis, so these constant coefficients define an *odd covariant relative tensor of weight* -1.

I have also seen it stated that these same coefficients define both an odd covariant relative tensor of weight -1 and an odd contravariant tensor of weight 1. That is because in these sources there is no notational distinction on the coefficients between "raised or lowered index" versions of a tensor, and what they mean is that when $\mathbf{a}$ is an orthonormal basis of the predetermined orientation, the n-vector

$$\epsilon^{i_1 \ldots i_n}\, a_{i_1} \otimes \cdots \otimes a_{i_n} \;=\; a_1 \wedge \cdots \wedge a_n$$

transforms to arbitrary bases with weighting function of weight 1.

This kind of thing is fine as long as you stay within the orthonormal bases, but when you explicitly intend to leave this realm, as you must if the statement is to have nontrivial interpretation, it is a **bad idea.**

If you want to consider the Cartesian tensor $\delta_{i\,j}$ or direct Cartesian tensor $\epsilon_{i_1,\ldots,i_k}$, or any such tensors whose coordinate formulae make sense in bases more general than their intended bases of application, you have two options.

You can insist on retaining the coordinate formulae in a more general base, in which case you may lose tensor character, and might not even have a relative tensor.

Alternatively, you can abandon the coordinate formulae outside their intended bases of application and allow coordinates to transform when moving to these bases by the usual coordinate change rules, thereby retaining tensor character.

25. Volume Elements and Cross Product

In this section we explore initial versions of ideas we explore in more depth later.

A **volume element on V** is simply a nonzero member of $\Lambda_n(V)$. This vocabulary is usually employed (obviously) when one intends to calculate volumes. An ordered basis $\mathbf{a}$ of V can be used to form a parallelogram or parallelepiped $\mathbf{Par}_{\mathbf{a_1},\ldots,\mathbf{a_n}}$ in V consisting of the following set of points:

$$Par_{a_1,\ldots,a_n} \text{ is formed from all } c^i a_i \text{ where each } c^i \text{ satisfies } 0 \le c^i \le 1.$$

We may decide that we want to regard the volume of this parallelepiped as a standard against which the volumes of other parallelepipeds are measured. In that case we would pick the volume element $\mathbf{Vol_a}$:

$$Vol_{\mathbf{a}} = a^1 \wedge \cdots \wedge a^n$$
$Vol_{\mathbf{a}}$ assigns value 1 to $(a_1, \ldots, a_n)$, interpreted as the volume of $Par_{a_1,\ldots,a_n}$.

More generally, for any ordered list of vectors $H_1, \ldots, H_n$ in V we interpret $|Vol_{\mathbf{a}}(H_1, \ldots, H_n)|$ to be the volume of the parallelepiped with edges $H_1, \ldots, H_n$.

This might be consistent with what you already know or think about volumes in light of the following facts:

- A sends each a_i to the corners of the unit cube in $\mathbb{R}^n$ which, barring evidence to the contrary, would seem to have volume 1.
- More generally, A sends each point in $Par_{a_1,\ldots,a_n} \subset V$ to the corresponding point in $Par_{e_1,\ldots,e_n} \subset \mathbb{R}^n$.
- Still more generally, if $H_1, \ldots, H_n$ is any ordered list of vectors in V then A sends each point in $Par_{H_1,\ldots,H_n}$ to $Par_{A^i(H_1)e_i,\ldots,A^i(H_n)e_i}$ which is a parallelepiped with edges $A^i(H_1)e_i, \ldots, A^i(H_n)e_i$.
- If H_1 is replaced by $2H_1$ the new parallelepiped *ought* to have twice the volume. Linearity of $Vol_{\mathbf{a}}$ in each coordinate reflects this.
- You might have calculated in dimensions 1, 2 and 3 that a parallelepiped with edges $A^i(H_1)e_i, \ldots, A^i(H_n)e_i$ has length, area or volume $|\det(H(\mathbf{a}))|$ and have reason to believe that this formula for volume makes sense in higher dimensions too. Justification could come by filling out larger or smaller shapes with translates of "rational multiples" of a standard cube and adding to get total volume: integration.
- We saw on page 41 in Section 15 that $|\det(H(\mathbf{a}))| = |Vol_{\mathbf{a}}(H_1, \ldots, H_n)|$ making it seem consistent to assign this number to the volume of $Par_{H_1,\ldots,H_n}$.

However you decide, once you come to the conclusion that $Par_{a_1,\ldots,a_n}$ should be your fiduciary volume standard and that $Vol_{\mathbf{a}}$ should be the volume-measurement tool it is easy to find $Vol_{\mathbf{a}}$ from any nonzero $\sigma \in \Lambda_n(V)$:

$$Vol_{\mathbf{a}} \; = \; \frac{1}{\sigma(a_1, \ldots, a_n)}\,\sigma \; = \; a^1 \wedge \cdots \wedge a^n.$$

But how would one come to that conclusion in the first place?

First, it might have arisen as the volume element $S^{\wedge n}$ for a symplectic form S, as outlined on page 62. This form can be used to decide an orientation for bases and has other physical significance related to the Hamiltonian formulation of mechanics on the even dimensional phase space, which we cannot explore here. **Contact forms** play a similar role on the extended phase space, which includes a time parameter and is therefore odd dimensional.

Second, we might have in hand a fixed inner product G. We saw that we can create an orthonormal ordered basis $\mathbf{a}$ for any inner product. This is where the volume element comes from in many situations: as $Vol_{\mathbf{a}}$ for orthonormal $\mathbf{a}$ with respect to inner product G.

Of course this is not *quite* enough to pin down a volume element. We saw in Section 23 that if $\mathbf{b}$ is another orthonormal ordered basis then $Vol_{\mathbf{a}}$ and $Vol_{\mathbf{b}}$ are different by a factor of the determinant of the matrix of transition between bases, which is ± 1. So to identify a unique volume element, we must also decide on one of the two orientations for our orthonormal ordered bases. Sometimes we are only interested in $|Vol_{\mathbf{a}}|$ rather than the signed or oriented volume provided by $Vol_{\mathbf{a}}$, and the issue is irrelevant. In other situations the distinction is crucial.

When given an inner product G and an orthonormal basis $\mathbf{a}$ of specified orientation the volume element $a^1 \wedge \cdots \wedge a^n$ can be written, in light of the formulae on pages 41 and 56, in terms of another basis $\mathbf{b}$ as

$$\begin{aligned} Vol_{\mathbf{a}} &= a^1 \wedge \cdots \wedge a^n = \frac{1}{det(\mathcal{B})}\, b^1 \wedge \cdots \wedge b^n \\ &= sign(\mathcal{B}) \sqrt{\frac{\det(\mathbf{G_b})}{\det(\mathbf{G_a})}}\, b^1 \wedge \cdots \wedge b^n \;=\; sign(\mathcal{B}) \sqrt{|\det(\mathbf{G_b})|}\, b^1 \wedge \cdots \wedge b^n \end{aligned}$$

where $\mathcal{B}$ is the matrix of transition to the new basis.

So that is the picture *if you are given an inner product on V from which you glean enough geometric information to understand volume in V.*

But this merely pushes our conundrum back one level. It remains: from which forehead springs this symplectic form or this metric tensor?

In an application, say from physics, a metric tensor can be thought of as the source itself, the most basic reflection and encapsulation of simple local physics, from which grows the idea of global distances measured by numerous little yardsticks trapped on locally flat pieces of a curvy universe. In other words, some theories about the world generate (locally) an inner product when phrased in geometrical language. These theories are satisfying to us because they tap into and salvage at least part of our natural internal models of how the external world should be.

Some of these theories *also* have the added advantage that they seem to be *true*, in the sense that repeated tests of these theories have so far turned up, uniformly, repeated confirmations.

When the metric tensor at each point of this universe is positive definite it is called **Riemannian** and induces the usual idea of distance on this curvy surface, a **Riemannian metric** on what is called a **Riemannian manifold.** Small pieces of a Riemannian manifold exhibit the ordinary **Euclidean geometry** of $\mathbb{R}^n$.

When the metric tensor defined at each point has **Lorentz signature**—that is, when diagonalized its matrix has one negative entry and three positive entries—we find ourselves on a **Lorentzian manifold**, the **space-time** of **Einstein's general relativity**. Small pieces of a Lorentzian manifold exhibit the **Minkowski geometry of special relativity.**

In that world with inner product G, some vectors satisfy $G(v,v) > 0$ and a v like that is said to indicate a "space-like" displacement. No massive or massless particle can move in this way. If $G(v,v) = 0$, v can indicate the a displacement of a massless particle such as light, and lies on the "light-cone." When $G(v,v) < 0$ the "time" coordinate of v is large enough in proportion to its "space" coordinates that the vector could indicate the displacement of a massive particle. The vector is called "time-like."

There is another construction one sees when given specific inner product G and orientation to yield $Vol_{\mathbf{a}}$ for orthonormal $\mathbf{a}$. This is a means of creating a vector in V, called cross product, from a given ordered list $H_1, \dots, H_{n-1}$ of members of V.

The cross product is defined in two steps as follows.

If you evaluate $Vol_{\mathbf{a}}$ using $H_1, \dots, H_{n-1}$ in its first $n-1$ slots, there is one slot remaining. Thus $Vol_{\mathbf{a}}(H_1, \dots, H_{n-1}, \_\,)$ is a member T of V^*, a different member of V^* for each choice of these various H_i. This is the first step.

For the second step, raise the index of T using $G^\star$ to yield a member of V, the **cross product**

$$T(\mathbf{a})^i\, a_i = T(\mathbf{a})_k\, G^{k,i}(\mathbf{a})\, a_i \in V.$$

The raised index version of the tensor T is often denoted $\mathbf{H_1} \times \cdots \times \mathbf{H_{n-1}}$. The case of $n = 3$ and $\mathbf{G_a} = \mathcal{I}$ is most familiar, yielding the usual cross product of pairs of vectors.

The covector (unraised) T has some interesting properties.

For example, it is linear when thought of as a function from V to V^*, where the member of V replaces H_i for any fixed i. Also, if H_i and H_j are switched for distinct subscripts, the new T becomes the negative of the old T. Because of this, if there is any linear relation among the H_i then $T = 0$. On the other hand, if the H_i are linearly independent then they can be extended to an ordered basis by adding one vector H_n and then $T(H_n) \neq 0$ which means $T \neq 0$. **Conclusion: $\mathbf{T = 0}$ exactly when $\mathbf{H_1, \dots, H_{n-1}}$ is a dependent list.**

The tensor T is calculated on any vector H_n as $\det(H(\mathbf{a}))$ and as we saw on page 40 this determinant can be expanded around the last column as

$$T(H_n) = \det(H(\mathbf{a})) = \sum_{i=1}^{n} (-1)^{i+n}\, A^i(H_n)\, \det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1} \right).$$

and so T is

$$T(\mathbf{a})_i\, a^i = \sum_{i=1}^{n} (-1)^{i+n}\, \det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1} \right)\, a^i \in V^*.$$

T is nonzero exactly when $H_1, \dots, H_{n-1}$ constitutes an independent collection of vectors.

An explicit formula for the raised-index version of T (the version in V) is

$$T(\mathbf{a})^i\, a_i = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} (-1)^{k+n}\, \det\left(H(\mathbf{a})^{1,\dots,k-1,k+1,\dots,n}_{1,\dots,n-1} \right)\, G^{k,i}(\mathbf{a}) \right) a_i.$$

If $\mathbf{a}$ has been chosen so that $\mathbf{G_a}$ is diagonal, this formula simplifies to

$$T(\mathbf{a})^i\, a_i = \left(\sum_{i=1}^{n} (-1)^{i+n}\, \det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1} \right)\, G^{i,i}(\mathbf{a}) \right) a_i.$$

Now suppose we let $H_n = H_1 \times \cdots \times H_{n-1}$ and calculate $\det(H(\mathbf{a}))$.

$$\begin{aligned}
\det(H(\mathbf{a}))& \\
&= \sum_{i=1}^{n} (-1)^{i+n}\, A^i(H_n)\, \det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1}\right) \\
&= \sum_{i=1}^{n} (-1)^{i+n}\, \det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1}\right) \\
&\qquad \cdot \left(\sum_{k=1}^{n} (-1)^{k+n}\, \det\left(H(\mathbf{a})^{1,\dots,k-1,k+1,\dots,n}_{1,\dots,n-1}\right)\, G^{k,i}(\mathbf{a})\right)
\end{aligned}$$

$$= \sum_{k=1}^{n}\sum_{i=1}^{n} (-1)^{i+n+k+n}\, G^{k,i}(\mathbf{a})\, \det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1}\right) \det\left(H(\mathbf{a})^{1,\dots,k-1,k+1,\dots,n}_{1,\dots,n-1}\right).$$

If $\mathbf{a}$ has been chosen so that $\mathbf{G_a}$ is diagonal the last line becomes

$$\det(H(\mathbf{a})) = \sum_{i=1}^{n} G^{i,i}(\mathbf{a})\, \left(\det\left(H(\mathbf{a})^{1,\dots,i-1,i+1,\dots,n}_{1,\dots,n-1}\right)\right)^2.$$

If $H_1, \dots, H_{n-1}$ constitutes an independent collection of vectors and the $G^{i,i}(\mathbf{a})$ all have the same sign, this is nonzero so $H_1, \dots, H_{n-1}, H_1 \times \cdots \times H_{n-1}$ constitutes an ordered basis of V. If the $G^{i,i}(\mathbf{a})$ are all positive, that ordered basis has the same orientation as $\mathbf{a}$.

Considering a different issue, if σ is any volume element and we change ordered basis from $\mathbf{a}$ to $\mathbf{b}$ then, as we saw earlier, the (single) coordinate changes according to $\sigma(\mathbf{b}) = \det(\mathcal{A})\sigma(\mathbf{a})$. This implies

$$a^1 \wedge \cdots \wedge a^n = \det(\mathcal{A})\ b^1 \wedge \cdots \wedge b^n.$$

So if $\det(\mathcal{A}) = 1$ then for any H_n

$$T(H_n) = \det(H(\mathbf{a})) = \det(\mathcal{A})\ \det(H(\mathbf{b})) = \det(H(\mathbf{b})).$$

Going back to the issue of the formation of cross product, we conclude that step one in the process of creating the coordinates of the cross product in an ordered basis is the same, up to numerical multiple, from basis to basis. When the matrix of transition has determinant 1, the volume element is the same for these two bases, so step one of the process has the same form regardless of other aspects of the new basis. Step two, however, involves the matrix $\overline{\mathbf{G_b}}$ which changes as $\mathcal{B}\ \overline{\mathbf{G_a}}\ \mathcal{B}^t$.

26. An Isomorphism Between $\Lambda^r(V)$ and $\Lambda_{n-r}(V)$

We have seen that the dimension of the space $\Lambda_r(V)$ is $\frac{n!}{r!\,(n-r)!}$ for any r with $0 \le r \le n$, which is the same as the dimensions of

$$\Lambda^r(V) = \Lambda_r(V^*) \quad \text{and} \quad \Lambda_{n-r}(V) \quad \text{and} \quad \Lambda^{n-r}(V) = \Lambda_{n-r}(V^*).$$

So they are all isomorphic as vector spaces.

Our goal in this section is to create a specific isomorphism between $\Lambda^r(V)$ and $\Lambda_{n-r}(V)$, and also between $\Lambda_r(V)$ and $\Lambda^{n-r}(V)$, and we want this construction to be independent of basis in some sense.

It is required that a volume element σ, a member of $\mathbf{\Lambda_n(V)}$, be specified for the construction to proceed.

If $\mathbf{a}$ is any basis of V with dual basis $\mathbf{a}^*$ the n-form $a^1 \wedge \cdots \wedge a^n$ is a multiple of σ. We presume that basis $\mathbf{a}$ has been chosen so that $a^1 \wedge \cdots \wedge a^n = \sigma$.

For a choice of $a_{i_1}, \ldots, a_{i_r}$ we define

$$\boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \otimes \cdots \otimes a_{i_r}) = \sigma(a_{i_1}, \ldots, a_{i_r}, \cdot, \ldots, \cdot) \colon V^{n-r} \to \mathbb{R}.$$

It has various interesting properties. It is multilinear and alternating, so it is a member of $\Lambda_{n-r}(V)$. Also, if P is any permutation of $\{1, \ldots, r\}$ then

$$\boldsymbol{Formit}_{\mathbf{a}}(a_{i_{P(1)}} \otimes \cdots \otimes a_{P(r)}) = sgn(P)\, \boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \otimes \cdots \otimes a_{i_r}).$$

So if there are any repeated terms among the a_{i_k} then $\boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \otimes \cdots \otimes a_{i_r}) = 0$.

Having defined $\boldsymbol{Formit}_{\mathbf{a}}$ on a basis of $\mathcal{T}^r_0(V)$ consisting of tensors of the form $a_{i_1} \otimes \cdots \otimes a_{i_r}$, we can extend the definition of $\boldsymbol{Formit}_{\mathbf{a}}$ by linearity to all of $\mathcal{T}^r_0(V)$. Any linear combination of members of $\Lambda_{n-r}(V)$ is also in $\Lambda_{n-r}(V)$, so the range of $\boldsymbol{Formit}_{\mathbf{a}}$ is in $\Lambda_{n-r}(V)$.

We note (evaluate on $(n-r)$-tuples $(a_{j_1}, \ldots a_{j_{n-r}})$ in V^{n-r}) that

$$\boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \otimes \cdots \otimes a_{i_r}) = \boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \wedge \cdots \wedge a_{i_r}) = sgn(P)\, \boldsymbol{Formit}_{\mathbf{a}}(a_{i_{P(1)}} \wedge \cdots \wedge a_{i_{P(r)}})$$

for any permutation P of $\{1, \ldots, r\}$ and any choice of $i_1, \ldots, i_r$.

We will only be interested here in $\boldsymbol{Formit}_{\mathbf{a}}$ evaluated on members of $\Lambda^r(V)$ and we officially restrict, therefore, to this domain:

$$\boldsymbol{Formit}_{\mathbf{a}} \colon \Lambda^r(V) \to \Lambda_{n-r}(V).$$

It is pretty easy to see that if $a_{i_1}, \ldots, a_{i_r}$ is a listing of basis vectors in increasing order and $a_{i_{r+1}}, \ldots, a_{i_n}$ is a listing of the basis vectors left unused among the $a_{i_1}, \ldots, a_{i_r}$ then

$$\boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \wedge \cdots \wedge a_{i_r})(a_{i_{r+1}}, \ldots, a_{i_n}) = \sigma(a_{i_1}, \ldots, a_{i_r}, a_{i_{r+1}}, \ldots, a_{i_n}) = \pm 1$$

while if any of the vectors $a_{i_{r+1}}, \ldots, a_{i_n}$ repeat a vector on the list $a_{i_1}, \ldots, a_{i_r}$ then

$$\boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \wedge \cdots \wedge a_{i_r})(a_{i_{r+1}}, \ldots, a_{i_n}) = \sigma(a_{i_1}, \ldots, a_{i_r}, a_{i_{r+1}}, \ldots, a_{i_n}) = 0.$$

We deduce from these facts that $\boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \wedge \cdots \wedge a_{i_r}) = \pm\, a^{i_{r+1}} \wedge \cdots \wedge a^{i_n}$.

Recall that for each increasing index choice $i_1, \ldots, i_r$ drawn from $\{1, \ldots, n\}$ there is exactly one shuffle permutation P in $\mathcal{S}_{r,n} \subset \mathcal{P}_n$ for which $P(k) = i_k$ for $k = 1, \ldots, r$. For this shuffle, we also have (by definition of shuffle) $P(r+1) < P(r+2) < \cdots < P(n)$. There are $n!$ members of $\mathcal{P}_n$ but only $\frac{n!}{r!(n-r)!}$ members of $\mathcal{S}_{r,n}$.

So for our shuffle permutation P we can see that

$$\boldsymbol{Formit}_{\mathbf{a}}(a_{P(1)} \wedge \cdots \wedge a_{P(r)}) = sgn(P)\, a^{P(r+1)} \wedge \cdots \wedge a^{P(n)}.$$

Suppose $T \in \Lambda^r(V)$ with

$$\begin{aligned} T &= T^{i_1,\dots,i_r}(\mathbf{a})\, a_{i_1} \wedge \cdots \wedge a_{i_r} \quad \text{(Sum on increasing indices.)} \\ &= \sum_{P \in \mathcal{S}_{r,n}} T^{P(1),\dots,P(r)}(\mathbf{a})\, a_{P(1)} \wedge \cdots \wedge a_{P(r)}, \end{aligned}$$

where the final sum above is only over the shuffle permutations. Then

$$\begin{aligned} \boldsymbol{Formit}_{\mathbf{a}}(T) &= T^{i_1,\dots,i_r}(\mathbf{a})\, \boldsymbol{Formit}_{\mathbf{a}}(a_{i_1} \wedge \cdots \wedge a_{i_r}) \quad \text{(Sum on increasing indices.)} \\ &= \sum_{P \in \mathcal{S}_{r,n}} sgn(P)\, T^{P(1),\dots,P(r)}(\mathbf{a})\ a^{P(r+1)} \wedge \cdots \wedge a^{P(n)} \end{aligned}$$

This map is obviously onto $\Lambda_{n-r}(V)$ and in view of the equality of dimension provides an isomorphism from $\Lambda^r(V)$ onto $\Lambda_{n-r}(V)$.

It is hardly pellucid that this isomorphism does not depend on specific choice of basis $\mathbf{a}$.

We let $\mathbf{b}$ be a second basis of V with $\sigma = b^1 \wedge \cdots \wedge b^n$ and form isomorphism $\boldsymbol{Formit}_{\mathbf{b}}$.

Suppose $m_1, \dots m_r$ is an increasing sequence and P is the shuffle permutation $P \in \mathcal{S}_{r,n}$ with $m_j = P(j)$ for $j = 1, \dots, r$.

By the change of basis formula for wedge product on page 41 we have

$$\begin{aligned} \boldsymbol{Formit}_{\mathbf{a}}(b_{m_1} \wedge \cdots \wedge b_{m_r}) &= \boldsymbol{Formit}_{\mathbf{a}}\left(\det\left(\mathcal{A}^{j_1,\dots,j_r}_{m_1,\dots,m_r}\right)\, a_{j_1} \wedge \cdots \wedge a_{j_r}\right) \\ &\quad \text{(Sum on the line above over increasing indices only.)} \\ &= \sum_{Q \in \mathcal{S}_{r,n}} sgn(Q)\ \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{m_1,\dots,m_r}\right)\ a^{Q(r+1)} \wedge \cdots \wedge a^{Q(n)} \\ &= \sum_{Q \in \mathcal{S}_{r,n}} sgn(Q)\ \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{P(1),\dots,P(r)}\right) \\ &\qquad\qquad \cdot \det\left(\mathcal{A}^{Q(r+1),\dots,Q(n)}_{k_1,\dots,k_{n-r}}\right)\, b^{k_1} \wedge \cdots \wedge b^{k_{n-r}}. \\ &\quad \text{(Sum on the line above over both shuffles and increasing indices.)} \end{aligned}$$

We would like to deduce that the n-r-form in the last two lines above is

$$\boldsymbol{Formit}_{\mathbf{b}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)}) \ = \ sgn(P)\, b^{P(r+1)} \wedge \cdots \wedge b^{P(n)}$$

and we will prove this by showing that $\boldsymbol{Formit}_{\mathbf{a}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})$ and $\boldsymbol{Formit}_{\mathbf{b}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})$ agree when evaluated on any $(n-r)$-tuple of vectors of the form $(b_{i_1}, \dots, b_{i_{n-r}})$ for integers $i_1, \dots, i_{n-r}$ between 1 and n.

First, if there is any duplication among these integers then alternating $\boldsymbol{Formit}_{\mathbf{a}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})$ and $b^{P(r+1)} \wedge \cdots \wedge b^{P(n)}$ must both be zero when evaluated at $(b_{i_1}, \dots, b_{i_{n-r}})$. So we may presume that all the $i_1, \dots, i_{n-r}$ are distinct.

Further, we can re-order the vectors in the n-r-tuple and both $\boldsymbol{Formit}_{\mathbf{a}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})(b_{i_1}, \dots, b_{i_{n-r}})$ and $\left(b^{P(r+1)} \wedge \cdots \wedge b^{P(n)}\right)(b_{i_1}, \dots, b_{i_{n-r}})$ will be multiplied by the signum of the reordering permutation. So we may presume, to confirm equality, that $i_1, \dots, i_{n-r}$ is an increasing sequence.

If $i_1, \dots, i_{n-r}$ contains any of the integers $P(1), \dots, P(r)$ then

$$\boldsymbol{Formit}_{\mathbf{b}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})(b_{i_1}, \dots, b_{i_{n-r}}) = \sigma(b_{P(1)}, \dots, b_{P(r)}, b_{i_1}, \dots, b_{i_{n-r}}) = 0.$$

Also in this case

$$\boldsymbol{Formit}_{\mathbf{a}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})(b_{i_1}, \dots, b_{i_{n-r}}) = \sum_{Q \in \mathcal{S}_{r,n}} sgn(Q)\ \det\left(\mathcal{A}^{Q(1),\dots,Q(r)}_{P(1),\dots,P(r)}\right)\ \det\left(\mathcal{A}^{Q(r+1),\dots,Q(n)}_{i_1,\dots,i_{n-r}}\right).$$

By the generalization of the Laplace expansion formula found on page 42, this is the determinant of a matrix related to $\mathcal{A}$ with columns reordered and with (at least) two columns duplicated. It is, therefore, also zero.

So it remains to show equality when $i_j = P(r+j)$ for $j = 1, \dots, n-r$. In that case

$$\boldsymbol{Formit}_{\mathbf{a}}(b_{P(1)} \wedge \cdots \wedge b_{P(r)})(b_{P(r+1)}, \dots, b_{P(n)}) = sgn(P)\ \det(\mathcal{A}) = sgn(P)$$

while we also have

$$\left(sgn(P)\, b^{P(r+1)} \wedge \cdots \wedge b^{P(n)}\right)(b_{P(r+1)}, \dots, b_{P(n)}) = sgn(P).$$

In other words, $\boldsymbol{Formit}_{\mathbf{a}}$ and $\boldsymbol{Formit}_{\mathbf{b}}$ agree when evaluated on a basis of V^{n-k} and so are equal. We no longer need to (and will not) indicate choice of basis as subscript on the function $\boldsymbol{Formit}_{\mathbf{a}}$, using instead $\boldsymbol{Formit}_{\sigma}$ with reference to the specified volume element σ. By the way, had the other orientation been given by choosing $-\sigma$ as volume element, we find that $\boldsymbol{Formit}_{\sigma} = -\boldsymbol{Formit}_{-\sigma}$.

We extend $\boldsymbol{Formit}_{\sigma}$ slightly and define

$\boldsymbol{Formit}_{\sigma} \colon \Lambda^0(V) \to \Lambda_n(V)$ to be the map sending a constant k to $k\sigma$.

We clarify the definition of $\boldsymbol{Formit}_{\sigma} \colon \Lambda^n(V) \to \Lambda_0(V)$. It is the map given by

$\boldsymbol{Formit}_{\sigma}\,(k\,\sigma) = k$ for constant k.

Recapitulating, we have defined isomorphisms

$$\boxed{\boldsymbol{Formit}_{\sigma} \colon \Lambda^r(V) \to \Lambda_{n-r}(V) \qquad \text{for } 0 \le r \le n.}$$

When $0 < r < n$ and for any basis $\mathbf{a}$ for which $\sigma = a^1 \wedge \cdots \wedge a^n$ we have

$$\boxed{\boldsymbol{Formit}_{\sigma}\left(a_{P(1)} \wedge \cdots \wedge a_{P(r)}\right) = sgn(P)\, a^{P(r+1)} \wedge \cdots \wedge a^{P(n)}}$$

for any shuffle permutation $P \in \mathcal{S}_{r,n}$.

More generally, if $\mathbf{b}$ is any basis at all,

$$\boxed{\boldsymbol{Formit}_{\sigma}\left(b_{P(1)} \wedge \cdots \wedge b_{P(r)}\right) = sgn(P)\sigma(b_1, \dots, b_n)\, b^{P(r+1)} \wedge \cdots \wedge b^{P(n)}}$$

for any shuffle permutation $P \in \mathcal{S}_{r,n}$.

Whenever $\mathbf{a}$ and $\mathbf{b}$ are bases for which $\sigma = a^1 \wedge \cdots \wedge a^n = b^1 \wedge \cdots \wedge b^n$ the change of basis formula tells us that the two n-vectors $a_1 \wedge \cdots \wedge a_n$ and $b_1 \wedge \cdots \wedge b_n$ are equal. This n-vector is related to and determined by σ, but is not the "raised index version of σ." There is no inner product here to raise or lower indices.

Using this n-vector we define, by means identical to those used above, isomorphisms

$$\boldsymbol{Vecit}_\sigma \colon \Lambda_r(V) \to \Lambda^{n-r}(V) \qquad \text{for } 0 \le r \le n.$$

When $0 < r < n$ and for basis $\mathbf{a}$ for which $\sigma = a^1 \wedge \cdots \wedge a^n$ we have

$$\boldsymbol{Vecit}_\sigma\left(a^{P(1)} \wedge \cdots \wedge a^{P(r)}\right) = sgn(P)\, a_{P(r+1)} \wedge \cdots \wedge a_{P(n)}$$

for any shuffle permutation $P \in \mathcal{S}_{r,n}$.

More generally, if $\mathbf{b}$ is any basis at all,

$$\boldsymbol{Vecit}_\sigma\left(b^{P(1)} \wedge \cdots \wedge b^{P(r)}\right) = \frac{sgn(P)}{\sigma(b_1, \ldots, b_n)}\, b_{P(r+1)} \wedge \cdots \wedge b_{P(n)}$$

for any shuffle permutation $P \in \mathcal{S}_{r,n}$.

One natural question is how $\boldsymbol{Vecit}_\sigma$ is related to $\boldsymbol{Formit}_\sigma^{-1}$ for each r.

Using basis $\mathbf{a}$ with $\sigma = a^1 \wedge \cdots \wedge a^n$

$$\boldsymbol{Vecit}_\sigma \circ \boldsymbol{Formit}_\sigma(a_1 \wedge \cdots \wedge a_r) = \boldsymbol{Vecit}_\sigma(a^{r+1} \wedge \cdots \wedge a^n) = sgn(Q)\, a_1 \wedge \cdots \wedge a_r$$

where Q is the permutation that converts the list

$$r+1, \ldots, n, 1, \ldots, r$$

to natural order. This is accomplished by $r(n-r)$ switches of consecutive members on the list, and we find that

$$\boldsymbol{Vecit}_\sigma \circ \boldsymbol{Formit}_\sigma(a_1 \wedge \cdots \wedge a_r) = (-1)^{r(n-r)}\, a_1 \wedge \cdots \wedge a_r.$$

To resolve the issue for a generic basis tensor $a_{i_1} \wedge \cdots \wedge a_{i_r}$ we can form a basis $\mathbf{c}$ by permutation of the members of the basis $\mathbf{a}$ and apply this calculation (watching signs if the permutation is odd) in this new basis $\mathbf{c}$. The same formula holds.

$\boldsymbol{Vecit}_\sigma \circ \boldsymbol{Formit}_\sigma$ is $(-1)^{r(n-r)}$ times the identity map.

So if r is even or n is odd this is always the identity map. A minus sign is introduced only in case n is even and r is odd.

27. The Hodge $*$ Operator and Hodge Duality

Our goal in this section is to identify explicit isomorphisms between $\Lambda_r(V)$ and $\Lambda_{n-r}(V)$ and also between $\Lambda^r(V)$ and $\Lambda^{n-r}(V)$ for each r.

In this section we presume given an inner product G on V and an orientation.

From these two specifications, we create the n-form $\sigma = \mathbf{a^1} \wedge \cdots \wedge \mathbf{a^n}$ where a is an orthonormal basis of requisite orientation.

This is the world of direct Cartesian tensors if the inner product happens to be positive definite, though in important cases it is not.

Our inner product has signature (p, q): that is, when diagonalized the matrix of G has p positive and $q = n - p$ negative entries.

For various orders r, we define "raising" isomorphisms $\mathbf{Raise}\colon \Lambda_r(V) \to \Lambda^r(V)$ and "lowering" isomorphisms $\mathbf{Lower}\colon \Lambda^r(V) \to \Lambda_r(V)$ by raising or lowering all indices on each tensor, as described on page 53. These isomorphisms, for a given r, are inverse to each other.

We define isomorphisms for various r values

$$\boldsymbol{Formit}_\sigma \circ \mathbf{Raise}\colon \Lambda_r(V) \to \Lambda_{n-r}(V) \text{ and } \boldsymbol{Vecit}_\sigma \circ \mathbf{Lower}\colon \Lambda^r(V) \to \Lambda^{n-r}(V)$$

which will *all* be denoted $*$ and called "the" **Hodge star operator**.

The $(n - r)$-form $*\tau \in \Lambda_{n-r}(V)$ will be called the **Hodge dual** of $\tau \in \Lambda_r(V)$. The $(n - r)$-vector $*x \in \Lambda^{n-r}(V)$ will be called the **Hodge dual** of $x \in \Lambda^r(V)$.

With our definition of $*$ on $\Lambda_r(V)$ we calculate that $*\,1 = \sigma$ and $*\,\sigma = (-1)^q\,1$.

As this hints, and as was the case with $\boldsymbol{Formit}_\sigma$ and $\boldsymbol{Vecit}_\sigma$, the Hodge operator $*$ is closely related to its inverse.

Our inner product G has signature (p, q), so q is the number of -1 entries on the diagonal of the matrix of $\mathbf{G_a} = \overline{\mathbf{G_a}}$ for orthonormal basis $\mathbf{a}$.

Suppose $0 < r < n$ and P is any shuffle permutation in $\mathcal{S}_{r,n}$ and that exactly k of the vectors $a_{P(1)}, \dots, a_{P(r)}$ satisfy $G\left(a_{P(i)}, a_{P(i)}\right) = -1$. So $q - k$ of the vectors $a_{P(r+1)}, \dots, a_{P(n)}$ satisfy that equation. It follows that

$$\begin{aligned}
&* \circ * \left(a_{P(1)} \wedge \cdots \wedge a_{P(r)}\right)\\
&\quad= \boldsymbol{Vecit}_\sigma \circ \mathbf{Lower} \circ \boldsymbol{Vecit}_\sigma \circ \mathbf{Lower}\left(a_{P(1)} \wedge \cdots \wedge a_{P(r)}\right)\\
&\quad= \boldsymbol{Vecit}_\sigma \circ \mathbf{Lower} \circ \boldsymbol{Vecit}_\sigma \left((-1)^k\, a^{P(1)} \wedge \cdots \wedge a^{P(r)}\right)\\
&\quad= \boldsymbol{Vecit}_\sigma \circ \mathbf{Lower}\left(sgn(P)\,(-1)^k\, a_{P(r+1)} \wedge \cdots \wedge a_{P(n)}\right)\\
&\quad= \boldsymbol{Vecit}_\sigma \left(sgn(P)\,(-1)^{q-k}\,(-1)^k\, a^{P(r+1)} \wedge \cdots \wedge a^{P(n)}\right)\\
&\quad= sgn(P)\,(-1)^{q-k}\,(-1)^k\,(-1)^{r(n-r)}\,sgn(P)\, a_{P(1)} \wedge \cdots \wedge a_{P(r)}\\
&\quad= (-1)^{r(n-r)+q}\, a_{P(1)} \wedge \cdots \wedge a_{P(r)}.
\end{aligned}$$

In other words,

$$* \circ * \text{ is } (-1)^{r(n-r)+q} \text{ times the identity map.}$$

The cases $r = 0$ and $r = n$, not covered above, are calculated directly with the same outcome. Also, the calculation of $* \circ * \left(a^{P(1)} \wedge \cdots \wedge a^{P(r)}\right)$ is very similar, again yielding the same conclusion.

The Hodge $*$ can be used to form an inner product on each $\Lambda^r(V)$ and each $\Lambda_r(V)$. We illustrate the construction on $\Lambda^r(V)$.

For each $x, y \in \Lambda^r(V)$ the n-vector $x \wedge *y$ is a multiple of the standard n-vector $a_1 \wedge \cdots \wedge a_n$. Define

$$< x, y >_r \text{ to be } k \text{ when } x \wedge *y = k\, a_1 \wedge \cdots \wedge a_n$$

where $\mathbf{a}$ is properly oriented and orthonormal.

$< x, y >_r$ is obviously linear in each factor separately.

$(a_{i_1} \wedge \cdots \wedge a_{i_r}) \wedge \left(a_{j_1} \wedge \cdots \wedge a_{j_{n-r}}\right)$ is nonzero exactly when all the subscripts are different. So $(a_{i_1} \wedge \cdots \wedge a_{i_r}) \wedge (* a_{j_1} \wedge \cdots \wedge a_{j_r})$ is nonzero exactly when all the subscripts $i_1, \ldots, i_r$ are different and constitute a reordering of $j_1, \ldots, j_r$. Assuming both are in increasing order, they must be the same indexing. The wedge product is of the form $\left(a_{P(1)} \wedge \cdots \wedge a_{P(r)}\right) \wedge \left(* a_{P(1)} \wedge \cdots \wedge a_{P(r)}\right)$ for $P \in \mathcal{S}_{r,n}$.

For any shuffle $P \in \mathcal{S}_{r,n}$ let

$$M_{\mathbf{a}}(P) = G\left(a_{P(1)}, a_{P(1)}\right) \cdots G\left(a_{P(r)}, a_{P(r)}\right).$$

In other words, $M_{\mathbf{a}}(P)$ is the product of the diagonal entries of the matrix $\mathbf{G_a}$ (all 1 or -1) which correspond to positions $P(j)$ along that diagonal for any $j = 1, \ldots, r$.

We calculate

$$\begin{aligned}
x \wedge *y \;&=\; \left(x^{i_1,\ldots,i_r}(\mathbf{a})\, a_{i_1} \wedge \cdots \wedge a_{i_r}\right) \wedge \left(* \, y^{j_1,\ldots,j_r}(\mathbf{a})\, a_{j_1} \wedge \cdots \wedge a_{j_r}\right) \\
&= \sum_{P \in \mathcal{S}_{r,n}} x^{P(1),\ldots,P(r)}(\mathbf{a})\, y^{P(1),\ldots,P(r)}(\mathbf{a})\, sgn(P)\, M_{\mathbf{a}}(P)\, a_{P(1)} \wedge \cdots \wedge a_{P(n)} \\
&= \left(\sum_{P \in \mathcal{S}_{r,n}} x^{P(1),\ldots,P(r)}(\mathbf{a})\, y^{P(1),\ldots,P(r)}(\mathbf{a})\, M_{\mathbf{a}}(P) \right) a_1 \wedge \cdots \wedge a_n
\end{aligned}$$

which is obviously symmetrical: that is, $< x, y >_r \,=\, < y, x >_r$.

It also follows by inspection of the sum above that we actually have an inner product: the bilinear function is nondegenerate and the standard basis of $\Lambda^r(V)$ for basis $\mathbf{a}$ is an orthonormal basis of $\Lambda^r(V)$ with respect to this inner product.

We note also that

$$\left\langle a_{P(1)} \wedge \cdots \wedge a_{P(r)},\; * \, a_{P(1)} \wedge \cdots \wedge a_{P(r)} \right\rangle_r \;=\; M_{\mathbf{a}}(P).$$

In particular, for any signature (p, q) of G we will have $G \;= \,< \cdot, \cdot >_1$.

28. The Grassmann Algebra

Given a vector space V of dimension $n \geq 2$ we define the **Grassmann algebra** $\mathcal{G}(V)$ for V to be the free sum

$$\boxed{\Lambda_0(V) \oplus \Lambda_1(V) \oplus \Lambda_2(V) \oplus \cdots \oplus \Lambda_n(V) \oplus \cdots}$$

together with wedge product as the multiplication on the algebra, calculated by distribution of the nonzero terms of the formal sums, and where wedge product against a scalar (a member of $\Lambda_0(V)$) is given by ordinary scalar multiplication.

The Grassmann algebra is also called **the exterior algebra,** with reference to the wedge or exterior product.

We note that all summands beyond $\Lambda_n(V)$ consist of the zero tensor only.

This means, first, that members of $\mathcal{G}(V)$ are all formal sums of the form $\theta = \sum_{i=0}^{\infty} \lceil \theta \rceil_i$ where each $\lceil \theta \rceil_i \in \Lambda_i(V)$. There can be at most $n+1$ nonzero summands.

$\lceil \theta \rceil_i$ is called the grade-i part of θ. This representation is unique for each θ: that is, two members of $\mathcal{G}(V)$ are equal if and only if their grade-i parts are equal for $i = 0, \ldots, n$.

Though $\Lambda_r(V)$ is not actually contained in $\mathcal{G}(V)$, an isomorphic copy of each $\Lambda_r(V)$ *is* in $\mathcal{G}(V)$. We do not normally distinguish, unless it is critical for some rare reason, between a member of $\Lambda_r(V)$ and the corresponding member of $\mathcal{G}(V)$ which is 0 except for a grade-r part.

$\mathcal{G}(V)$ is a vector space with scalar multiplication given by ordinary scalar multiplication distributed across the formal sum to each grade, while addition is calculated by adding the parts of corresponding grade. The member of $\mathcal{G}(V)$ which is the zero form at each grade acts as additive identity.

If $\mathbf{a}$ is a basis of V then 1 together with the set of all

$$a^{i_i} \wedge \cdots \wedge a^{i_r}$$

for all r between 1 and n and increasing sequences $i_1, \ldots, i_r$ forms a basis for $\mathcal{G}(V)$, which therefore has dimension

$$1 + n + \cdots + \frac{n!}{r!(n-r)!} + \cdots + n + 1 = 2^n.$$

Looking at individual grades, $\theta \wedge \tau$ for θ, $\tau \in \mathcal{G}(V)$ is given by

$$\lceil \theta \wedge \tau \rceil_i = \sum_{r=0}^{i} \lceil \theta \rceil_r \wedge \lceil \tau \rceil_{i-r}$$

for each i between 0 and n, while $\lceil \theta \wedge \tau \rceil_i$ is the zero i-form for $i > n$.

Sometimes formulae produce reference to a negative grade. Any part or coefficient involved is interpreted to be the appropriate form of zero, just as one would do for a reference to a grade larger than n.

If θ is nonzero only in grade i while τ is nonzero only in grade j then $\theta \wedge \tau$ can only be nonzero in grade $i + j$. By virtue of this property, $\mathcal{G}(V)$ is called a **graded algebra.**

Members of $\mathcal{G}(V)$ are called **multi-forms.**

Replacing V by V^* produces the Grassmann algebra $\mathcal{G}(V^*)$, whose members are called **multi-vectors** and which also forms a graded algebra with wedge product.

If the field we are working with is $\mathbb{R}$, we note that the Grassmann algebra of multi-forms $\mathcal{G}(V)$ is

$$\mathbb{R} \oplus V^* \oplus \Lambda_2(V) \oplus \cdots \oplus \Lambda_n(V) \oplus \{0\} \oplus \cdots$$

while the Grassmann algebra of multi-vectors $\mathcal{G}(V^*)$ is

$$\mathbb{R} \oplus V \oplus \Lambda^2(V) \oplus \cdots \oplus \Lambda^n(V) \oplus \{0\} \oplus \cdots .$$

When a volume element σ is specified, the two families of isomorphisms $\boldsymbol{Formit}_\sigma$ and $\boldsymbol{Vecit}_\sigma$ from Section 26 induce (vector space) isomorphisms (when applied at each grade) between $\mathcal{G}(V)$ and $\mathcal{G}(V^*)$.

Given the additional structure of an inner product G on V and a choice of orientation we can, similarly, use the Hodge $*$ map to induce (vector space) automorphisms on $\mathcal{G}(V)$ and on $\mathcal{G}(V^*)$.

The lowering and raising isomorphisms **Lower** and **Raise** induce vector space isomorphisms between $\mathcal{G}(V)$ and $\mathcal{G}(V^*)$.

Further, at the end of Section 27 we saw how to use G to create inner products on each $\Lambda^r(V)$ and, using $G^\star$, on each $\Lambda_r(V)$.

By defining for each multi-vector x and y in $\mathcal{G}(V^*)$ the sum

$$< x, y >= \sum_{i=0}^{n} \langle \, \lceil x \rceil_i, \; \lceil y \rceil_i \, \rangle_i$$

these inner products induce an inner product on all of $\mathcal{G}(V^*)$.

An analogous procedure using the dual inner product $G^\star$ gives us an inner product on the space $\mathcal{G}(V)$ of multi-forms.

The interior product on s-forms and s-vectors induces in the obvious way an **interior product** on $\mathcal{G}(V^*)$ and on $\mathcal{G}(V)$, lowering the highest nonzero grade (if any) of a multi-vector or multi-form by one.

For instance, if $v \in V$ and $\theta \in \mathcal{G}(V)$ we define

$$v \lrcorner \theta \;\; = \;\; \sum_{i=0}^{n} v \lrcorner \lceil \theta \rceil_i .$$

The eponymous Hermann Grassmann, around and after 1840, wrote a series of works within which he essentially invented linear and multilinear algebra as we know it, including the first use of vectors. His writing style was considered to be unmotivated and opaque by the few who read his work. In this, as in other matters, he was a bit ahead of his time.

The value of his ideas did not become widely known for seventy years, and then mostly by reference, through the work of other mathematicians and physicists such as Peano, Clifford, Cartan, Gibbs and Hankel.

Grassmann interpreted nonzero wedge products of vectors such as $W = x_1 \wedge \cdots \wedge x_r$ to represent an r dimensional subspace of V, in light of the fact that this wedge product is characterized (up to constant multiple) by the property that $W \wedge y = 0$ and $y \in V$ if and only if y is in the span of $x_1, \ldots, x_r$.

He thought of the various operations on the Grassmann algebra as constituting a calculus of subspaces or "extensions," combining and giving information about geometrical objects and associated physical processes in a very direct and intuitive way. For instance

$$(\, x_1 \wedge \cdots \wedge x_r \,) \wedge (\, y_1 \wedge \cdots \wedge y_s \,)\, ,$$

will be nonzero if and only if $x_1, \ldots, x_r, y_1, \ldots, y_s$ form a linearly independent set of vectors. In that case the wedge product determines the combined span of the two spaces the factors represent, which will have dimension $r + s$.

On the other hand, if the combined span of the two spaces is all of V but neither of the two tensors is, individually, 0 and if V is given an inner product and orientation, the multi-vector

$$*\left(\left(*\,x_1 \wedge \cdots \wedge x_r\right) \wedge \left(*\,y_1 \wedge \cdots \wedge y_s\right)\right)$$

determines the intersection of the two spaces which will have dimension $r+s-n$.

His work, reformulated and clarified with modern notation, plays a large and increasing role in many areas, from current theoretical physics to practical engineering.

References

[1] Abraham, R. and Marsden, J., *Foundations of Mechanics*. Addison-Wesley Publishing Co., Reading, MA, 1985

[2] Aczél, J., *Lectures on Functional Equations and Their Applications*. Dover Publications, Inc., New York, NY, 2006

[3] Akivis, M. and Goldberg, V., *An Introduction to Linear Algebra and Tensors*. Dover Publications, Inc., New York, NY, 1972

[4] Arnold, V., *Mathematical Methods of Classical Mechanics*. Springer-Verlag New York, Inc., New York, NY, 1980

[5] Bishop, R. and Crittenden, R., *Geometry of Manifolds*. Academic Press, New York, NY, 1964

[6] Bishop, R. and Goldberg, S., *Tensor Analysis on Manifolds*. Dover Publications, Inc., New York, NY, 1980

[7] Flanders, H., *Differential Forms with Applications to the Physical Sciences*. Dover Publications, Inc., New York, NY, 1989

[8] Hoffman, K. and Kunze, R., *Linear Algebra*. Prentice-Hall Inc., Inglewood Cliffs, NJ, 1971

[9] Hungerford, T., *Algebra*. Springer-Verlag Inc., New York, NY, 1974

[10] José, J., and Saletan, E., *Classical Dynamics A Contemporary Approach*. Cambridge University Press, Cambridge, U.K., 1998

[11] Misner, C. and Thorne, K. and Wheeler, J., *Gravitation*. W. H. Freeman and Co., San Fancisco, CA, 1973

[12] Nelson, E., *Tensor Analysis*. Princeton University Press, Princeton, NJ, 1967

[13] Spivak, M., *Calculus on Manifolds*. W.A. Benjamin, Inc., Menlo Park, CA, 1965

[14] Spivak, M., *Differential Geometry Vol. I*. Publish or Perish, Inc., Berkeley, CA, 1970

[15] Stacy, B. D., *Tensors in Euclidean* 3-*Space from a Course Offered by J. W. Lee*. Notes transcribed and organized by B. D. Stacy, 1981

[16] Sussman, G., and Wisdom, J., and Mayer, M., *Structure and Interpretation of Classical Mechanics*. The MIT Press, Cambridge, MA, 2001

[17] Synge, J. L., and Schild, A., *Tensor Calculus*. Dover Publications, Inc., New York, NY, 1978

[18] Wald, R., *General Relativity*. University of Chicago Press, Chicago, IL, 1984

[19] Weinreich, G., *Geometrical Vectors*. University of Chicago Press, Chicago, IL, 1998

Index