# Theory and application of Grassmann Algebra

by

**William C. Schulz**

August 31, 2011

Transgalactic Publishing Company
Flagstaff, Vienna

# Contents


**1 Introduction** **1**

**2 Linear Algebra** **5**
2.1 Introduction . . . . . . . . 6
2.2 Vector Spaces . . . . . . . . 7
2.3 Bases in a Vector Space . . . . . . . . 9
2.4 The Dual Space $V^*$ of $V$ . . . . . . . . 15
2.5 Inner Products on $V$ . . . . . . . . 19
2.6 Linear Transformations . . . . . . . . 25

**3 Tensor Products of Vector Spaces** **33**
3.1 Introduction . . . . . . . . 34
3.2 Multilinear Forms and the Tensor Product . . . . . . . . 35
3.3 Grassmann Products from Tensor Products . . . . . . . . 40

**4 Grassmann Algebra on the Vector Space $V$** **47**
4.1 Introduction . . . . . . . . 48
4.2 Axioms . . . . . . . . 49
4.3 Permutations and Increasing Permutations . . . . . . . . 53
4.4 Determinants . . . . . . . . 61
4.5 Extending a Linear Transformation to the Grassmann Algebra and the Cauchy–Binet Theorem . . . . . . . . 71
4.6 The Equivalence of Axioms 4a and 4b . . . . . . . . 75
4.7 Products and Linear Dependence . . . . . . . . 79

**5 Grassmann Algebra on the Space $V^*$** **81**
5.1 Introduction . . . . . . . . 82
5.2 Grassmann's Theorem by Fiat . . . . . . . . 83
5.3 Grassmann's Theorem by Tensor Products . . . . . . . . 84
5.4 Grassmann's Theorem by Use of a Basis . . . . . . . . 85
5.5 The Space of Multilinear Functionals . . . . . . . . 87
5.6 The Star Operator and Duality . . . . . . . . 90
5.7 The $\delta$ systems and $\epsilon$ systems . . . . . . . . 96

**6 Inner Products on $V$ and $V^*$** **103**
6.1 Introduction . . . 104
6.2 Exporting the Inner Product on $V$ to $V^*$, $\Lambda^r(V)$ and $\Lambda^r(V^*)$ . . 105
6.3 The Unit Boxes $\Omega_0$ and $\Omega_0^*$ . . . 114
6.4 $*$ Operators Adapted to Inner Products. . . . 118
6.5 Coordinate formulas for $*$-Operators . . . 124
6.6 Formulas for Orthogonal bases . . . 128

**7 Regressive Products** **135**
7.1 Introduction and Example . . . 136
7.2 Definition of the Regressive Product . . . 139
7.3 Change of Basis . . . 143
7.4 Dot Notation . . . 147

**8 Applications to Vector Algebra** **149**
8.1 Introduction . . . 150
8.2 Elementary $n$-dimensional Vector Algebra . . . 151
8.3 Standard Interpretation of the Grassmann Algebra $\Lambda^r(V)$ of a Real Vector Space $V$ . . . 158
8.4 Geometrical Interpretation of $V^*$ . . . 166
8.5 Geometrical Interpretation of $* : \Lambda^r(V^*) \to \Lambda^{n-r}(V)$ and $* : \Lambda^r(V) \to \Lambda^{n-r}(V^*)$ . . . 177

**9 Applications to Projective Geometry** **183**
9.1 Introduction . . . 184
9.2 Standard Projective Geometry . . . 185
9.3 Weight Functions . . . 193

# Chapter 1

# Introduction

## 1.1 Introduction

Heinrich Günther Grassmann published his book Die Lineale Ausdehnungslehre in 1842. The book contained and exposition of $n$–dimensional linear algebra, an alternating product, inner products and a duality operator, among many other things. The style of the book was roughly comparable to the more abstract writings of the 1930's and was perceived in its own day as being so wildly abstract as to be incomprehensible. The book was not well received, although it has always been admired by a limited number of enthusiasts.

Many of Grassmann's ideas have been subsequently rediscovered, but generally in piecemeal fashion, and Grassmann's imposing edifice has never received the recognition it deserves. In addition, mathematicians are generally unaware of how much of modern mathematics traces back to Grassmann's ideas.

The purpose of this book is to lay out some of Grassmann's major ideas in a modern formulation and notation. The most serious departure from Grassmann's own presentation is the recognition of the complementary roles of the vector space $V$ and its dual space $V^*$, perhaps the only part of modern linear algebra with no antecedents in Grassmann's work.

Certain technical details, such as the use of increasing permutations or the explicit use of determinants also do not occur in Grassmann's original formulation. I have, with some reluctance, used the modern $\wedge$ instead of Grassmann's notation, although in some circumstances I revert to Grassmann's notation when it is clearly more convenient for calculation.

Another departure from Grassmann is the primacy given to vectors over points in the present exposition. In chapter eight I show that this is merely cosmetic; the same abstract structure admits two apparently different geometric interpretations. I there show how the two interpretations are related and how to move back and forth between them. The motivation for departing from Grassmann's point–based system and using vectors is the desire to introduce Grassmann's ideas in the most familiar possible setting. The vector interpretation is more useful for applications in differential geometry and the point interpretation is more suited for projective geometry.

One of the goals of this book is to lay out a consistent notation for Grassmann algebra that encompasses the majority of possible consumers. Thus we develop the theory for indefinite (but non degenerate) inner products and complex scalars. The additional effort for this level of generality over real scalars and positive definite inner products is very modest and the way is cleared for the use of the material in modern physics and the lower forms of algebraic geometry. We simply must be a little careful with factors of $(-1)^s$ and conjugate bars. I have, with reluctance, not developed the theory over commutative rings, because that level of generality might obscure Grassmann's ideas.

While it is not possible to eliminate bases altogether from the development, I have made a great effort to use them as little as possible and to make the proofs as invariant as I could manage. In particular, I have given here an invariant treatment of the $*$ duality operator which allows the algorithmic computation of $*\circ*$ without use of bases, and this has a lot of methodological advantages. I have also found ways to express $*$ algebraically so that it seems a lot more pleasant and natural than is generally the impression. Also, I never prove anything with orthonormal bases, which I consider a great source of confusion however convenient they sometimes prove. The algorithmic treatment given here also helps to avoid errors of the lost-minus-sign type.

I have adopted a modular methodology of introducing the Grassmann product. In chapter two we introduce tensor products, define Grassmann products in terms of them, and prove certain fundamental laws about Grassmann products. These fundamental laws then are used as Axioms in chapter three to develop the fundamental theory of the Grassmann product. Thus a person satisfied to go from the axioms can skip chapter two and go directly to chapter three. Chapter two is used again only very late in the book, in a sketch of differential geometry.

For much of the foundational material the plan is to develop the theory invariantly, then introduce a basis and see how the theory produces objects related to the basis, and finally to discuss the effect of changing the basis. Naturally some of this material is a trifle dull, and the reader should feel free to skip it and return when motivated to learn more about it.

None of the material in this book is deep in a mathematical sense. Nevertheless, I have included a vast amount of detail so that (I hope) the book is easy to read. Feel free to skim through computations if you find them dull. I have always erred on the side of including more rather than less detail, so the book would be easy to read for the less experienced. My apologies to the more experienced, who must be prepared to skip. One of the reasons for including vast numbers of explicit formulas is for the convenience of persons who may wish to implement one aspect or another on computers.

One further point: while none of the material in this book is deep, it is possible to make a horrible hash of it by incorrect technique. I try to provide the reader with suitable tools with which he can make computations easily and happily. In particular, the $*$–operator is often underutilized because the more generally known formulas for it are often infelicitous, and one of the purposes of this book is to remedy this defect.

I have tried to make the pace of the book leisurely. In order to make the book easily readable for persons who just dip in anywhere, I have often repeated calculations rather than referring to previous occurrences. Another reason for doing this is to familiarize the reader with certain tricks and devices, so these are sometimes repeated rather than referring the reader back to previous occurrences.

In the latter part of the book which deals with applications certain compromises had to be made to keep the size of the book within bounds. These consist of intuitive descriptions of the manifold concept and some of the analytic apparatus. To do otherwise would have meant including entire textbooks on manifolds and analysis, which was not practical. I have tried to give the reader a good intuitive description of what is going on in areas where great detail was clearly inappropriate, for example the use of Sobolev spaces in the section on harmonic forms. I apologize in advance to the cognoscenti in these areas for the omission of favorite technical devices. The desire always was to make the book as accessible as possible to persons with more elementary backgrounds.

With regard to motivation, the book is perhaps not optimally organized. One always has to choose between systematic exposition in which things are logically organized and fully developed and then applied versus a style which mixes applications in with theory so that the motivation for each development is clear. I have gone with the first alternative in order to lay out the theory more systematically.

No single critical feature in the exposition is my own invention. The two most important technical tools are the increasing permutations and the duality $* : \Lambda^r(V) \to \Lambda^{n-r}(V^*)$. The first was developed by Professor Alvin Swimmer of Arizona State University. I do not know the ultimate origin of the second item but I encountered it in the book [Sternberg]. It is my pleasure to express a vast debt to Professor Swimmer for introducing me to Grassmann algebra and allowing me the use of his own unpublished manuscript on the subject.

# Chapter 2

# Linear Algebra

## 2.1 Introduction

**2.1 Introduction**

In this chapter we wish to sketch the development of linear algebra. To do this in detail would of course take a whole book. We expect that the reader has some experience with linear algebra and we will use this chapter to remind her of the basic definitions and theorems, and to explain the notation we will be using. We will use standard tensor notation with superscripts and subscripts and we will explain the rules for this.

We are greatly interested in the correlation between the objects like vectors and the columns of numbers that represent the vector in a given basis. We are also interested in formulas for change of basis. In order to give this chapter at least some interest, we will lay out an interesting way of quickly deriving these formulas, even though it is perhaps mathematically odd and may offend the squeamish. I have noticed others using this notation in the last couple of decades, and I do not lay any claim to novelty.

## 2.2 Vector Spaces

### 2.2 Vector spaces

A *vector space* is a mathematical system involving two kinds of entities, the scalars and the vectors. The scalars are to be thought of as numbers, and the mathematical way of expressing this is to say that they form a field. A field is a mathematical system with two operations + (Addition) and $\cdot$ (Multiplication, generally omitted; $\alpha \cdot \beta = \alpha\beta$) in which every element $\alpha$ has an additive inverse $-\alpha$ and in which every non-zero element $\alpha$ has an additive inverse $\alpha^{-1}$ and in which there is an additive identity 0 and a multiplicative identity $1 \neq 0$ and which satisfies the following laws.

| | | |
|---|---|---|
| A1 | $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ | Associative law |
| A2 | $\alpha + 0 = 0 + \alpha = \alpha$ | Identity law |
| A3 | $\alpha + (-\alpha) = (-\alpha) + \alpha = 0$ | Inverse law |
| A4 | $\alpha + \beta = \beta + \alpha$ | Commutative law |
| M1 | $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ | Associative law |
| M2 | $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$ | Identity law |
| M3 | If $\alpha \neq 0$ then $\alpha\alpha^{-1} = \alpha^{-1}\alpha = 1$ | Inverse law |
| M4 | $\alpha\beta = \beta\alpha$ | Commutative law |
| D1 | $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ | Left Distributive law |
| D2 | $(\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha$ | Right Distributive law |

These laws are not all independent; for example D2 can be proved from D1 and M4. They are chosen as they are for symmetry and certain other considerations.

It is easy to prove from these laws that 0 and 1 are the unique identities and that each element has a unique additive inverse and each $x \neq 0$ has a unique multiplicative inverse.

We will generally think of the scalars as being the real or the complex numbers in this book, but much of what we do will work of any field that is not of characteristic 2, that is fields in which we do not have $1 + 1 = 0$. Places where difficulties may occur are noted in passing.

In using the complex numbers is often necessary to deal with the conjugate $\overline{z}$ of $z$. This too can be generalized; the properties of conjugate we require are

$$\begin{aligned} \overline{x+y} &= \overline{x} + \overline{y} \\ \overline{x \cdot y} &= \overline{x} \cdot \overline{y} \\ \overline{\overline{x}} &= x \\ \overline{x}x &> 0 \text{ for } x \neq 0 \end{aligned}$$

For this to make sense, $\overline{x}x$ must be in a subset of the field which has a linear ordering on it, just as in the complex numbers.

Practically speaking, I suggest the reader think in terms of the fields of real and complex numbers, and ignore the conjugate bar in the case of real numbers or other fields of scalars.

Now we can define a *Vector Space* . A vector space is a mathematical system with two sorts of objects, the field of scalars and the vectors. We will use Greek letters $\alpha, \beta, \gamma, \ldots$ for the scalars and Latin letters $u, v, w, \ldots$ for the vectors. The vectors may be added $u + v$ and vectors may be multiplied by scalars $\alpha v$. (It would be methodologically more correct to write the scalars to the right of the vectors $v\alpha$, but such is not the common usage.) There is an additive identity 0 which is distinct from the 0 of the field of scalars but is customarily written with the same symbol. There is an additive inverse $-v$ for each vector $v$. The addition and scalar multiplication satisfy the following laws.

| | | |
|---|---|---|
| V1 | $u + (v + w) = (u + v) + w$ | Associative law |
| V2 | $v + 0 = 0 + v$ | Identity law |
| V3 | $v + (-v) = (-v) + v = 0$ | Inverse law |
| V4 | $v + w = w + v$ | Commutative law |
| D1 | $\alpha(v + w) = \alpha v + \alpha w$ | Distributive law (scalars over vectors) |
| D2 | $(\alpha + \beta)v = \alpha v + \beta v$ | Distributive law (vectors over scalars) |
| U1 | $1 \cdot v = v$ | Unitary law |

The Unitary law U1 has the job of preventing $\alpha v = 0$ for all $\alpha$ and $v$, a pathology not prevented by the other laws.

We define $v - w = v + (-w)$. From the above basic laws for a vector space, the following may easily be derived.

$0 \cdot v = 0$ first 0 is scalar 0, second 0 is vector 0
$\alpha \cdot 0 = 0$
$(-1) \cdot v = -v$
If $\alpha v = 0$ then $\alpha = 0$ or $v = 0$

It is also easy to prove that there is a unique additive identity 0 and it is the solution $w$ of $v + w = v$ for any $v$. Also there is a unique additive inverse $-v$ for each $v$ and it is the solution $w$ of $v + w = 0$, and is unique.

## 2.3 Bases in a Vector Space

**2.3 Bases in a Vector Space**

In this section we study the concepts of span, linear independence, basis, representation in a basis, and what happens to the representation when we change the basis. We begin with some critical definitions.

**Def** If $v_1, v_2, \dots v_r$ is a set of vectors in a vector space $V$, then a linear combination of the $v_i$ is any expression of the form

$$\lambda^1 v_1 + \lambda^2 v_2 + \dots + \lambda^r v_r$$

Here the $\lambda^i$ are scalars, and the indices $i$ that look like exponents are really just labels, like subscripts. They are called superscripts. There is a sense to whether the labels are superscripts or subscripts, and we will eventually explain how the position is meaningful. For the moment we want to note that in almost all cases a sum will consist of a summation index written twice, once up, once down. The above expression could be written

$$\sum_{i=1}^{r} \lambda^i v_i$$

Albert Einstein discovered that one could conveniently leave off the sum sign as long as the range of the summation, 1 to $r$, stays the same for the various computations, which is normally the case. Hence Einstein suggested leaving the summation sign off and writing just

$$\lambda^i v_i$$

where the summation is indicated by the presence of the same index $i$ written once in an up position and once in a down position. This is called the *Einstein summation convention* and it is extremely convenient. We will use it throughout the book. However, it will turn out that summing from 1 to $r$ is a special case of summing over a certain kind of permutation, and we will extend the summation convention in chapter three to cover this also.

Next we want to define

**Def** The *span* of a set of vectors $v_1, v_2, \dots, v_r$ is the set of all linear combinations

$$\lambda^i v_i$$

of the $v_i$ where the $\lambda^i$ run through all elements of the field of scalars. The span of any (nonempty) set of vectors is a subspace of the vector space $V$. We denote it by $[v_1, v_2, \dots, v_r]$ in this chapter. We then have

**Def** A set of vectors $\{v_1, v_2, \dots, v_r\}$ is *linearly independent* if it satisfies the following condition. If $\lambda^1, \dots, \lambda^r$ are scalars and

$$\sum_{i=1}^{r} \lambda^i v_i = 0$$

then

$$\lambda^1 = \lambda^2 = \ldots = \lambda^r = 0\,.$$

This is the most difficult concept of elementary vector space theory to understand, so we will talk a little about the concept. Suppose we are in $\mathbf{R}^3$ and suppose we have

$$2v_1 + 3v_2 - 5v_3 = 0\,.$$

Then we can solve this equation for and of the $v_i$, for example $v_3$ to get

$$v_3 = \frac{2}{5}v_1 + \frac{3}{5}v_2\,.$$

If we think of the way vectors are added in $\mathbf{R}^3$ we see that this means $v_3$ is in the plane determined by $v_1$ and $v_2$. Similarly, in $\mathbf{R}^5$ the equation

$$2v_1 + 3v_2 - 5v_3 + 0v_4 + 2v_5 = 0$$

means that

$$v_2 = -\frac{2}{3}v_1 + \frac{5}{3}v_3 - \frac{2}{3}v_5$$

so that $v_2$ is in the space spanned by $v_1, v_3, v_5$. (Notice that from this equation we can say nothing about $v_4$.) Conversely, if some $v_i$ is a linear combination of vectors $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_r$ then we will have an equation of type

$$\lambda^i v_i = 0 \qquad \text{SUMMATION CONVENTION IN FORCE!}$$

in which not all the $\lambda^i$ are 0. Thus linear independence is a condition that requires that no $v_i$ in the set $\{v_1, \ldots, v_r\}$ is a linear combination of the remaining ones, and so each $v_i$ is not in the span of the remaining ones. Geometrically speaking, each $v_i$ "sticks out" of the linear subspace generated by the remaining ones.

Now we can define

**Def** A set of vectors $\{e_1, \ldots, e_n\}$ is a *basis* of the vector space $V$ if it is

1. Linearly independent and
2. Spans the space $V$

A vector space is said to be *finite dimensional* if it has a basis with a finite number $n$ of vectors. We have the

**Theorem** The number of a vectors in a basis of a finite dimensional vector space $V$ is the same no matter which basis is chosen.

**Def** The *dimension* of a finite dimensional vector space $V$ is the number of vectors in any basis of the $V$.

In this section the dimension of $V$ will always be $n$. We now want to consider the representation of an element of a vector space in terms of a basis. First we have

**Theorem** the representation $v = \xi^i e_i$ of $v \in V$ in terms of a basis is unique.

**Proof** First, $v$ *has* a representation in terms of a basis since the set $e_1, \dots, e_n$ spans the space $V$. Suppose now that it has two such representations

$$\begin{aligned} v &= \xi^i e_i \\ v &= \eta^i e_i \end{aligned}$$

Then

$$0 = (\xi^i - \eta^i) e_i \,.$$

Since the $e_i$ are linearly independent, $\xi^i - \eta^i = 0$ so $\xi^i = \eta^i$

We will often find it convenient to place the $\xi^i$ in a column, that is an $n \times 1$ matrix, with entries

$$\begin{pmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \xi^n \end{pmatrix}$$

in the field of scalars. There is always a danger of confusing this column with the vector itself, which is analogous to confusing a shoe size, which *measures* a shoe, with the shoe itself. The column vector *measures* the vector, where the $e_i$ are somewhat analogous to units of measure. (This analogy should not be pushed too far.)

Now just as we can change measuring units from inches to centimeters, we can change the basis in a vector space.. Suppose $\{e_1, \dots, e_n\}$ are the original basis vectors, which we will call the *old basis,* and that $\{\tilde{e}_1, \dots, \tilde{e}_n\}$ is another basis, which we will call the *new basis.* then the vector $v \in V$ can be written in either basis:

$$\begin{aligned} v &= \xi^i e_i \\ v &= \tilde{\xi}^i e_i \end{aligned}$$

and we would like the connection between the new coordinates and $\{\tilde{\xi}^1, \dots, \tilde{\xi}^n\}$ the old coordinates $\{\xi^1, \dots, \xi^n\}$. For this, we express the new basis $\{\tilde{e}_1, \dots, \tilde{e}_n\}$ in terms of the old basis $\{e_1, \dots, e_n\}$. We have

$$\begin{aligned} \tilde{e}_1 &= \alpha_1^1 e_1 + \alpha_1^2 e_2 + \ldots + \alpha_1^n e_n \\ \tilde{e}_2 &= \alpha_2^1 e_1 + \alpha_2^2 e_2 + \ldots + \alpha_2^n e_n \\ &\vdots \\ \tilde{e}_n &= \alpha_n^1 e_1 + \alpha_n^2 e_2 + \ldots + \alpha_n^n e_n \end{aligned}$$

This can be nicely digested as

$$\tilde{e}_i = \alpha_i^j e_j$$

We can then put the $(\alpha_i^j)$ into a matrix as follows

$$\mathcal{C} = (\alpha_i^j) = \begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \dots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \dots & \dots & \dots & \dots \\ \alpha_1^n & \alpha_2^n & \dots & \alpha_n^n \end{pmatrix} \qquad \text{NOTE TRANSPOSITION FROM ABOVE ARRAY!}$$

The matrix elements do not come in the order they do in $\tilde{e}_i = \alpha_i^j e_j$; rows there have changed to columns in the matrix.

Some inconvenience of this sort will always occur no matter how things are arranged, and the way we have done it is a quite common (but not universal) convention. The matrix $\mathcal{C}$ will be referred to as the *change of basis matrix.* It will now be easy to find the relationship between $\xi^i$ and $\tilde{\xi}^i$:

$$\xi^j e_j = v = \tilde{\xi}^i \tilde{e}_i = \tilde{\xi}^i \alpha_i^j e_j\,.$$

Since representation in terms of a basis is unique we have

$$\xi^j = \alpha_i^j \tilde{\xi}^i$$

which can be written in matrix form as

$$\begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} = (\alpha_i^j) \begin{pmatrix} \tilde{\xi}^1 \\ \vdots \\ \tilde{\xi}^n \end{pmatrix} = \mathcal{C} \begin{pmatrix} \tilde{\xi}^1 \\ \vdots \\ \tilde{\xi}^n \end{pmatrix}\,.$$

Notice that this is not the same as for the basis vectors. One must remember this! Generally speaking, there are just two ways things change when bases are changed; either like the $e_i$ (called *covariant change*) or like the $\xi^j$ called *contravariant change*). The indices are placed up or down according to the way the object changes. (Historically, "covariant" means "varies like the basis vectors" and "contravariant" means "varies the other way." It has been (repeatedly) suggested that the terminology is opposite to the way it should be. However if the terminology *were* opposite it would probably generate the identical suggestion.)

We now want to exploit matrix multiplication to derive the basis change rule in a new way. We are using matrices here as a formal convenience and some readers will find it uncomfortable. Fortunately, it will never be necessary to use this technique of one does not like it.

First, recall that if $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are matrices with the following dimensions

$$\begin{array}{lcll} \mathcal{A} & = & (\alpha_j^i) & m \times n \\ \mathcal{B} & = & (\beta_j^i) & n \times p \\ \mathcal{C} & = & (\gamma_j^i) & n \times p \end{array}$$

and $\mathcal{C} = \mathcal{A}\mathcal{B}$, then we have

$$\gamma_k^i = \alpha_j^i \beta_k^j \qquad \text{SUMMATION CONVENTION IN FORCE!}$$

which we can also write out explicitly as a matrix equation

$$(\gamma^i_k) = (\alpha^i_j)(\beta^j_k) .$$

Now, using matrix multiplication we can write

$$v = e_j \xi^j = (e_1, \ldots, e_n) \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} .$$

(Here is the source of the discomfort; the first matrix has vectors for entries. If this upsets you, remember it is only for mnemonic convenience; one can always default back to $v = \xi^i e_i$.)

The change of basis in terms of the old basis $\{e_1, \ldots, e_n\}$. We have

$$\begin{array}{rcl} \tilde{e}_1 & = & \alpha^1_1 e_1 + \alpha^2_1 e_2 + \ldots + \alpha^n_1 e_n \\ \tilde{e}_2 & = & \alpha^1_2 e_1 + \alpha^2_2 e_2 + \ldots + \alpha^n_2 e_n \\ & & \vdots \\ \tilde{e}_n & = & \alpha^1_n e_1 + \alpha^2_n e_2 + \ldots + \alpha^n_n e_n \end{array}$$

can now be written as

$$\begin{array}{rcl} (\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n) & = & (e_1, e_2, \ldots, e_n) \begin{pmatrix} \alpha^1_1 & \ldots & \alpha^1_n \\ \alpha^2_1 & \ldots & \alpha^2_n \\ \vdots & \vdots & \vdots \\ \alpha^n_1 & \ldots & \alpha^n_n \end{pmatrix} \\ & = & (e_1, e_2, \ldots, e_n)\, \mathcal{C} \end{array}$$

We then have

$$(e_1, \ldots, e_n) \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} = v = (\tilde{e}_1, \ldots, \tilde{e}_n) \begin{pmatrix} \tilde{\xi}^1 \\ \vdots \\ \tilde{\xi}_n \end{pmatrix} = (e_1, \ldots, e_n)\, \mathcal{C} \begin{pmatrix} \tilde{\xi}^1 \\ \vdots \\ \tilde{\xi}_n \end{pmatrix}$$

so that (since representation in a basis is unique)

$$\begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} = \mathcal{C} \begin{pmatrix} \tilde{\xi}^1 \\ \vdots \\ \tilde{\xi}_n \end{pmatrix} .$$

This kind of shorthand is very convenient for quick derivations once one gets used to it. It is also fairly clear that it could be made rigorous with a little additional effort.

## 2.4 The Dual Space $V^*$ of $V$

**2.4 The Dual Space $V^*$ of $V$**

We will now define the space $V^*$ of linear functionals on the vector space $V$, show how to represent them with a basis and calculate what happens when the basis changes. Let $\mathbb{F}$ be the field of scalars of our vector space. A *linear functional* on the vector space $V$ is a function $f : V \to \mathcal{F}$ satisfying the property

**Def** For all $u, v \in V,\ \ \alpha, \beta \in \mathcal{F}$

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)\,.$$

this is equivalent to the two properties

$$\begin{aligned} f(u+v) &= f(u) + f(v) \\ f(\alpha u) &= \alpha f(u) \end{aligned}$$

The property is called *linearity* and $f$ is said to be *linear on* $V$.

The set of linear functionals on $V$ is itself a vector space, denoted by $V^*$. Addition is defined by

$$\big(f+g\big)(v) = f(v) + g(v)$$

and scalar multiplication by

$$\big(\alpha f\big)(v) = \alpha \cdot \big(f(v)\big)\,.$$

Since $V^*$ is a vector space we naturally want to find a basis. Define a linear functional $e^i$ by the rule, for $v = \xi^j e_j \in V$,

$$\begin{aligned} e^i(v) = e^i(\xi^j e_j) &= e^i(\xi^1 e_1 + \ldots \xi^{i-1} e_{i-1} + \xi^i e_i + \xi^{i+1} ei + 1 + \ldots + \xi^n e_n \\ &= \xi^i\,. \end{aligned}$$

It is trivial to verify that $e^i$ is a linear functional and that

$$e^i(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}\ .$$

This situation occurs so frequently that it is useful to have a notation for it:

**Def**
$$\delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}\ .$$

This $\delta^i_j$ is called the *Kronecker delta,* after the German mathematician Leopold Kronecker (1823-1891), and we may now write

$$e^i(e_j) = \delta^i_j \qquad\qquad \text{for } e_j \in V,\ \ e^i \in V^*\,.$$

We claim the set $\{e^1, \dots e^n\}$ is a basis of $V^*$. Indeed, suppose $f \in V^*$ and $f(e_j) = \lambda_j$. Then

$$\begin{aligned}
\lambda_j e^j(v) &= \lambda_j e^j(\xi^i e_i) = \lambda_j \xi^i e^j(e_i) = \lambda_j \xi^i \delta^i_j \\
&= \lambda_j \xi^j = f(e_j)\xi^j = f(\xi^j e_j) \\
&= f(v)\,.
\end{aligned}$$

Thus $\lambda_j e^j$ and $f$ have the same value on any vector $v \in V$, and thus $\lambda_j e^j = f$. Hence the set $\{e^1, \dots, e^n\}$ spans $V^*$. Now suppose $\lambda_j e^j = 0$. The for any $v \in V$, $\lambda_j e^j(v) = 0$. Hence we have

$$\begin{aligned}
\lambda_j e^j(e_i) &= 0 \qquad\qquad i = 1, \dots, n \\
\lambda_j \delta^j_i &= 0 \\
\lambda_i &= 0
\end{aligned}$$

Thus $\{e^1, \dots, e^n\}$ is a linearly independent set and therefore is a basis of $V^*$. The basis $\{e^1, \dots, e^n\}$ of $V^*$ has a very special relationship with the basis $\{e_1, \dots, e_n\}$ of $V$ given by $e^i(e_j) = \delta^i_j$. We define

**Def** The set of linear functionals $\{e^1, \dots, e^n\}$ defined above and satisfying

$$e^i(e_j) = \delta^i_j$$

is called the *dual basis* of $V^*$.

It will play a supremely important role in our work.

As we saw above, any $f \in v^*$ can be represented in the dual basis as

$$f = \lambda_j e^j \qquad \text{where } \lambda_j = f(e_j)\,.$$

We will represent $f$ by the $1 \times n$ matrix

$$(\lambda_1, \lambda_2, \dots, \lambda_n)\,.$$

The value of $f(v)$ can now be found from the representatives of $f$ and $v$ (in the dual basis of $V^*$ and $V$) by matrix multiplication:

$$\begin{aligned}
f(v) &= (\lambda_i e^j)(\xi^j e_j) = \lambda_i \xi^j e^j(e_j) \\
&= \lambda_i \xi^j \delta^j_i = \lambda_i \xi^i \\
&= (\lambda_1, \dots, \lambda_n) \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}
\end{aligned}$$

In tensor theory $\lambda_i \xi^i$(summing over a repeated index) is called *contraction*.

Naturally we want to know how $(\lambda_1, \dots, \lambda_n)$ changes when we change the basis. The productive way to approach this is via the question: when the basis

$\{e_1, \dots, e_n\}$ is changed to $\{\tilde{e_1}, \dots, \tilde{e_n}\}$, how do $\{e^1, \dots, e^n\}$ and $(\lambda_1, \dots, \lambda_n)$ change? We have

$$\tilde{\lambda}_i = f(\tilde{e}_i) = f(\alpha_i^j e_j) = \alpha_i^j f(e_j) = \alpha_i^j \lambda_j$$

which shows how the $(\lambda_1, \dots, \lambda_n)$ changes:

$$(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n) = (\lambda_1, \dots, \lambda_n)(\alpha_i^j) = (\lambda_1, \dots, \lambda_n)(\alpha_i^j)\, \mathcal{C}$$

so that the representation of a linear functional changes *exactly like the basis vectors,* that is, *covariantly.*

To find the formula for the change in the dual basis, recall that if $f(\tilde{e}_i) = \tilde{\lambda}_i$ then $f = \tilde{\lambda}_i \tilde{e}^i$. Now

$$\begin{aligned} e^j(\tilde{e}_i &= e^j(\alpha_i^k e_k) = \alpha_i^k\, e^j(e_k) \\ &= \alpha_i^k \delta_k^j = \alpha_i^j \end{aligned}$$

so

$$e^j = \alpha_i^j\, \tilde{e}^i$$

and *the dual basis vectors change contravariantly.* We can write this in matrix form as

$$\begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = (\alpha_i^j) \begin{pmatrix} \tilde{e}^1 \\ \vdots \\ \tilde{e}^n \end{pmatrix} = \mathcal{C} \begin{pmatrix} \tilde{e}^1 \\ \vdots \\ \tilde{e}^n \end{pmatrix}$$

which is a contravariant change.

The reader will by now have noted that when the indices are high they count by the row, and when low they count by column. For example

$$\begin{matrix} \text{first row} & \longrightarrow \\ \text{second row} & \longrightarrow \\ \vdots & \vdots \\ n^{\text{th}} \text{ row} & \longrightarrow \end{matrix} \begin{pmatrix} e^1 \\ e^2 \\ \vdots \\ e^n \end{pmatrix}$$

An object with a single high index will then be written as a column and an object with a single low index will be written as a row.

We will now introduce a method of obtaining the change of basis equations by matrix methods. To do this we introduce an action of the linear functional $f \in V^*$ on a row matrix of vectors

$$f(v_1, \dots, v_r) = \big(f(v_1), \dots, f(v_r)\big)\,.$$

Then we have

$$\begin{aligned} (\tilde{\lambda}_1, \dots, \tilde{\lambda}_n) &= \big(f(\tilde{e}_1), \dots, f(\tilde{e}_n)\big) = f(\tilde{e}_1, \dots, \tilde{e}_n) \\ &= f(e_1, \dots, e_n)\, \mathcal{C} = \big(f(e_1), \dots, f(e_n)\big)\, \mathcal{C} \\ &= (\lambda_1, \dots, \lambda_n)\, \mathcal{C}\,. \end{aligned}$$

This is the same result we previously obtained. Note the quiet use here of the associativity of matrix multiplication which corresponds to the use of linearity in the original derivation.

From this we can easily derive the the change for the $e^i$ with this method; for $f \in V^*$

$$(\lambda_1, \ldots, \lambda_n) \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = f = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n) \begin{pmatrix} \tilde{e}^1 \\ \vdots \\ \tilde{e}^n \end{pmatrix} = (\lambda_1, \ldots, \lambda_n)\, \mathcal{C} \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix}$$

and since this is true for all $\lambda_1, \ldots, \lambda_n$ we must have

$$\begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = \mathcal{C} \begin{pmatrix} \tilde{e}^1 \\ \vdots \\ \tilde{e}^n \end{pmatrix} .$$

## 2.5 Inner Products on $V$

**2.5 Inner Products on $V$**

This section is a little more complex than the last two because we wish to simultaneously accommodate the symmetric and Hermitian inner products. To do this we require conjugation on the base field $\mathbb{F}$. Conjugation commutes with addition and multiplication

$$\begin{aligned} \overline{\alpha+\beta} &= \overline{\alpha}+\overline{\beta} \\ \overline{\alpha\beta} &= \overline{\alpha}\,\overline{\beta} \end{aligned}$$

and we assume it to involutive

$$\overline{\overline{\alpha}} = \alpha$$

and we also assume it is one to one and onto, which has as a consequence that

$$\alpha \neq 0 \Rightarrow \alpha\overline{\alpha} \neq 0\,.$$

The obvious example is conjugation for complex numbers; another example is in the field of rational numbers with $\sqrt{3}$ adjoined, called $\mathbb{Q}[\sqrt{3}]$ where

$$\overline{\alpha+\beta\sqrt{3}} = \alpha-\beta\sqrt{3} \qquad\qquad \alpha,\beta\in\mathbb{Q}\,.$$

This example is important in number theory. The most important example which is *not* the complex numbers is the case of an arbitrary field $\mathbb{F}$ where conjugation does not have any effect; $\overline{\alpha} = \alpha$. For example, this is the natural definition for the real numbers $\mathbb{F} = \mathbb{R}$ and and the rational numbers $\mathbb{F} = \mathbb{Q}$

With conjugation under control we proceed to the definition of the inner product

**Def** An inner product on a vector space $V$ is a function $(\ ,\ ) : V \times V \to \mathbb{F}$ satisfying

**1.** $(u, v+w) = (u,v) + (u,w)$
$(u+v, w) = (u,w) + (v,w)$

**2.** $(u, \alpha v) = \alpha(u,v)$
$(\alpha u, v) = \overline{\alpha}(u,v)$

**3.** $(v,u) = \overline{(u,v)}$

**4.** if $(v,u) = 0$ for all $u \in V$ then $v = 0$ .

If conjugation does nothing, ($\overline{\alpha} = \alpha$), then numbers 1 and 2 are called *bilinearity.* They are also sometimes called bilinearity when conjugation has an effect, and sometimes by a similar name like *semi–linearity* or *Hermitian linearity.* Number 3 is called *symmetric* when $\overline{\alpha} = \alpha$ and *the Hermitian property* when conjugation has an effect. Number 4 is called *non–degeneracy.*

A slight variant of the above definition where number two is replaced by

**1.** $(u, v + w) = (u, v) + (u, w)$

**2′.** $(u, \alpha v) = \overline{\alpha}(u, v)$

**3.** $(\alpha u, v) = \alpha(u, v)$

is also called an inner product. The difference is cosmetic but we must be careful about it because one way of "exporting" an inner product on $V$ to an inner product on $V^*$ gives $2'$ on $V^*$.

A special case of an inner product is one that replaces number 4 by

**1.** $(u, v + w) = (u, v) + (u, w)$

**4′.** $u \neq 0 \Rightarrow (u, u) > 0$ Positive Definite

Clearly for $4'$ to function the subset $\{(u, u) |\, u \in V\} \subseteq \mathbb{F}$ must have an order relation $<$ on it. This is the case when $\mathbb{F} = \mathbb{R}$ or when $\mathbb{F} = \mathbb{C}$ and conjugation is ordinary complex conjugation. Clearly $4' \Rightarrow 4$ so that this is indeed a special case. We say when $4'$ holds that we have a *positive definite* inner product. (*Definite* here means that if $u \neq 0$ then $(u, u)$ cannot be 0, so the word is a bit redundant in this context, but customary.)

Inner products are often used to introduce a concept with some of the properties of length into a vector space. If the inner product is positive definite and the field has square roots of positive elements then this length has the properties we expect of length, and it is defined by

$$\|u\| = \sqrt{(u, u)}\,.$$

If the inner product is not positive definite but the field $\mathbb{F}$ is $\mathbb{C}$ and the inner product has the property that $(u, u)$ is real then we may define "length" by

$$\|u\| = \sqrt{|(u, u)|}\,.$$

This "length" is used in the theory of relativity and though it has some unusual properties (there are non-zero vectors whose "lengths" are 0) it is still quite useful. However, we will not make extensive use of this length concept in this book.

We now wish to represent the inner product in terms of matrices. We first introduce the operation $*$ matrices:

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}^* = \begin{pmatrix} \overline{\alpha_{11}} & \overline{\alpha_{21}} & \cdots & \overline{\alpha_{m1}} \\ \overline{\alpha_{12}} & \overline{\alpha_{22}} & \cdots & \overline{\alpha_{m2}} \\ \vdots & \vdots & \dots & \vdots \\ \overline{\alpha_{1n}} & \overline{\alpha_{2n}} & \cdots & \overline{\alpha_{mn}} \end{pmatrix}$$

so that $*$ results in both the transposing and conjugation of the elements of the matrix. (If $\overline{\alpha} = \alpha$ then $*$ is merely the transpose.) It is easy to check that

$$(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$$

We now define

**Def** $\mathcal{A}$ is Hermitian $\iff \mathcal{A}^* = \mathcal{A}$

We remind the reader that if conjugation does nothing ($\overline{\alpha} = \alpha$) then *Hermitian* means simply *symmetric* ($\mathcal{A}^T = \mathcal{A}$). We now form the matrix

$$(g_{ij}) = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}$$

of the inner product with respect to the basis $\{e_1, \cdots, e_n\}$ of $V$ by

$$g_{ij} = (e_i, e_j)\,.$$

Note that

$$\overline{g_{ji}} = \overline{(e_j, e_i)} = (e_i, e_j) = g_{ij}$$

so that $(g_{ij})$ is a Hermitian (or symmetric) matrix. Note also that our former convention whereby the row of a matrix is counted by an upper index is here not applicable. For the matrix $(g_{ij})$ the *first index* counts the *row* and the *second index* counts the *column*.

Let now $v = \xi^i e_i$ and $w = \eta^j e_j$. We then have

$$\begin{aligned} (v, w) &= (\xi^i e_i, \eta^j e_j) \\ &= \overline{\xi^i}\eta^j (e_i, e_j) \\ &= g_{ij}\overline{\xi^i}\eta^j \end{aligned}$$

which gives the inner product $(v, w)$ in terms of the coefficients $\xi^i$ and $\eta^j$ of the vectors $v$ and $w$ in the basis representation and the matrix representation $(g_{ij})$ of the inner product in the same basis. This will be most important for the entire book. In matrix form we can write this as

$$(u, v) = (\,\overline{\xi_1}, \cdots, \overline{\xi_n}\,)\,(g_{ij}) \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix}\,.$$

If we wish, we can compress this further by setting

$$\mathcal{G} = (g_{ij}) \qquad \underline{\xi} = \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} \qquad \underline{\eta} = \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix}$$

and then

$$(u, v) = \underline{\eta}^* \mathcal{G} \underline{\xi}\,.$$

To find what happens to $(g_{ij})$ under change of basis we recall

$$\tilde{e}_i = \alpha_i^j e_j$$

so that

$$\begin{aligned} \tilde{g}_{ij} &= (\tilde{e}_i, \tilde{e}_j) = (\alpha_i^k e_k, \alpha_j^\ell e_\ell) \\ &= \overline{\alpha}_i^k \alpha_j^\ell (e_k, e_\ell) = \overline{\alpha}_i^k \alpha_j^\ell g_{kl} \,. \end{aligned}$$

When the field is $\mathbf{R}$ and $\overline{\alpha} = \alpha$ this looks like a covariant index change but when the field is $\mathbf{C}$ things are perturbed slightly and the first index has a conjugation. We can write this in matrix form as

$$(\tilde{g}_{ij}) = \mathcal{C}^*(g_{ij})\mathcal{C} \,.$$

The transpose is necessary because in $\mathcal{C}^*$ the summing index k counts rows in $(g_{kl})$ and rows in $\overline{\alpha}_i^k$, so for the matrix multiplication to function correctly the rows of $\overline{\mathcal{C}}$ must be switched to columns. More explicitly we need

$$\begin{pmatrix} \tilde{g}_{11} & \tilde{g}_{12} & \cdots & \tilde{g}_{1n} \\ \tilde{g}_{21} & \tilde{g}_{22} & \cdots & \tilde{g}_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{g}_{n1} & \tilde{g}_{n2} & \cdots & \tilde{g}_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} \overline{\alpha}_1^1 & \overline{\alpha}_1^2 & \cdots & \overline{\alpha}_1^n \\ \overline{\alpha}_2^1 & \overline{\alpha}_2^2 & \cdots & \overline{\alpha}_2^n \\ \vdots & \vdots & \cdots & \vdots \\ \overline{\alpha}_n^1 & \overline{\alpha}_n^2 & \cdots & \overline{\alpha}_n^n \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \cdots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_1^n & \alpha_n^2 & \cdots & \alpha_n^n \end{pmatrix}$$

We can get this basis change formula more simply by matrix methods. If $v = \xi^i e_i = \tilde{\xi}^k \tilde{e}_k$ and $w = \eta^j e_j = \tilde{\eta}^\ell \tilde{e}_\ell$ then

$$\begin{aligned} \underline{\tilde{\xi}}^*(\tilde{g}_{kl})\underline{\tilde{\eta}} &= (v, w) = \underline{\xi}^*(g_{ij})\underline{\eta} && (2.1) \\ &= (\mathcal{C}\underline{\tilde{\xi}})^*(g_{ij})(\mathcal{C}\underline{\tilde{\eta}}) && (2.2) \\ &= \underline{\tilde{\xi}}^*(\mathcal{C}^*(g_{ij})\mathcal{C})\underline{\tilde{\eta}} && (2.3) \end{aligned}$$

and since this much be true for all $\underline{\tilde{\xi}}$ and $\underline{\tilde{\eta}}$, we must have

$$(\tilde{g}_{kl}) = \mathcal{C}^*(g_{ij})\mathcal{C} \,.$$

Next we must investigate certain facts about the matrix $(g_{ij})$ of the inner product. We use here certain facts from the theory of determinants. We will develop these facts systematically and in detail in Chapter 2 which is independent of the inner product concept, but probably most readers are familiar with them already.

The most important property of $(g_{ij})$ is

$$\det(g_{ij}) \neq 0\,,$$

which means that it $(g_{ij})$ has an inverse, (which we will denote by $(g^{kl})$ ). To prove $\det(g_{ij}) \neq 0$, let us assume $\det(g_{ij}) = 0$. This means, as we show in chapter 2, that the columns of $(g_{ij})$ are linearly dependent: $\exists \xi^1, \dots, \xi^n$ not all 0 for which

$$\begin{pmatrix} g_{11} \\ g_{21} \\ \vdots \\ g_{n1} \end{pmatrix} \xi^1 + \begin{pmatrix} g_{12} \\ g_{22} \\ \vdots \\ g_{n2} \end{pmatrix} \xi^2 + \cdots \begin{pmatrix} g_{1n} \\ g_{2n} \\ \vdots \\ g_{nn} \end{pmatrix} \xi^n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

or

$$(g_{ij}) \begin{pmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \xi^n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Now let $v = \xi^i e_i$ and $u = \eta^j e_j$ be any vector in $V$. We have

$$(u, v) = \underline{\eta}^* (g_{ij}) \underline{\xi} = \underline{\eta}^* \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$$

Thus $(u, v) = 0$ for all $u$ and a non-zero $v$, which contradicts number 4 in the definition of an inner product.

We can extract a little more out of this result. Suppose $\{v_1, \dots, v_r\}$ is a set of vectors in $V$. Consider the determinant of the matrix of $(v_i, v_j)$:

$$\det \begin{pmatrix} (v_1, v_1) & \cdots & (v_1, v_r) \\ \vdots & \cdots & \vdots \\ (v_r, v_1) & \cdots & (v_r, v_r) \end{pmatrix}.$$

If the $v_i$ are linearly dependent, then for some $\xi^i$ not all 0 we have $\xi^i v_i = 0$. This will force a linear dependence

$$\xi^1 \begin{pmatrix} (v_1, v_1) \\ \vdots \\ (v_r, v_1) \end{pmatrix} + \xi^2 \begin{pmatrix} (v_1, v_2) \\ \vdots \\ (v_r, v_2) \end{pmatrix} + \cdots + \xi^r \begin{pmatrix} (v_1, v_r) \\ \vdots \\ (v_r, v_r) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

and hence the determinant is 0. If, however, $\{v_1, \dots, v_r\}$ are linearly *independent,* then $W = (span of \{v_1, \dots, v_r\})$ is a subspace of $V$ and it inherits the inner product. However, sadly, number 4 (nondegeneracy) in the definition of inner product may fail on $W$. However, if the inner product remains nondegenerate on $W$, as must happen for example if it is positive definite, then $\{v_1, \ldots, v_r\}$ is a basis for $W$ and by our previous result

$$\det((v_i, v_j)) \neq 0\,.$$

This determinant is called the *Grassmanian* of $\{v_1, \ldots, v_r\}$. Digesting, we have

**Theorem** If the inner product restricted to span$[v_1, \ldots, v_r]$ is non-degenerate then

$$\{v_1, \ldots, v_r\} \text{ is linearly independent } \iff \det((v_i, v_j)) \neq 0\,.$$

A basic theorem in the theory of inner product spaces is the following:

**Theorem** Let $f : V \to \mathcal{F}$ be a linear functional on $V$. Then there is a unique $u \in V$ for which, for all $v \in V$,

$$f(v) = (u, v)\,.$$

**Proof** We will prove this using coordinates. A coordinate free proof can be found in section 5.1. Let $f = \lambda_i e^i \in V^*$ and set $\xi^j = g^{ji}\overline{\lambda_i}$ where $(g^{kl}) = (g_{ij})^{-1}$. Then setting $u = \xi^j e_j$ we have, with $v = \eta^j e_j$,

$$\begin{aligned}
(u, v) &= g_{jk}\overline{\xi^j}\eta^k \\
&= g_{jk}\overline{g^{ji}\overline{\lambda_i}}\,\eta^k \\
&= g_{jk}\overline{g^{ji}}\lambda_i\eta^k \\
&= g^{ij}g_{jk}\lambda_i\eta^k \\
&= \delta^i_k\lambda_i\eta^k \\
&= \lambda_k\eta^k \\
&= f(v)
\end{aligned}$$

Thus the required $u$ exists. If there were two such $u$ we would have

$$(u_1, v) = f(v) = (u_2, v)$$

for all $v \in V$, and then

$$(u_1 - u_2, v) = 0 \qquad \text{for all } v \in V\,.$$

Thus $u_1 - u_2 = 0$ by non-degeneracy, and $u_1 = u_2$.

**Corollary** The mapping $\Phi(u) = f$ using the $u$ and $f$ of the last theorem is an anti-isomorphism from $V$ to $V^*$. (Anti-isomorphism means $\Phi(\alpha u) = \overline{\alpha}\Phi(u)$.)

Thus, an inner product sets up a canonical anti-isomorphism (canonical means it does not depend upon the basis) between $V$ and $V^*$. We can see it is canonical because it is fully specified by the equation

$$(u, v) = [\Phi(u)](v)\,.$$

Conversely, some persons like to define an inner product by starting with such a canonical anti-isomorphism and defining an inner product by this formula.

## 2.6 Linear Transformations

**2.6 Linear Transformations**

Let $V$ and $W$ be vector spaces with $\dim(V) = n$ and $\dim(W) = m$. A *Linear Transformation* or *Linear Operator* is a function from $V$ to $W$ satisfying

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$$

There is a huge amount of theory about linear operators in finite dimensional spaces and we are going to present only some elementary results, and some results without proofs. For more details the reader should consult any good linear algebra book, for example, [Gelfand]

The *Range*

$$R[T] = \{w \in W | (\exists u \in V) w = T(v) \subseteq W$$

of the linear operator is a subspace of $W$ and we denote its dimension by

$$r(t) = \dim R[T] \,.$$

We define the *nullspace* or *kernel* of $T$ as the set

$$N(T) = \{v \in V | Tv = 0\} \,.$$

$N(T)$ is a subspace of $V$ and we denote its dimension by

$$n(T) = \dim N(T) \,.$$

The operator $\tilde{T}$ defined on the factor space $v/N(T)$ onto $R[T]$ is an isomorphism and is defined by

$$\tilde{T}(v + N(T))) = T(v) \,.$$

Counting dimensions we have

$$n(T) + r(T) = \dim(V) \,.$$

Next we define the *Matrix* of $T$. Let $e_1, \dots, e_n$ be a basis in $V$ and $f_1, \dots, f_n$ be a basis in $W$. Then there exits unique scalars $\tau^i_j$ so that

$$Te_j = \tau^i_j f_i$$

or, more explicitly,

$$\begin{array}{lllllllll}
Te_1 & = \tau^1_1 f_1 & + & \tau^2_1 f_2 & + & \cdots & + & \tau^m_1 f_m \\
Te_2 & = \tau^1_2 f_1 & + & \tau^2_2 f_2 & + & \cdots & + & \tau^m_2 f_m \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
Te_n & = \tau^1_n f_1 & + & \tau^2_n f_2 & + & \cdots & + & \tau^m_n f_m \,.
\end{array}$$

Then *the matrix of* $T$ in the bases $e_1, \ldots, e_n$ of $V$ and $f_1, \ldots, f_n$ of $W$ is

$$(\tau^i_j) = \begin{pmatrix} \tau^1_1 & \tau^1_2 & \cdots & \tau^1_n \\ \tau^2_1 & \tau^2_2 & \cdots & \tau^2_n \\ \vdots & \vdots & \vdots & \vdots \\ \tau^m_1 & \tau^m_2 & \cdots & \tau^m_n \end{pmatrix}$$

Carefully note that the matrix of $T$ is the transpose of the array of coefficients in $Te_j = \tau^i_j f_i$. Much of the theory of linear operators and matrices consists in finding bases in which $(\tau^i_j)$ has some desirable form, for example upper triangular or diagonal. This amounts to basis change in $V$ and $W$. Suppose $\tilde{e}_1, \ldots \tilde{e}_n$ is a new basis in $V$ and $\tilde{f}_1, \ldots \tilde{f}_m$ is a new basis in $W$, so that

$$\begin{aligned} \tilde{e}_i &= \gamma^j_i e_j \,, & \mathcal{C} = (\gamma^j_i) \\ \tilde{f}_k &= \partial^j_i e_j \,, & \mathcal{D} = (\partial^j_i) \end{aligned}$$

These basis change matrices are invertible, so let

$$\mathcal{D}^{-1} = (\zeta^i_j)\,.$$

We now have

$$T\tilde{e}_j = \tilde{\tau}^i_j \tilde{f}_i \qquad \text{new bases}$$

and we want the relationship between $(\tau^i_j)$ and $(\tilde{\tau}^i_j)$. This is easily computed:

$$\begin{aligned} T(\tilde{e}_k) &= \tilde{\tau}^\ell_k \tilde{f}_\ell \\ T(\gamma^j_k e_j) &= \tilde{\tau}^\ell_k \partial^i_\ell f_i \\ \gamma^j_k T(e_j) &= \\ \gamma^j_k \tau^i_j f_i &= \end{aligned}$$

Since representation in the basis $\{f_1 \ldots f_m\}$ is unique,

$$\tau^i_j \gamma^j_k = \partial^i_\ell \tilde{\tau}^\ell_k$$

which can be written in matrix form as

$$(\tau^i_j)\mathcal{C} = \mathcal{D}(\tilde{\tau}^\ell_k)$$

so

$$\mathcal{D}^{-1}(\tau^i_j)\mathcal{C} = (\tilde{\tau}^\ell_k)$$

As an application of this, the Gauss reduction process gives a process whereby

$$(\tau^i_j) \rightarrow \text{Gauss Reduction} \rightarrow (\tilde{\tau}^i_j) \text{ in reduced row eschelon form.}$$

Since each action in the Gauss reduction process can be accomplished by multiplication on the left by an invertible elementary matrix $\mathcal{E}_i$ we have

$$\mathcal{E}_r \mathcal{E}_{r-1} \cdots \mathcal{E}_2 \mathcal{E}_1 (\tau^i_j) = (\tilde{\tau}^i_j)$$

and setting

$$\mathcal{D} = (\xi_r \dots \xi_1)^{-1} = \xi_1^{-1} \dots \xi_r^{-1} = (\partial_i^j)$$

we have

$$\mathcal{D}^{-1}(\tau_j^i) = (\tilde{\tau}_j^i)\,.$$

Thus there is a new basis $\tilde{f}_1 \dots \tilde{f}_m$ in $W$ where

$$\tilde{f}_i = \partial_i^j f_j$$

relative to which $T$ has a matrix in row echelon form. If $w \in W$ is expressed in this basis it is trivial to solve $Tv = w$, if it is solvable.

**Example** In the $e_i$ and $f_j$ bases $T$ is represented by

$$\begin{pmatrix} 1 & -2 & -1 \\ -1 & -1 & 3 \\ 5 & -1 & -11 \end{pmatrix}, \qquad v \text{ by } \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix}, \quad \text{and } w \text{ by } \begin{pmatrix} \eta^1 \\ \eta^2 \\ \eta^3 \end{pmatrix}$$

Then $Tv = w$ becomes

$$\begin{array}{cccccc} \xi^1 & -2\xi^2 & -\xi^3 & = & \eta^1 \\ -\xi^1 & -\xi^2 & +3\xi^3 & = & \eta^2 \\ 5\xi^1 & -\xi^2 & -11\xi^3 & = & \eta^3 \end{array}$$

Following the Gauss Reduction process we find $\mathcal{D}^{-1}$ to be

$$\mathcal{D}^{-1} = \begin{pmatrix} 1 & -2 & 0 \\ -1/3 & -1/3 & 0 \\ -2 & 3 & 1 \end{pmatrix}$$

and

$$\mathcal{D}^{-1} \begin{pmatrix} 1 & 2 & -1 \\ -1 & -1 & 3 \\ 5 & -1 & -11 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -7/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus with the new basis for $W$ and the old basis (no change) for $V$ we have

$$\begin{pmatrix} 1 & 0 & -7/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \mathcal{D}^{-1} \begin{pmatrix} \eta^1 \\ \eta^2 \\ \eta^3 \end{pmatrix} = \begin{pmatrix} \tilde{\eta}^1 \\ \tilde{\eta}^2 \\ \tilde{\eta}^3 \end{pmatrix}.$$

This is solvable if and only if $\tilde{\eta}^3 = -2\eta^1 + 3\eta^2 + 1\eta^3 = 0$ and the solution is then

$$\begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \begin{pmatrix} \tilde{\eta}^1 \\ \tilde{\eta}^2 \\ 0 \end{pmatrix} + \xi^3 \begin{pmatrix} 7/3 \\ 2/3 \\ 1 \end{pmatrix}$$

An important special case is an operator that goes from $V$ to itself. In this case, there is (usually) only one basis change involved, from old $e_i$ to new $\tilde{e}_j$, so that $\mathcal{D} = \mathcal{C}$ and the basis change rule has the form

$$(\tilde{\tau}_l^k) = \mathcal{C}^{-1}(\tau_j^i)\mathcal{C}\,.$$

It is much more difficult to arrange a "nice" matrix by selecting the new $\tilde{e}_j$ under these circumstances. The standard method is to use eigenvalues and eigenvectors, and we will now give a sketch of this process and its various pitfalls.

**Def** An *eigenvalue* and *eigenvector* of $T$ are a pair $\lambda \in \mathbb{F}$ and $v \in V$ so that

$$Tv = \lambda v\,.$$

**Def** The *eigenspace* of $\lambda$ is $N(T - \lambda I)$.

This is the space of eigenvectors for the eigenvalue $\lambda$

The eigenvalues are solutions of $\det(T - \lambda I) = 0$. This is a polynomial equation with coefficients in $\mathbb{F}$ and so the solutions may not be in $\mathbb{F}$. However, assume they are in $\mathbb{F}$, which is the case for example when $\mathbb{F} = \mathbb{C}$, the complex numbers. In this case we have $n$ not necessarily distinct eigenvalues. Eigenvectors belonging to distinct eigenvalues are linearly independent, so if we do happen to have $n$ distinct eigenvalues then the corresponding eigenvectors form a basis for the space.

If the eigenvalues are *not* distinct then the eigenvectors span the space (that is there are still $n$ linearly independent eigenvectors) if and only if, for each multiple eigenvalue, we have

$$N((T - \lambda I)^2) = N((T - \lambda I)\,.$$

(This condition is automatically satisfied for those $\lambda$ which are not repeated, that is $\lambda$ is a simple root of $\det(T - \lambda I) = 0$.)

Suppose the condition to be fulfilled, and $\{v_1, \dots, v_n\}$ to be a basis of eigenvectors with $Tv_i = \lambda_i v_i$, then with

$$v_i = \gamma_i^j e_j \qquad \text{and} \qquad \mathcal{C} = (\gamma_i^j)$$

we have

$$\mathcal{C}^{-1}(\tau_j^i)\,\mathcal{C} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

a diagonal matrix. This is the optimal situation. Under suboptimal conditions where we have all the roots $\lambda_1, \dots, \lambda_n$ of $\det(T - \lambda I) = 0$ in $\mathbb{F}$ but not enough eigenvectors, we may find "generalized eigenvectors" which are in $N(T - \lambda I)^k$ and arrange a $(\tilde{\tau}_j^i)$ consisting of blocks

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

with ones above the diagonal. This is the Jordan Canonical Form. If the roots $\lambda_1, \dots, \lambda_n$ are not in the field of scalars $\mathcal{F}$ then more complex blocks are necessary. We refer the reader to advanced books on linear algebra, for example [Malcev] for the details, which are extremely interesting.

To complete this section we would like to show how to get the change of basis equations by matrix methods. To do this it is necessary to introduce matrices whose elements are themselves vectors. For example, for a set of vectors $v_1, \dots, v_r$ we may create the row "matrix" (in an extended sense of the word matrix)

$$(v_1, v_2, \dots v_r)$$

and it would be possible to create a column of vectors also, though not natural to do so in the present context. We now define an "action" of the linear operator $T$ on the row of vectors by

$$T(v_1, v_2, \dots v_r) = (Tv_1, Tv_2, \dots Tv_r)\,.$$

We then have, for the change of bases in $V$ and $W$,

$$\begin{aligned}
(\tilde{e}_1, \dots, \tilde{e}_n) &= (e_1, \dots, e_n)\,\mathcal{C} \\
(\tilde{f}_1, \dots, \tilde{f}_m) &= (f_1, \dots, f_m)\,\mathcal{D}
\end{aligned}$$

and then

$$\begin{aligned}
T(e_1, \dots, e_n) &= (Te_1, \dots, Te_n) \\
&= (\tau_1^j f_j, \dots, \tau_n^j f_j) \\
&= (f_1, \dots, f_m)(\tau_i^j)
\end{aligned}$$

and similarly

$$T(\tilde{e}_1, \dots, \tilde{e}_n) = (\tilde{f}_1, \dots, \tilde{f}_m)(\tilde{\tau}_k^l)\,.$$

Next we put these together

$$\begin{aligned}
T(\tilde{e}_1, \dots, \tilde{e}_n) &= (\tilde{f}_1, \dots, \tilde{f}_m)(\tilde{\tau}_k^l) \\
T(e_1, \dots, e_n)\,\mathcal{C} &= (f_1, \dots, f_m)\,\mathcal{D}\,(\tilde{\tau}_k^l) \\
(f_1, \dots, f_m)(\tau_i^j)\,\mathcal{C} &= (f_1, \dots, f_m)\,\mathcal{D}\,(\tilde{\tau}_k^l)\,.
\end{aligned}$$

Since the $f_1, \dots, f_m$ are linearly independent, we must have

$$(\tau_i^j)\,\mathcal{C} = \mathcal{D}\,(\tilde{\tau}_k^l)$$

and thus

$$\mathcal{D}^{-1}(\tau_i^j)\,\mathcal{C} = (\tilde{\tau}_k^l)\,.$$

We now turn to the *conjugate operator.* Our treatment is not quite standard because we will write the matrices representing linear functionals as row matrices.

We also want to write $f(v)$ in a new way to emphasize the symmetry between $V$ and $V^*$. Henceforth, we will often write

$$< f, v > \quad \text{instead of} \quad f(v)\,.$$

If $T : V \to W$ is a linear operator, then there is a dual operator $T^* : W^* \to V^*$ defined by the equation

$$< T^* g, v >=< g, Tv > \ .$$

Let now $e_1, \ldots e_n$ be a basis in $V$ and $e^1 \ldots e^n$ be the dual basis in $V^*$. Notice that, for $f \in V^*$ and $f = \lambda_j e^j$

$$< f, e_i >=< \lambda_j e^j, e_i >= \lambda_j < e^j, e_i >= \lambda_j \delta_i^j = \lambda_i$$

so that the coefficient $\lambda_i$ of $e^i$ in $f = \lambda_j e^j$ may be found by taking $f$'s value on $e_i$. Now let $f_1, \ldots f_m$ be a basis of $W$ and $f^1, \ldots f^m$ the dual basis in $W^*$. We have

$$T^* f^j = \rho_i^j e^i$$

for some coefficients $\rho_i^j$ and we would like the relationship between the matrix $\rho_i^j$) and the matrix $(\tau_i^j)$ of $T$ in the bases $e_1, \ldots e_n$ of $V$ and $f_1, \ldots f_m$ of $W$. We have

$$\begin{aligned}
\rho_i^j &= \rho_k^j \delta_i^k = \rho_k^j < e^k, e_k >=< \rho_k^j e^k, e_k > \\
&= < T^* f^j, e_i >=< f^j, Te_i > \\
&= < f^j, \tau_i^k f_k >= \overline{\tau_i^k} < f^j, f_k >= \overline{\tau_i^k} \delta_k^j = \overline{\tau_i^j}\ .
\end{aligned}$$

Thus $T^*$ and $T$ have matrices which are conjugates of one another. (If one writes the matrices representing linear functionals as *columns,* then there would also be a transpose of the matrix of $T$ involved, but this is not convenient for us.)

We can now represent the action of $T^*$ on an element of $g \in W^*$ in the usual way; if $g = \lambda_j f_i^j$ and $T^* g = \mu_k e^k \in V$, we have

$$\mu_k e^k = T^* g = T^*(\lambda_j f^j) = \lambda_j T^*(f^j) = \lambda_j \overline{\tau_k^j} e^k$$

so

$$\mu_k = \lambda_j \overline{\tau_k^j}$$

or in matrix form

$$(\mu_1, \ldots, \mu_n) = (\lambda_1, \ldots, \lambda_m)(\overline{\tau_k^j})\,.$$

Once again, no transpose is involved because we are representing the elements of the dual spaces $W^*$ and $V^*$ as *row matrices.* If we were to use columns for this representation, then the above would be written

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = (\overline{\tau_k^j})^\top \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$$

Notice that if we have $v = \xi^i e_i \in V$ and $g = \lambda_j f^j \in W^*$ then we calculate $< g, Tv >= g(Tv)$ using matrix representatives by

$$(\overline{\lambda_1}, \dots, \overline{\lambda_m})(\tau^j_k) \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}$$

and this is exactly what one would use to calculate $< T^* g, v >= [T^*(g)](v)$;

$$\overline{(\lambda_1, \dots, \lambda_m)(\overline{\tau^j_k})} \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} .$$

(P.A.M. Dirac used a notation

$$<g|T|v>$$

for $< g, Tv >$ and $< T^* g, v >$, removing the need to distinguish $T^*$ and $T$; $Tv$ is then written as $T|v>$ and $T^* g$ as $<g|T$ thus indicating the proper one of $T^*$ and $T$ by the symbolism.)

# Chapter 3

# Tensor Products of Vector Spaces

## 3.1 Introduction

Tensor products are not at all necessary for the understanding or use of Grassmann Algebra. As we shall show, it is possible to build Grassmann Algebra using tensor products as a tool, but this is by no means necessary. If follows that the reader may completely skip this chapter if he has no interest in tensor products.

Then why do we include this chapter? There are several reasons which we discuss in the following paragraphs.

First, many people, especially in differential geometry, *like* to build Grassmann Algebra from tensor products. This is, after all, a matter of taste, and we want persons of this persuasion to feel at home in this book.

Second, for purposes of generalization in algebra, for example to modules over a commutative ring, the method has advantages, in that tensor products are well understood in that context.

Third, in differential geometry there are many contexts in which tensor products are the natural mode of expression. In such a context it is natural to want to know how tensor products and Grassmann products interact. If Grassmann products are defined as certain combinations of tensor products, the interaction becomes clear.

There is a mild use of permutations and their signs $\text{sgn}(\pi)$ in this chapter. Readers completely unfamiliar with permutations might profitably read a portion of section 4.3 on permutations (up to *increasing* permutations). Section 4.3 is independent of other material and may be read at any time.

## 3.2 Multilinear Forms and the Tensor Product

One way to build the Grassmann Algebra is by use of the tensor product, whose theory we develop in this section. We will also need some of the theory of multilinear algebra, which we construct simultaneously.

There are several ways to construct the tensor product. The one we use here is the usual method in commutative algebra.

Let $V_i$, $i = 1, \ldots, r$ be vector spaces. We construct $V_1 \otimes, V_2 \otimes, \ldots, \otimes V_r$, the tensor product of the vector spaces, as follows. We form the (very large) vector space $V(V_1, \ldots, V_r)$ with basis the elements of $V_1 \times, V_2 \times, \ldots, \times V_r$. Elements of this space may be represented by

$$\sum_{i=1}^{l} \alpha^i (v_{1i}, v_{2i}, \ldots, v_{ri})$$

where $v_{ji} \in V_j$. This space is an infinite dimensional space. We will now form a subspace $V_0(V_1, \ldots, V_r)$ generated by all elements of the form

$$(v_1, \ldots, v_{i-1}, \alpha v_i, v_{i+1}, \ldots, v_r) - \alpha(v_1, \ldots, v_r)$$
$$(v_1, .., v_{i-1}, u + w, v_{i+1}, .., v_r) - (v_1, .., v_{i-1}, u, v_{i+1}, .., v_r) - (v_1, .., v_{i-1}, w, v_{i+1}, .., v_r)$$

If one prefers a single type of generator one could define $V_0(V_1, \ldots, V_r)$ to be generated by elements of the form

$$(v_1, .., v_{i-1}, \alpha u + \beta w, v_{i+1}, .., v_r) - \alpha(v_1, .., v_{i-1}, u, v_{i+1}, .., v_r) - \beta(v_1, .., v_{i-1}, w, v_{i+1}, .., v_r)$$

The image of $(v_1, \ldots, v_r)$ in the space $V(V_1, \ldots, V_r)/V_0(V_1, \ldots, V_r)$ will be denoted by $v_1 \otimes, v_2 \otimes, \ldots, \otimes v_r$, and

$$\bigotimes_{i=1}^{r} V_i = V_1 \otimes, V_2 \otimes, \ldots, \otimes V_r = V(V_1, \ldots, V_r)/V_0(V_1, \ldots, V_r)$$

will denote the Factor Space. Because of the form of the generators of $V_0(V_1, \ldots, V_r)$ we will have

$$v_1 \otimes \ldots \otimes v_{i-1} \otimes \alpha v_i \otimes v_{i+1} \otimes \ldots \otimes, v_r = \alpha(v_1 \otimes \ldots \otimes v_r)$$

(since the first type of generator is sent to 0 by the factoring process). Also, we will have

$$\begin{aligned}(v_1 \otimes \ldots \otimes v_{i-1} \otimes (u + w) \otimes v_{i+1} \otimes \ldots \otimes v_r) &= (v_1 \otimes \ldots \otimes v_{i-1} \otimes u \otimes v_{i+1} \otimes \ldots \otimes v_r) \\ &+ (v_1 \otimes \ldots \otimes v_{i-1} \otimes v \otimes v_{i+1} \otimes \ldots \otimes v_r)\end{aligned}$$

because of the second type of generator.

The fundamental abstract principle underlying tensor products concerns their interaction with multilinear functionals, which we now define. Let $V_1, \ldots, V_r, W$ be vector spaces over a given field $\mathbb{F}$. Then

**Def** Let $F : V_1 \times \ldots \times V_r \to W$. $F$ is *multilinear* if and only if it has the following two properties:

$$\begin{aligned}
F(v_1, \ldots, v_{i-1}, \alpha v_i, v_{i+1}, \ldots, v_r) &= \alpha F(v_1, \ldots, v_r) \\
F(v_1, \ldots, v_{i-1}, u + w, v_{i+1}, \ldots, v_r) &= F(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_r) \\
&+ F(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_r)
\end{aligned}$$

The connection between multilinear functionals and tensor products is through the following basic theorem.

**Theorem** Given a multilinear functional $F : V_1 \times \ldots \times V_r \to W$ there exists a UNIQUE mapping $\tilde{F} : V_1 \otimes \ldots \otimes V_r \to W$ so that the following diagram commutes:

$$\begin{array}{ccc}
V_1 \times \ldots \times V_r & \longrightarrow & W \\
\searrow & & \nearrow \\
& V_1 \otimes \ldots \otimes V_r &
\end{array}$$

**Proof** The elements $(v_1, \ldots, v_r)$ are a basis of $V(V_1, \ldots, V_r)$ and the multilinear functional $F$ extends to a linear functional $F_1 : V(V_1, \ldots, V_r) \to W$ by defining $F_1(\,(v_1, \ldots, v_r)\,) = F(v_1, \ldots, v_r)$ and extending by linearity. We further note the $F_1$ is identically 0 on $V_0(V_1, \ldots, V_r)$. For example,

$$\begin{aligned}
F_1(\,(v_1, \ldots \alpha v_i, \ldots, v_r) - \alpha(v_1, \ldots, v_r)\,) &= F_1(\,(v_1, \ldots \alpha v_i, \ldots, v_r)\,) - \alpha F_1(\,(v_1, \ldots, v_r)\,) \\
&= F(v_1, \ldots \alpha v_i, \ldots, v_r) - \alpha F(v_1, \ldots, v_r) \\
&= \alpha F(v_1, \ldots v_i, \ldots, v_r) - \alpha F(v_1, \ldots, v_r) \\
&= 0
\end{aligned}$$

and the same for the other type of generator.

By the fundamental theorem on factor spaces we know that there is a mapping

$$\tilde{F} : V(V_1 \ldots V_r)/V_0(V_1 \ldots V_r) \to W$$

defined by

$$\tilde{F}(v_1 \otimes \ldots \otimes v_r) = F(v_1 \ldots v_r)$$

as desired.

The mapping is clearly *unique*, because elements of the form $v_1 \otimes \ldots \otimes v_r$ generate $V_1 \otimes \ldots \otimes V_r$ and the previous equation determines the value of $\tilde{F}$ on these elements.

We will now, in s somewhat mystical manner, explain the significance of the last theorem. This paragraph is not part of the logical development and may be skipped with no loss of continuity. It is intended for psychological orientation only. The purpose of the tensor product is to give a product of vectors subject *only* to the bilinear restrictions

$$\begin{aligned}
v_1 \otimes \ldots \otimes \alpha v_i \otimes \ldots \otimes, v_r &= \alpha(v_1 \otimes \ldots \otimes v_r) \\
v_1 \otimes \ldots \otimes (u + w) \otimes \ldots \otimes v_r &= v_1 \otimes \ldots \otimes u \otimes \ldots \otimes v_r \\
&+ v_1 \otimes \ldots \otimes w \otimes \ldots \otimes v_r.
\end{aligned}$$

These restrictions are essential for a meaningful product of vectors, and we want *no other* algebraic rules than these. For example, we do not want $v_1 \otimes \ldots \otimes v_r = 0$ unless some $v_i = 0$. If such a thing were to happen (as indeed it may when the vector spaces are generalized to modules, we might say we have "collapse" of the product. But we want no collapse except that which occurs through the action of the above two laws (which means *none whatever* for vector spaces.) The theorem is supposed to guarantee this lack of collapse; if an element is 0, then *no* multilinear functional applied to it is non–zero. This insures that the tensor product is "big enough" to accommodate the action of it all multilinear functionals. We say then that the tensor product has a *universal* property with respect to multilinear functionals. In situations like vector spaces this whole matter can be simply controlled by finding a basis, but in the relatively simple generalization to modules the basis method is generally not available, and the universal construction is the best tool we have. Naturally we will show the equivalence of the two approaches in our case of vector spaces. We return now to the systematic development.

Since $V_1 \otimes \ldots \otimes V_r$ is a vector space we will eventually have to deal with its dual space. Here we will define an interaction between $V_1 \otimes \ldots \otimes V_r$ and $V_1^* \otimes \ldots \otimes V_r^*$ which will eventually be used to show that one is the dual of the other. We begin with an action of $V(V_1^* \ldots V_r^*)$ on $V(V_1 \ldots V_r)$ defined on basis elements $(f^1, \ldots, f^r)$, $f^i \in V_i^*$ and $(v_1, \ldots, v_r)$, $v_i \in V_i$ by

**Def** $$(f^1, \ldots, f^r)(v_1, \ldots, v_r) = f^1(v_1)f^2(v_2)\ldots f^r(v_r).$$

The action is extended from the basis elements to both spaces by linearity:

$$\begin{aligned}
(\sum_i \alpha^i (f_i^1, \ldots, f_i^r)(\sum_j \beta_j (v_1^j, \ldots, v_r^j) &= \sum_{ij} \alpha^i \beta_j (f_i^1, \ldots, f_i^r)(v_1^j, \ldots, v_r^j) \\
&= \sum_{ij} \alpha^i \beta_j f_i^1(v_1^j)\ldots f_i^r(v_r^j)
\end{aligned}$$

we now notice that the action of any element of $V(V_1^* \ldots V_r^*)$ is multilinear on $V(V_1 \ldots V_r)$. It suffices to check this on a basis:

$$\begin{aligned}
(\sum_i \alpha^i (f_i^1, \ldots, f_i^r)(\sum_j \beta_j (v_1^j, \ldots, v_r^j) &= \sum_{ij} \alpha^i \beta_j (f_i^1, \ldots, f_i^r)(v_1^j, \ldots, v_r^j) \\
&= \sum_{ij} \alpha^i \beta_j f_i^1(v_1^j)\ldots f_i^r(v_r^j)
\end{aligned}$$

Thus there is a mapping, again denoted by $(f^1, \ldots, f^r)$ from $V_1 \otimes \ldots \otimes V_r$ to the field given by

$$(f^1, \ldots, f^r)(v_1 \otimes \ldots \otimes v_r) = f^1(v_1)\ldots f^r(v_r).$$

Next, we note that the mapping

$$F_{v_1 \otimes \ldots \otimes v_r}(f^1 \otimes \ldots \otimes f^r)$$

is bilinear; the proof being repeat of the previous one, so that there is a mapping, again denoted by $F_{v_1\otimes\ldots\otimes v_r}$, of $V_1^*\otimes\ldots\otimes V_r^*$ to the field given by

$$F_{v_1\otimes\ldots\otimes v_r}(f^1\otimes\ldots\otimes f^r)=f^1(v_1)\ldots f^r(v_r).$$

We now regard this as a pairing between elements of a vector space and its dual. That is, we regard $F_{v_1\otimes\ldots\otimes v_r}$ to be an element of $(V_1^*\otimes\ldots\otimes V_r^*)^*$, and define the interaction by

**Def** $$(f^1\otimes\ldots\otimes f^r)(v_1\otimes\ldots\otimes v_r)=f^1(v_1)\ldots f^r(v_r)$$

which the above analysis shows is well defined.

We now ask the natural question, since $V_1\otimes\ldots\otimes V_r$ is a vector space what is its dimension and what is a basis for it? First, it is clear that $v_1\otimes\ldots\otimes v_r$, $v_i\in V_i$ generate $V_1\otimes\ldots\otimes V_r$, and the, since $v_i=\alpha_i^j e_{ij}$ (where $e_{i1},e_{i2},\ldots,e_{in_i}$ is a basis for $V_i$,) we have $e_{1,j_1}\otimes\ldots\otimes e_{r,j_r}$, $1\le j_i\le n_i=\dim V_i$ is also a set of generators for $V_1\otimes\ldots\otimes V_r$. So everything turns on the linear independence of all elements of the form

$$e_{1,j_1}\otimes\ldots\otimes e_{r,j_r}\qquad\begin{cases}e_{i1},e_{i2},\ldots,e_{in_i} & \text{basis of } V_i\\ 1\le j_i\le n_i\end{cases}$$

To establish this we assume a linear dependence

$$\sum_{j_1,j_2\ldots j_r}\alpha^{j_1,j_2\ldots j_r}e_{1,j_1}\otimes\ldots\otimes e_{r,j_r}=0$$

and show the coefficients are 0. To this end, we consider the multilinear functional $F$ defined by

$$F(v_1\ldots v_r)=e^{1i_1}(v_1)e^{2i_2}(v_2)\ldots e^{ri_r}(v_r)$$

where $e^{ji_j}$ is selected from the dual basis $e^{j1},e^{j1},\ldots,e^{jn_j}$ of $V_j^*$. By the basic theorem there is an

$$\tilde F; V_1\otimes\ldots\otimes V_r\to\text{Field}$$

satisfying

$$\tilde F(v_1\otimes\ldots\otimes v_r)=F(v_1,\ldots,v_r)\quad\text{for all } v_i\in V_i.$$

Applying the linear transformation $\tilde F$ to the above supposed linear dependence, we have

$$\sum_{j_1,j_2\ldots j_r}\alpha^{j_1,j_2\ldots j_r}\tilde F(e_{1j_1}\otimes\ldots\otimes e_{rj_r})=0$$

$$\sum_{j_1,j_2\ldots j_r}\alpha^{j_1,j_2\ldots j_r}e^{1i_1}(e_{1j_1})e^{2i_2}(e_{2j_2})\ldots e^{ri_1}(e_{rj_1})=0$$

$$\alpha^{i_1,i_2\ldots i_r}=0$$

because in the penultimate equation all terms will be 0 except the single term in which $j_1=i_1$, $j_2=i_2$, $\ldots j_r=i_r$, by the definition of the dual basis. This proves the linear independence of the indicated terms.

From this it is clear that

$$\dim(V_1 \otimes \ldots \otimes V_r) = \dim(V_1) \cdot \dim(V_2) \cdots \dim(V_r)$$

and that the elements

$$e_{1j_1} \otimes \ldots \otimes e_{rj_r}$$

form a basis of $V_1 \otimes \ldots \otimes V_r$.

## 3.3 Grassmann Products from Tensor Products

The intent of this section is to derive the basic laws of Grassmann algebra from the definition of Grassmann products in terms of tensor products. We start with a vector space $V$ and form the $r^{th}$ tensor power

$$\underbrace{V^r = V \otimes \ldots \otimes V}_{r \text{ terms}}$$

We will now define a projection operator $\Pi : V^r \to V^r$ by the formula

$$\Pi(v_1 \otimes \ldots \otimes v_r) = \frac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi)\, v_{\pi(1)} \otimes v_{\pi(2)} \otimes \ldots \otimes v_{\pi(r)}$$

Here, $\mathcal{S}_r$ is the symmetric group of all permutations of $r$ letters and $\operatorname{sgn}(\pi) = +1$ or $-1$ according to whether $\pi$ is an even or odd permutations. (Readers unfamiliar with these concepts may read the initial part of section 3.2 where the exposition has been crafted to be readable at this point. Read up to the beginning of *increasing* permutations and then return to this point.)

We now give some examples

$$\begin{aligned}
\Pi(v_1 \otimes v_2) &= \tfrac{1}{2}(v_1 \otimes v_2 - v_2 \otimes v_1) \\
\Pi(v_2 \otimes v_1) &= \tfrac{1}{2}(v_2 \otimes v_1 - v_1 \otimes v_2) = -\Pi(v_1 \otimes v_2) \\
\Pi\Pi(v_1 \otimes v_2) &= \tfrac{1}{2}\Pi(v_1 \otimes v_2) - \tfrac{1}{2}\Pi(v_2 \otimes v_1) \\
&= \tfrac{1}{2}\Pi(v_1 \otimes v_2) + \tfrac{1}{2}\Pi(v_1 \otimes v_2) = \Pi(v_1 \otimes v_2)
\end{aligned}$$

(This property, $\Pi^2 = \Pi$, is the reason $\Pi$ was referred to as a *projection*). Continuing now with products of three vectors

$$\begin{aligned}
\Pi(v_1 \otimes v_2 \otimes v_3) &= \tfrac{1}{6}(v_1 \otimes v_2 \otimes v_3 - v_1 \otimes v_3 \otimes v_2 - v_2 \otimes v_1 \otimes v_3 \\
&\quad + v_2 \otimes v_3 \otimes v_1 - v_3 \otimes v_2 \otimes v_1 + v_3 \otimes v_1 \otimes v_2) \\
\Pi(v_2 \otimes v_3 \otimes v_5) &= \tfrac{1}{6}(v_2 \otimes v_3 \otimes v_5 - v_2 \otimes v_5 \otimes v_3 - v_3 \otimes v_2 \otimes v_5 \\
&\quad + v_3 \otimes v_5 \otimes v_2 - v_5 \otimes v_3 \otimes v_2 + v_5 \otimes v_2 \otimes v_3)
\end{aligned}$$

This last example has been included to make an important point; the permutations act on the **slot index** (the position of the element in the row of tensored vectors) and not the index to the element that happens to be in the slot. Thus, for a $\sigma \in \mathcal{S}_3$, we have

$$\begin{aligned}
\Pi(v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes v_{\sigma(3)}) &= \tfrac{1}{6}\big(\, v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes v_{\sigma(3)} - v_{\sigma(1)} \otimes v_{\sigma(3)} \otimes v_{\sigma(2)} \\
&- \quad v_{\sigma(2)} \otimes v_{\sigma(1)} \otimes v_{\sigma(3)} + v_{\sigma(2)} \otimes v_{\sigma(3)} \otimes v_{\sigma(1)} \\
&- \quad v_{\sigma(3)} \otimes v_{\sigma(2)} \otimes v_{\sigma(1)} + v_{\sigma(3)} \otimes v_{\sigma(1)} \otimes v_{\sigma(2)}\big) \\
&= \tfrac{1}{6} \sum_{\pi \in \mathcal{S}_3} \operatorname{sgn}(\pi)\, v_{\sigma(\pi(1))} \otimes v_{\sigma(\pi(2))} \otimes v_{\sigma(\pi(3))}
\end{aligned}$$

where is the last line $\pi \in \mathcal{S}_3$. *Note the order of the permutations in the indices.*

We are now in a position to prove that in general $\Pi^2 = \Pi$, so that $\Pi$ is a projection. Recall that if $\sigma, \pi \in \mathcal{S}_r$ then $\text{sgn}(\sigma\pi) = \text{sgn}(\sigma) \cdot \text{sgn}(\pi)$ and that $\text{sgn}(\sigma) = \pm$ so $[\text{sgn}(\sigma)]^2 = +1$. Also note that, since $\mathcal{S}_r$ is a group, if $\sigma$ is a fixed element of $\mathcal{S}_r$ and $\pi$ runs through all the elements of $\mathcal{S}_r$ once each, then $\sigma\pi$ runs through all the elements of $\mathcal{S}_r$ once each. Keeping all this in mind, we first have, for $\sigma, \pi \in \mathcal{S}_r$

$$\begin{aligned}
\Pi(v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}) &= \tfrac{1}{r!}\sum_{\pi} \text{sgn}(\pi)\, v_{\sigma\pi(1)} \otimes \ldots \otimes v_{\sigma\pi(r)} \\
&= \tfrac{1}{r!}[\text{sgn}(\sigma)]^2 \sum_{\pi} \text{sgn}(\pi)\, v_{\sigma\pi(1)} \otimes \ldots \otimes v_{\sigma\pi(r)} \\
&= \text{sgn}(\sigma)\tfrac{1}{r!}\sum_{\pi} \text{sgn}(\sigma\pi)\, v_{\sigma\pi(1)} \otimes \ldots \otimes v_{\sigma\pi(r)} \\
&= \text{sgn}(\sigma)\Pi(v_1 \otimes \ldots \otimes v_r).
\end{aligned}$$

Next we see $\Pi^2 = \Pi$, for

$$\begin{aligned}
\Pi\Pi(v_1 \otimes \ldots \otimes v_r) &= \Pi(\tfrac{1}{r!}\sum_{\sigma \in \mathcal{S}_r} \text{sgn}(\sigma) v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}) \\
&= \tfrac{1}{r!}\sum_{\sigma} \text{sgn}(\sigma)\Pi(v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}) \\
&= \tfrac{1}{r!}\sum_{\sigma} [\text{sgn}(\sigma)]^2 \Pi(v_1 \otimes \ldots \otimes v_r) \\
&= \Pi(v_1 \otimes \ldots \otimes v_r)
\end{aligned}$$

We now define $\Lambda^r(V)$, the $r^{th}$ exterior power of $V$.

**Def**
$$\Lambda^r(V) = \big\{ A \in \bigotimes_{i=1}^{r} V \mid \Pi A = A \big\}$$

In fact, this is simply the *range* of $\Pi$, since $\Pi$ is a projection.

Since $\Lambda^r(V)$ is the image of $\bigotimes_{i=1}^{r} V$ under $\Pi$, $\Lambda^r(V)$ is generated by elements of the form $\Pi(v_1 \otimes \ldots \otimes v_r)$. However, before we can go on a technical consideration intrudes.

It is most important for us that we be able to treat the Grassmann Algebras, which we are about to define, on a vector space and it's dual space in a wholly symmetric manner. In general, this is not completely possible, because it requires the introduction of $\sqrt{1/r!}$ into the formulas at this point and this quantity may not exist in the base field being used. For most of the book we are going to insist that the base field contain this quantity, but for the rest of *this* section we are going to compromise in order to define the Grassmann Algebra over any field. Thus we introduce a function $S(\alpha)$ where $S : \mathbb{F} \to \mathbb{F}$ from the field into itself satisfying

$$S(\alpha) \cdot S(\beta) = S(\alpha\beta)$$

The two most common choices for $S$ are

$$S(\alpha) = \alpha$$

or

$$S(\alpha) \quad = \quad \sqrt{\alpha}$$

and after the end of this section we will definitely settle on the second alternative. The first alternative is useful because it is applicable to any field, but as we mentioned it makes it impossible to maintain a complete symmetry between a vector space and its dual.

Having, in any specific circumstance, chosen an $S$-function, we can now proceed to define the wedge or exterior product.

**Def**

$$\begin{aligned} v_1 \wedge \ldots \wedge v_r &= S(r!)\Pi(v_1 \otimes \ldots \otimes v_r) \\ &= \frac{S(r!)}{r!} \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)} \end{aligned}$$

These elements clearly generate $\Lambda^r(V)$.

It is worth noting at this point that if the number of elements in the product $v_1 \wedge \ldots \wedge v_r$ exceeds the characteristic of the field, the exterior product will be 0 for most choices of $S$.

Next note that

$$\begin{aligned} v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(r)} &= S(r!)\Pi(v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}) \\ &= \operatorname{sgn}(\sigma) S(r!)\Pi(v_1 \otimes \ldots \otimes v_r) \\ &= \operatorname{sgn}(\sigma) v_1 \wedge \ldots \wedge v_r \end{aligned}$$

We now wish to define a product

$$\wedge : \Lambda^r(V) \times \Lambda^s(V) \to \Lambda^{r+s}(V).$$

We do this as follows

**Def** For $f \in \Lambda^r(V)$ and $g \in \Lambda^s(V)$

$$f \wedge g = \frac{S((r+s)!)}{S(r!s!)} \Pi(f \otimes g)$$

To clarify this definition, we prove first that for $f \in \bigotimes_{i=1}^r(V)$ and $g \in \bigotimes_{i=1}^s(V)$ we have

$$\Pi(f \otimes g) = \Pi(\Pi f \otimes g) = \Pi(f \otimes \Pi g) = \Pi(\Pi f \otimes \Pi g).$$

It suffices to prove the first equality where $f$ and $g$ are generators. Let $f = v_1 \otimes \ldots \otimes v_r$, $g = v_{r+1} \otimes \ldots \otimes v_{r+s}$. Then

$$\Pi(\Pi f \otimes g) = \tfrac{1}{r!}\Pi\big[\big(\sum_{\pi \in \mathcal{S}_r} v_{\pi(1)} \otimes \ldots \otimes v_{\pi(r)}\big) \otimes v_{\pi(r+1)} \otimes \ldots \otimes v_{\pi(r+s)}\big].$$

Now define $\tilde{\pi} \in \mathcal{S}_{r+s}$ for $\pi \in \mathcal{S}_r$ by the equations

$$\tilde{\pi}(i) = \begin{cases} \pi(i) & \text{if } i = 1, \dots, r \\ i & \text{if } i = r+1, \dots, r+s \end{cases}$$

Then, clearly, $\operatorname{sgn} \tilde{\pi} = \operatorname{sgn} \pi$, and

$$\begin{aligned}
\Pi(\Pi f \otimes g) &= \tfrac{1}{r!}\Pi\Big\{\sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\tilde{\pi})\, v_{\tilde{\pi}(1)} \otimes \dots \otimes v_{\tilde{\pi}(r)} \otimes v_{\tilde{\pi}(r+1)} \otimes \dots \otimes v_{\tilde{\pi}(r+s)}\Big\} \\
&= \tfrac{1}{r!}\tfrac{1}{(r+s)!} \sum_{\pi \in \mathcal{S}_r} \sum_{\sigma \in \mathcal{S}_{r+s}} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tilde{\pi})\, v_{\tilde{\pi}\sigma(1)} \otimes \dots \otimes v_{\tilde{\pi}\sigma(r)} \otimes \dots \otimes v_{\tilde{\pi}\sigma(r+s)} \\
&= \tfrac{1}{r!}\tfrac{1}{(r+s)!} \sum_{\pi \in \mathcal{S}_r} \sum_{\sigma \in \mathcal{S}_{r+s}} \operatorname{sgn}(\tilde{\pi}\sigma) v_{\tilde{\pi}\sigma(1)} \otimes \dots \otimes v_{\tilde{\pi}\sigma(r+s)} \\
&= \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \Pi(v_1 \otimes \dots \otimes v_{r+s}) \\
&= \Pi(v_1 \otimes \dots \otimes v_{r+s}) \\
&= \Pi(f \otimes g)
\end{aligned}$$

The equality $\Pi(f \otimes g) = \Pi(f \otimes \Pi g)$ is proved in the same way and the last equality follows from the first two.

We may now compute $f \wedge g$ when $f$ and $g$ are generators of $\Lambda^r(V)$ and $\Lambda^s(V)$. For we have

$$\begin{aligned}
f = v_1 \wedge \dots \wedge v_r &= S(r!)\Pi(v_1 \otimes \dots \otimes v_r) \\
g = v_{r+1} \wedge \dots \wedge v_{r+s} &= s(s!)\Pi(v_{1+1} \otimes \dots \otimes v_{r+s})
\end{aligned}$$

and thus

$$\begin{aligned}
f \wedge g &= \frac{S((r+s)!)}{S(r!) \cdot S(s!)}\Pi(S(r!)\Pi(v_1 \otimes \dots \otimes v_r) \otimes S(s!)\Pi(v_{r+1} \otimes \dots \otimes v_{r+s}) \\
&= S((r+s)!)\Pi((v_1 \otimes \dots \otimes v_r) \otimes (v_{r+1} \otimes \dots \otimes v_{r+s}) \\
&= S((r+s)!)\Pi(v_1 \otimes \dots \otimes v_r \otimes v_{r+1} \otimes \dots \otimes v_{r+s}) \\
&= v_1 \wedge \dots \wedge v_r \wedge v_{r+1} \wedge \dots \wedge v_{r+s}
\end{aligned}$$

From the above we separate out specifically

**Corollary** For elements $v_1 \wedge \dots \wedge v_r \in \Lambda^r(V)$ and $v_{r+1} \wedge \dots \wedge v_{r+s} \in \Lambda^s(V)$

$$(v_1 \wedge \dots \wedge v_r) \wedge (v_{r+1} \wedge \dots \wedge v_{r+s}) = v_1 \wedge \dots \wedge v_r \wedge v_{r+1} \wedge \dots \wedge v_{r+s}$$

Based on the above, we are now in a position to prove the basic Axioms of Grassmann Algebra. We refer to these results as Axioms because they can be used to build the remainder of the theory of Grassmann Algebra on an axiomatic basis with no further use of the tensor product. This is more elegant than falling back on the tensor definition from time to time.

**Axiom 1** The Grassmann product is associative.

**Proof** Let $f = v_1 \wedge \ldots \wedge v_r \in \Lambda^r(V)$, $g = v_{r+1} \wedge \ldots \wedge v_{r+s} \in \Lambda^s(V)$ and $h = v_{r+s+1} \wedge \ldots \wedge v_{r+s+t} \in \Lambda^t(V)$. Then, utilizing the previous equation,

$$\begin{aligned}(f \wedge g) \wedge h &= \big[(v_1 \wedge \ldots \wedge v_r) \wedge (v_{r+1} \wedge \ldots \wedge v_{r+s})\big] \wedge \big[v_{r+s+1} \wedge \ldots \wedge v_{r+s+t}\big] \\ &= \big[v_1 \wedge \ldots \wedge v_r \wedge v_{r+1} \wedge \ldots \wedge v_{r+s}\big] \wedge \big[v_{r+s+1} \wedge \ldots \wedge v_{r+s+t}\big] \\ &= v_1 \wedge \ldots \wedge v_r \wedge v_{r+1} \wedge \ldots \wedge v_{r+s} \wedge v_{r+s+1} \wedge \ldots \wedge v_{r+s+t}\end{aligned}$$

and clearly we will also have $f \wedge (g \wedge h)$ equal to the same value, so that

$$(f \wedge g) \wedge h = f \wedge g(\wedge h)$$

and the Grassmann Algebra is associative on its generators, and hence associative.

**Axiom 2** The Grassmann Product is multilinear:

$$v_1 \wedge \ldots \wedge (\alpha^1 u_1 + \alpha^2 u_2 \wedge \ldots \wedge v_r = \alpha^1 (v_1 \wedge \ldots \wedge u_1 \wedge \ldots \wedge) + \alpha^2 (v_2 \wedge \ldots \wedge u_2 \wedge \ldots \wedge)$$

**Proof** This is trivially true because the tensor product is multilinear and $\Pi$ is a linear transformation.

**Axiom 3** For any $v \in V$, $v \wedge v = 0$

**Proof**

$$\begin{aligned}v_1 \wedge v_2 &= S(2!)\Pi(v_1 \otimes v_2) \\ &= \frac{S(2!)}{2!} \sum_{\pi \in \mathcal{S}_2} \operatorname{sgn}(\pi) v_{\pi(1)} \otimes v_{\pi(1)} \\ &= \frac{S(2!)}{2!} (v_1 \otimes v_2 - v_2 \otimes v_1)\end{aligned}$$

Hence, substituting $v_1 = v_2 = v$,

$$\begin{aligned}v \wedge v &= \frac{S(2!)}{2!} (v \otimes v - v \otimes v) \\ &= 0\end{aligned}$$

We next formulate Axiom 4a. Axiom 4 comes in two equivalent forms, and in this section we introduce only 4a. We will discuss the equivalent form 4b later.

To formulate Axiom 4a, we must introduce the concept of an alternating multilinear functional, which is a multilinear functional satisfying one additional property.

**Def** A multilinear functional is *alternating* if and only if

$$F(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, v_{i+2}, \ldots, v_r) = -F(v_1, \ldots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \ldots, v_r)$$

for $i = 1, \ldots, r-1$.

**Note** The condition for a multilinear functional to be alternating is that interchanging two adjacent arguments changes the sign of the functional. In fact, it is easily proved that interchanging any two arguments, adjacent or not, changes the sign of the functional. Since we go into this matter in detail in section 4.3 we will not discuss it further here, but the reader is urged to try a couple of examples to see how it works.

**Axiom 4a** Given an alternating multilinear functional $F(v_1, \ldots, v_r)$, there exists a linear functional $\tilde{F} : \Lambda^r(V) \to W$ so that the following diagram commutes:

$$\begin{array}{ccc} V_1 \times \ldots \times V_r & \xrightarrow{F} & W \\ \Pi \searrow & & \nearrow \tilde{F} \\ & \Lambda^r(V) & \end{array}$$

which is to say that $\tilde{F}(v_1 \wedge \ldots \wedge v_r) = F(v_1, \ldots, v_r)$ for every $v_1, \ldots, v_r \in V$.

**Proof** This is easy. Since $F$ is multilinear there is a multilinear functional $F_1 : V \otimes \ldots \otimes V \to W$ satisfying $F_1(v_1 \otimes \ldots \otimes v_r) = F(v_1, \ldots, v_r)$. Since $\Lambda^r(V) \subseteq V \otimes \ldots \otimes V$, there is a restriction of $F_1$ to $\tilde{F} : \Lambda^r(V) \to W$ which naturally remains linear. The reader will easily convince himself that for any permutation the alternating property implies that

$$F(v_{\sigma(1)}, \ldots, v_{\sigma(r)}) = \text{sgn}(\sigma)\, F(v_1, \ldots, v_r)$$

since $\text{sgn}(\sigma)$ is equal to $-1$ raised to a power equal to the number of adjacent interchanges necessary to restore the sequence $\sigma(1), \sigma(2), \ldots, \sigma(r)$ to the sequence $1, \ldots, r$ (This is handled more exhaustively in section 3.2) Hence

$$\begin{aligned}
\tilde{F}(v_1 \wedge \ldots \wedge v_r) &= F_1(v_1 \otimes \ldots \otimes v_r) \\
&= F_1\big(\tfrac{1}{r!} \sum_{\sigma \in \mathcal{F}_r} \text{sgn}(\sigma) v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}\big) \\
&= \tfrac{1}{r!} \sum_{\sigma \in \mathcal{F}_r} \text{sgn}(\sigma) F_1(v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}) \\
&= \tfrac{1}{r!} \sum_{\sigma \in \mathcal{F}_r} \text{sgn}(\sigma) F(v_{\sigma(1)}, \ldots, v_{\sigma(r)}) \\
&= \tfrac{1}{r!} \sum_{\sigma \in \mathcal{F}_r} \text{sgn}(\sigma)\text{sgn}(\sigma) F(v_1, \ldots, v_r) \\
&= \tfrac{1}{r!} \sum_{\sigma \in \mathcal{F}_r} F(v_1, \ldots, v_r) \\
&= F(v_1, \ldots, v_r)
\end{aligned}$$

as desired.

# Chapter 4

# Grassmann Algebra on the Vector Space $V$

## 4.1 Introduction

In this chapter we develop the elementary theory of Grassmann Algebra on an axiomatic basis. The axioms we will use were proved as theorems in Chapter 2 on the basis of the tensor product, but we do not wish tensor products to play a role in the systematic development of Grassmann Algebra in this chapter, and therefore base our development on the axioms. This has the effect of breaking the theory into modular units.

As we will see, determinants appear naturally as the coefficients in Grassmann Algebra, and this accounts for the tendency of determinants to appear sporadically throughout mathematics. As a general rule, the presence of a determinant signals an underlying Grassmann Algebra which is seldom exploited to its full potential.

There is an alternate way of realizing the Grassmann Algebra by building it on Cartesian products in analogy to the way tensor products are built as factor spaces of the vector space generated by $V_1, \ldots, V_r$. There are some cumbersome features to this method but many people like to do it this way and it is important that it can be done, so we will lay out this construction in detail in the last section of this chapter. This may be read immediately after section 4.2 .

## 4.2 Axioms

Let $V$ be a vector space over a field $\mathbb{F}$. The Grassmann algebra of $V$, denoted by $\Lambda(V)$ or $\Lambda V$, is the linear span (over $\mathbb{F}$) of products

$$v_1 \wedge v_2 \wedge \ldots \wedge v_r$$

where $v_i \in V$. Each term in a member of $\Lambda V$ has a *degree,* which is the number of vectors in the product: $\deg(v_1 \wedge v_2 \wedge \ldots \wedge v_r) = r$ We agree that, by definition, the elements of the field $F$ will have degree equal to 0. We subject the product to the following laws or Axioms.

**Axiom 1** The product is associative: for $1 < r < s < t$

$$\begin{aligned}\big((v_1 \wedge \ldots \wedge v_r) \wedge (v_{r+1} \wedge \ldots \wedge v_s)\big) \wedge (v_{s+1} \wedge \ldots \wedge v_t)\\ = (v_1 \wedge \ldots \wedge v_r) \wedge \big((v_{r+1} \wedge \ldots \wedge v_s) \wedge (v_{s+1} \wedge \ldots \wedge v_t)\big)\end{aligned}$$

so that each of the above terms may be written

$$v_1 \wedge \ldots \wedge v_r \wedge v_{r+1} \wedge \ldots \wedge v_s \wedge v_{s+1} \wedge \ldots \wedge v_t.$$

**Axiom 2** The product is bilinear:

$$\begin{aligned}v_1 \wedge \ldots \wedge v_{r-1} \wedge (\alpha^1 u_1 + \alpha^2 u_2) \wedge v_{r+1} \wedge \ldots \wedge v_s\\ = \alpha^1 (v_1 \wedge \ldots \wedge v_{r-1} \wedge u_1 \wedge v_{r+1} \wedge \ldots \wedge v_s)\\ + \alpha^2 (v_1 \wedge \ldots \wedge v_{r-1} \wedge u_2 \wedge v_{r+1} \wedge \ldots \wedge v_s)\end{aligned}$$

for any $r, s$ with $1 < r < s$.

**Axiom 3** The product is very nilpotent:

$$v \wedge v = 0 \qquad \text{for all } v \in V$$

As a consequence of Axiom 3 and bilinearity, we have

$$\begin{aligned}(v + w) \wedge (v + w) &= v \wedge v + v \wedge w + w \wedge v + w \wedge w\\ 0 &= 0 + w \wedge w + w \wedge v + 0\end{aligned}$$

so that

$$v \wedge w = -w \wedge v \qquad \text{for all } v, w \in V.$$

We will refer to this as the anti-commutativity property.

We remark that Axiom 3 is preferable to the equation $v \wedge w = -w \wedge v$ as an axiom because, for a field of characteristic 2, the axiom implies the equation but not conversely. For any characteristic other than 2, the axiom and the equation $v \wedge w = -w \wedge v$ are equivalent.

The terms of the Grassmann Algebra can always be rewritten into sums of terms of homogeneous degree; for example

$$v_1 + 2v_1 \wedge v_2 + v_2 \wedge v_3 + v_4 - 3v_1 \wedge v_3 \wedge v_4 + 7$$

can be rewritten as

$$7 + (v_1 + v_4) + (2v_1 \wedge v_2 + v_2 \wedge v_3) + (-3v_1 \wedge v_3 \wedge v_4)$$

with terms of degree 0,1,2 and 3. The Grassmann Algebra is a vector space and the set of products of degree r form a subspace $\Lambda^r(V)$.

Up to now, there has been nothing to prevent the complete or partial "collapse" of the Grassmann Algebra; for example the axioms would all be true if $v \wedge w = 0$ for all $v, w \in V$, or this might be true for some $v, w$ and not others. We wish this phenomenon to be reduced to a minimum. There are two equivalent ways to do this. The first form of Axiom 4 is methodologically preferable because it does not involve a basis, whereas the second form is psychologically preferable because the content is clear. We will eventually prove the two are equivalent.

To formulate Axiom 4 in a basis free way, we define the concept of an alternating multilinear functional. (We did this in chapter 2 also, but we want to keep this chapter independent of chapter 2.)

**Def** Let V and W be vector spaces. A function

$$G : \underbrace{V \otimes V \otimes \ldots \otimes V}_{s \text{ factors}} \to W$$

is an alternating multilinear functional if and only if

$$\begin{aligned} G(v_1, \ldots, v_{r-1}(\alpha^1 u_1 + \alpha^2 u_2), v_{r+1}, \ldots, v_s) &= \alpha^1 G(v_1, \ldots, v_{r-1}, u_1, v_{r+1}, \ldots, v_s) \\ &+ \alpha^2 G(v_1, \ldots, v_{r-1}, u_2, v_{r+1}, \ldots, v_s) \\ G(v_1, \ldots, v_r, v_{r+1}, \ldots, v_s) &= -G(v_1, \ldots, v_{r+1}, v_r, \ldots, v_s). \end{aligned}$$

We may now formulate Axiom 4 quite simply.

**Axiom 4a** Let $G(v_1, \ldots, v_r)$ be an alternating multilinear function from $V \otimes \ldots \otimes V$ to $W$. Let $\Phi$ be the map $\Phi : V \otimes \ldots \otimes V \to \Lambda^r(V)$ given by

$$\Phi(v_1, \ldots, v_r) = v_1 \wedge \ldots \wedge v_r$$

(which is alternating and multilinear by Axioms 2 and 3.) Then there exists a unique map $\tilde{G} \colon \Lambda^r(V) \to W$ so that the following diagram commutes:

$$\begin{array}{ccc} V_1 \times \ldots \times V_n & \xrightarrow{G} & W \\ {\scriptstyle \Phi}\searrow & & \nearrow{\scriptstyle \tilde{G}} \\ & \Lambda^r(V) & \end{array}$$

The commutativity of the the diagram says that $G = \tilde{G} \circ \Phi$.

In section 3.2 we explained how a similar condition forced tensor products to be "as large as possible." Axiom 4 insured the the Grassmann Algebras is as large as possible, consistent with Axioms 1,2,3. This same end can be achieved more simply but less elegantly by introducing a basis into V. Let $e_1, \ldots, e_n$ be a basis for $V$. Then we may achieve the same end by the axiom

**Axiom 4b** The set of all products

$$e_{i_1} \wedge \ldots \wedge e_{i_r} \qquad \begin{cases} 1 \le r \le n \\ i_1 < i_2 \ldots < i_r \end{cases}$$

is linearly independent.

The equivalence of Axioms 4a and 4b is not immediately obvious but will be demonstrated in due course.

We now address the point that the expressions of Axiom 4b form a basis for $\Lambda(V)$. The question of involves only the spanning property of a basis; line independence is guaranteed by Axiom 4b. It is sufficient to show that expressions of the form

$$e_{i_1} \wedge \ldots \wedge e_{i_r} \qquad r \text{ fixed}, \qquad i_1 < i_2 \ldots < i_r$$

span $\Lambda^r(V)$. To see this, let $v_1, \ldots, v_r$ be given in terms of the basis by

$$v_j = \alpha^i_j e_i$$

Then by Axioms 1–3,

$$v_1 \wedge \ldots \wedge v_r = \alpha^{i_1}_1 \alpha^{i_2}_2 \ldots \alpha^{i_r}_r \, e_{i_1} \wedge \ldots \wedge e_{i_r}.$$

The product on the right hand side may be rearranged by the anticommutative property so that the indices increase in each term (with possible sign changes). This shows that $v_1 \wedge \ldots \wedge v_r$ is a linear combination of the terms of the specified form, proving that these terms span $\Lambda^r(V)$

By taking the direct sum of the vector spaces $\Lambda^r(V)$ we get an algebra. If $A \in \Lambda^r(V)$ and $B \in \Lambda^s(V)$ then $A \cdot B \in \Lambda^{r+s}(V)$. To complete the Grassmann Algebra, however, we must put in the basement. We define $\Lambda^0(V)$ to be the Field of constants, and we define for $\alpha \in \Lambda^0(V)$ and $A \in \Lambda^r(V)$, the Grassmann product $\alpha \wedge A$ to be be simply scalar multiplicaton $\alpha A$.

The reader may be bothered by the fact that multiplication $\alpha \wedge \beta$ for $\alpha, \beta \in \Lambda^0(V) = \mathbb{F}$ is not anticommutative. This gives us an opportunity to point out that while multiplication of elements of $V$ is anticommutative, $v_2 \wedge v_1 = -v_1 \wedge v_2$, this does not hold in general for the Grassmann Algebra. Indeed, consider in a four dimensional space the element $A = e_1 \wedge e_2 + e_3 \wedge e_4 \in \Lambda^2(V)$. Then we have

$$\begin{aligned}
A \wedge A &= (e_1 \wedge e_2 + e_3 \wedge e_4) \wedge (e_1 \wedge e_2 + e_3 \wedge e_4) \\
&= (e_1 \wedge e_2) \wedge (e_1 \wedge e_2) + (e_1 \wedge e_2) \wedge (e_3 \wedge e_4) \\
&+ (e_3 \wedge e_4) \wedge (e_1 \wedge e_2) + (e_3 \wedge e_4) \wedge (e_3 \wedge e_4) \\
&= 0 + (e_1 \wedge e_2) \wedge (e_3 \wedge e_4) + (e_3 \wedge e_4) \wedge (e_1 \wedge e_2) + 0 \\
&= e_1 \wedge e_2 \wedge e_3 \wedge e_4 + e_3 \wedge e_4 \wedge e_1 \wedge e_2 \\
&= e_1 \wedge e_2 \wedge e_3 \wedge e_4 + e_1 \wedge e_2 \wedge e_3 \wedge e_4 \\
&= 2\, e_1 \wedge e_2 \wedge e_3 \wedge e_4
\end{aligned}$$

where we have used $e_3 \wedge e_4 \wedge e_1 \wedge e_2 = -e_3 \wedge e_1 \wedge e_4 \wedge e_2 = +e_1 \wedge e_3 \wedge e_4 \wedge e_2 = -e_1 \wedge e_3 \wedge e_2 \wedge e_4 = +e_1 \wedge e_2 \wedge e_3 \wedge e_4$. Thus, while it is true that if $A = v_1 \wedge \ldots \wedge v_r$

then $A \wedge A = 0$, this need not be true when $A$ is not a pure product, or when $A \in \Lambda^0(V)$.

## 4.3 Permutations and Increasing Permutations

For efficient computations in the Grassmann Algebra and for use in the theory of determinants we must develop some efficient notations for permutations. We regard a permutation as a one to one onto function from the set $1, \ldots, r$ onto itself. We can diagram this by a symbol that puts the ordered pairs of the permutation function in vertical columns:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix}$$

When the top row of the matrix (the domain) is arranged in increasing order it is clearly redundant. However, omitting it would clash with another popular way of writing permutations, (cycle notation, which we will not use,) and this way of writing may also have slight benefits in clarity. Also we have the option of rearranging the upper row, for example,

$$\sigma = \begin{pmatrix} 4 & 3 & 2 & 5 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

which can occasionally be useful. In either case, the symbol represents the function $\sigma$ whose values are:

$$\begin{array}{lll} \sigma(1) = 5 & \sigma(2) = 3 & \sigma(3) = 2 \\ \sigma(4) = 1 & \sigma(5) = 4 & \end{array}$$

All the permutations of $n$ letters (called *permutations of order $n$)* form a group, the symmetric group $\mathcal{S}_n$. If we use the $\sigma$ above and introduce here:

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 5 & 3 \end{pmatrix}$$

we can then form the permutation

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 4 & 2 \end{pmatrix}$$

The permutations are here composed as functions would be; first $\tau$ and then $\sigma$. Thus we have $\sigma\tau(4) = \sigma(5) = 4$. It is most important to understand the order used here, especially since some people use the opposite order.

It is easy to see that the permutations form a group. Since they compose like functions and function composition is associative, we know the composition is associative. The identity is clearly

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

and the inverse of $\sigma$ can be found by rearranging $\sigma$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 2 & 5 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

and then "swapping the rows":

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 5 & 1 \end{pmatrix}.$$

The order of the group $\mathcal{S}_n$ is clearly $n!$, for there are $n$ choices for $\sigma(1)$, and then $n-1$ choices for $\sigma(2)$, $n-2$ choices for $\sigma(3), \ldots$, and finally 1 choice for $\sigma(1)$, giving a total of $n \cdot (n-1) \cdot (n-2) \ldots 1 = n!$ choices in total.

The next concept we introduce is the sign of a permutation which is absolutely critical in all that follows. Let $f(x_1, \ldots, x_n)$ be given by

$$f(x_1, \ldots, x_n) = \prod_{1 \le i < j \le r} (x_j - x_i)$$

Then define $\sigma f$, for $\sigma \in \mathcal{S}_n$, by

**Def** $$(\sigma f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

and then define $\operatorname{sgn}(\sigma)$ by

**Def** $$(\sigma f)(x_1, \ldots, x_n) = \operatorname{sgn}(\sigma) \cdot f(x_1, \ldots, x_n).$$

This makes sense, and $\operatorname{sgn}(\sigma) = \pm 1$, because

$$\begin{aligned} (\sigma f)(x_1, \ldots, x_n) &= f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \\ &= \prod_{1 \le i < j \le n} (x_{\sigma(j)} - x_{\sigma(i)}) \\ &= \operatorname{sgn}(\sigma) \cdot f(x_1, \ldots, x_n) \end{aligned}$$

and the product in the second line contains the same entries $(x_j - x_i)$ as the product for $f$ except for possible reversals of order. Each reversal of order contributes a $(-1)$ to the product, so the final result is the same as the original product except possibly for sign.

We are now going to look at an example so there will be no confusion about the way products of permutations act on $f$. We regard $\sigma$ as acting on $f$ to give $\sigma f$, and $\tau$ as acting similarly on $\sigma f$ to give $\tau(\sigma f)$ which ideally should be $(\tau\sigma)f$. To be concrete, let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$$

so that

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

Now the action of $\sigma$ on the arguments of $f$ is for all arguments except the first to march forward in the line, and for the last to go to the end of the line:

$$(\sigma f)(x_1, x_2, x_3, x_4) = f(x_{\sigma(1)}, x_{\sigma(1)}, x_{\sigma(4)}, \ldots, x_{\sigma(4)}) = f(x_2, x_3, x_4, x_1)$$

and similarly $\tau$ reverses the first two arguments of $f$:

$$(\tau f)(x_1, x_2, x_3, x_4) = f(x_2, x_1, x_3, x_4).$$

From this we can see that

$$(\tau(\sigma f))(w, x, y, z) = (\sigma f)(x, w, y, z) = f(w, y, z, x).$$

On the other hand,

$$((\tau\sigma)f)(x_1, x_2, x_3, x_4) = f(x_{\tau\sigma(1)}, x_{\tau\sigma(1)}, x_{\tau\sigma(4)}, \ldots, x_{\tau\sigma(4)}) = f(x_1, x_3, x_4, x_2)$$

so that, making the substitution,

$$((\tau\sigma)f)(w, x, y, z) = f(w, y, z, x)$$

which coincides with what we found above for $(\tau(\sigma f))(w, x, y, z)$. Hence we have shown that indeed we have $\tau(\sigma f) = (\tau\sigma)f$ as we desired. We have gone into this in such detail because experience shows the likelihood of confusion here.

To return to the general case then, we have

$$\begin{aligned}
(\tau(\sigma f)(x_1, \ldots, x_n) &= (\sigma f)(x_{\tau(1)}, \ldots, x_{\tau(n)}) \\
&= f(x_{\tau(\sigma(1))}, \ldots, x_{\tau(\sigma(n))}) \\
&= f(x_{(\tau\sigma)(1)}, \ldots, x_{(\tau\sigma)(n)}) \\
&= ((\tau\sigma)f)(x_1, \ldots, x_n).
\end{aligned}$$

The surprising thing here is the way the $\sigma$ jumps inside the $\tau(\cdot)$ in the second line, but this is the way it must work, because $\sigma$ rearranges the arguments according to their slot, not according to what is in them. Hence $\sigma$ must operate on the slot argument inside the $\tau(\cdot)$. If the reader finds this confusing, he should compare the general calculation with the example above.

With these details under control, we now have, using $\tau(\sigma f) = (\tau\sigma)f$,

$$\begin{aligned}
[\text{sgn}(\tau\sigma)]f &= (\tau\sigma)f \\
&= \tau(\sigma f) \\
&= \tau(\text{sgn}(\sigma)f) \\
&= \text{sgn}(\sigma) \cdot (\tau f) \\
&= \text{sgn}(\sigma) \cdot \text{sgn}(\tau) \cdot f
\end{aligned}$$

so we have

$$\text{sgn}(\tau\sigma) = \text{sgn}(\tau) \cdot \text{sgn}(\sigma)$$

and sgn is a homomorphism of $\mathcal{S}_n$ onto $\{1, -1\}$ regarded as a multiplicative group.

Next note that since $\sigma \cdot \sigma^{-1} = \text{identity}$, we have $\text{sgn}(\sigma) \cdot \text{sgn}(\sigma^{-1}) = \text{sgn}(\text{identity}) = 1$. Thus we have

$$\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$$

While the foregoing definition of sgn($\sigma$) is elegant and useful for theory, we will need a more convenient method for the calculation of sgn($\sigma$) for actual cases. To this end we first note that a permutation that has the form

$$\sigma_r = \begin{pmatrix} 1 & 2 & 3 & \dots & r-1 & r & r+1 & r+2 & \dots & n \\ 1 & 2 & 3 & \dots & r-1 & r+1 & r & r+2 & \dots & n \end{pmatrix}$$

has $\text{sgn}(\sigma_r) = -1$. Indeed, in the function $f$ used to define sgn($\sigma$), the terms $(x_i - x_r)$ and $(x_i - x_{r+1})$ will be exchanged into one another by $\sigma_r$ for $r+2 \le i \le n$ and similarly for the terms $(x_r - x_i)$, $(x_{r+1} - x_i)$, $1 \le i \le r-1$. The only real effect will be through the term $x_{r+1} - x_r$ which will be transformed into $x_r - x_{r+1}$, and this one sign reversal gives $\text{sgn}(\sigma_r) = -1$.

Next we are concerned with the permutation

$$\sigma_{ij} = \begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}$$

which exchanges the $i^{th}$ and $j^{th}$ entries. The permutations $\sigma_i, \sigma_{i+1}, \dots, \sigma_{j-1}$ will successively move $i$ to positions $i+1$, $i+2$, $\dots, j$ forcing $j$ into position $i-1$. Then permutation $\sigma_{j-2}, \sigma_{j-3}, \dots, \sigma_i$ will back up $j$ to positions $j-2$, $j-3, \dots, i$, completing the exchange of $i$ and $j$. The permutations strung out in the proper order are

$$\sigma_{ij} = \underbrace{\sigma_i \dots \sigma_{j-3}\sigma_{j-2}}_{j-i-1\,\text{terms}} \cdot \underbrace{\sigma_{j-1} \dots \sigma_{i+1}\sigma_i}_{j-i\,\text{terms}}$$

so that

$$\begin{aligned} \text{sgn}(\sigma_{ij}) &= \text{sgn}(\sigma_i) \dots \text{sgn}(\sigma_{j-3})\,\text{sgn}(\sigma_{j-2}) \cdot \text{sgn}(\sigma_{j-1} \dots \text{sgn}(\sigma_{i+1})\,\text{sgn}(\sigma_i) \\ &= (-1)^{j-i-1} \cdot (-1)^{j-i} = -1^{2(j-i)-1} = -1. \end{aligned}$$

Finally, for any permutation $\sigma$

$$\sigma' = \sigma\sigma_{ij} =$$

$$\begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ \sigma(1) & \dots & \sigma(i-1) & \sigma(i) & \sigma(i+1) & \dots & \sigma(j-1) & \sigma(j) & \sigma(j+1) & \dots & \sigma(n) \end{pmatrix}$$
$$\times \begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}$$
$$= \begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ \sigma(1) & \dots & \sigma(i-1) & \sigma(j) & \sigma(i+1) & \dots & \sigma(j-1) & \sigma(i) & \sigma(j+1) & \dots & \sigma(n) \end{pmatrix}$$

so that $\sigma'$ is almost $\sigma$ but the $i^{th}$ and $j^{th}$ entries are interchanged. Thus

$$\begin{aligned} \text{sgn}(\sigma') &= \text{sgn}(\sigma) \cdot \text{sgn}(\sigma_{ij}) \\ \text{sgn}(\sigma') &= -\text{sgn}(\sigma) \end{aligned}$$

Thus, if any two elements of a permutation are interchanged, it reverses the sign of the permutation. Now any permutation can, by means of interchanges (of adjacent elements if you like), be brought back to the identity permutation, whose sign is +1. Hence, the sign of the permutation is equal to the number of interchanges which return it to the identity. Here are some examples.

$$\begin{aligned}\sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix} \to \\ &\to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}\end{aligned}$$

or, using only adjacent interchanges, which takes longer

$$\begin{aligned}\sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 5 & 1 \end{pmatrix} \to \\ &\to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 1 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix} \to \\ &\to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}\end{aligned}$$

so that

$$\operatorname{sgn}(\sigma) = (-1)^3 = (-1)^7.$$

Thus we can find the sign of a permutation by counting the number of (possibly but not necessarily) adjacent interchanges necessary to return the permutation to the identity and raising $(-1)$ to that power. The various ways of doing the interchanges will always produce the same final result.

Having dealt with these preliminary general considerations, we turn our attention to the *increasing* permutations. The reason for our interest in these will appear shortly.

**Def** Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & r & r+1 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(r) & \sigma(r+1) & \dots & \sigma(n) \end{pmatrix}$$

Then $\sigma$ is an increasing $r-$permutation if and only if

$$\sigma(1) < \sigma(2) < \sigma(3) < \dots < \sigma(r)$$

and

$$\sigma(r+1) < \sigma(r+2) < \dots < \sigma(n).$$

We will denote the set of all increasing $r-$permutations by $\mathcal{S}_{n,r}$ and we will use the notation

$$\left(\begin{array}{ccccc|ccc} 1 & 2 & 3 & \dots & r & r+1 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(r) & \sigma(r+1) & \dots & \sigma(n) \end{array}\right)$$

Here are some examples:

$$\left(\begin{array}{cc|ccc} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{array}\right) \qquad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{array}\right) \qquad \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 3 & 5 \end{array}\right)$$

$$\left(\begin{array}{c|cccc} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{array}\right) \qquad \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 1 & 2 & 4 & 5 & 3 \end{array}\right)$$

Our interest in increasing permutations is due to their role acting as the indices for the basis elements in the Grassmann Algebra.

For example, the basis elements for $\Lambda^3(V)$ where $V$ is a vector space of dimension 5 with basis $\{e_1, e_2, e_3, e_4, e_5\}$ are

$$\begin{array}{cccc} e_1 \wedge e_2 \wedge e_3 & e_1 \wedge e_2 \wedge e_4 & e_1 \wedge e_2 \wedge e_5 & e_1 \wedge e_3 \wedge e_4 \\ e_1 \wedge e_3 \wedge e_5 & e_1 \wedge e_4 \wedge e_5 & e_2 \wedge e_3 \wedge e_4 & e_2 \wedge e_3 \wedge e_5 \\ & e_2 \wedge e_4 \wedge e_5 & e_3 \wedge e_4 \wedge e_5 & \end{array}$$

and we will write these as

$$e_\sigma = e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge e_{\sigma(3)}$$

where $\sigma \in \mathcal{S}_{5,3}$. The corresponding $\sigma$ are given by:

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 & 4 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{array}\right)$$

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 2 & 4 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & 2 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 1 & 5 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 1 & 4 \end{array}\right)$$

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{array}\right)$$

This is the method we will use to index basis elements of the Grassmann Algebra for the remainder of the book. Although it may look cumbersome at first, it is in fact quite efficient and elegant in practise. The indexing problem has always caused difficulty in the use of Grassmann Algebra and this method essentially removes the difficulty. It has been known for a long time, but the knowledge has not been widely disseminated, and it is greatly superior to the many horribly messy systems used when people invent an indexing system on site, so to speak.

We note in passing that an alternate way of forming the basis of $\Lambda^r(V)$ which has certain advantages is to use

$$\operatorname{sgn}(\sigma)\, e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)} \quad \text{instead of} \quad e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)}$$

However, we will not do this in this book except in certain special circumstances.

Note that when $\sigma(1), \ldots, \sigma(r)$ are known for $\sigma \in \mathcal{S}_{n,r}$, then $\sigma(r+1), \ldots, \sigma(n)$ are uniquely determined. Since there are $n$ elements to choose from and r are chosen to form $\sigma$, we see that $\mathcal{S}_{n,r}$ contains $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ elements. Thus

$$\text{The dimension of} \quad \Lambda^r(V) \quad \text{is} \quad \binom{n}{r}$$

We note that the scalers $\Lambda^0$ are one-dimensional; $\dim \Lambda^0 = 1$. Thus the total dimension of the entire Grassmann Algebra is

$$\begin{aligned}\sum_{r=0}^{n} \dim \Lambda^r(V) &= 1 + \binom{n}{1} + \binom{n}{2} + \dots \binom{n}{n-1} + \binom{n}{n} \\ &= (1+1)^n \\ &= 2^n\end{aligned}$$

Our next project is determining the sign of an increasing permutation which is considerably easier than determining the sign of an arbitrary permutation. First it is clear that $\sigma(j) \geq j$ for $j \leq r$, because in

$$\left(\begin{array}{ccccc|ccc} 1 & 2 & 3 & \dots & r & r+1 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(r) & \sigma(r+1) & \dots & \sigma(n) \end{array}\right)$$

we have $\sigma(1) \geq 1$ and $\sigma(1) < \sigma(2) < \ldots < \sigma(r)$. Consider now $\sigma(r)$. It is in position $r$, and we would like it returned to position $\sigma(r)$. This requires $\sigma(r) - r$ adjacent interchanges. Moreover, the elements $\sigma(r+1), \dots, \sigma(n)$, whatever their positions, retain their increasing order. Now repeat the process with $\sigma(r-1), \sigma(r-2), \dots, \sigma(1)$. These elements having returned to their positions in the identity permutation, and $\sigma(r+1), \dots, \sigma(n)$ having remained in increasing order, the final result must be the identity. The total number of interchanges is

$$\sum_{j=1}^{r} (\sigma(j) - j) = \sum_{j=1}^{r} \sigma(j) - \sum_{j=1}^{r} j = \sum_{j=1}^{r} \sigma(j) - T_r$$

where $T_r = \sum_{j=1}^{r} j = \frac{r(r+1)}{2}$ is the $r^{th}$ triangular number. Thus we have

$$\text{if} \quad \sigma \in \mathcal{S}_{n,r} \quad \text{then} \quad \operatorname{sgn}(\sigma) = (-1)^{\sum_{j=1}^{r} \sigma(j) - T_r}.$$

Examples:

$$\begin{aligned}
\sigma &= \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 1 & 3 & 6 \end{array}\right) & \operatorname{sgn}(\sigma) &= (-1)^{2+4+5-T_3} = (-1)^{11-6} = (-1)^5 = -1 \\
\sigma &= \left(\begin{array}{cc|cccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 1 & 2 & 4 & 5 \end{array}\right) & \operatorname{sgn}(\sigma) &= (-1)^{3+6-T_2} = (-1)^{9-3} = (-1)^6 = +1 \\
\sigma &= \left(\begin{array}{c|ccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 2 & 3 & 5 & 6 \end{array}\right) & \operatorname{sgn}(\sigma) &= (-1)^{4-T_1} = (-1)^{4-1} = (-1)^3 = -1
\end{aligned}$$

Our next concept is the *reverse* of a permutation, which is important when dealing with the dualizing operator $*$. Let $\sigma$ be given as usual by

$$\sigma = \left(\begin{array}{cccc|ccc} 1 & 2 & \dots & r & r+1 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(r) & \sigma(r+1) & \dots & \sigma(n) \end{array}\right) \in \mathcal{S}_{n,r}.$$

The reverse of $\sigma$, which we denote by $\tilde{\sigma}$, is given by

**Def** $$\tilde{\sigma} = \left(\begin{array}{cccc|ccc} 1 & 2 & \dots & n-r & n-r+1 & \dots & n \\ \sigma(r+1) & \sigma(r+1) & \dots & \sigma(n) & \sigma(1) & \dots & \sigma(r) \end{array}\right) \in \mathcal{S}_{n,n-r}.$$

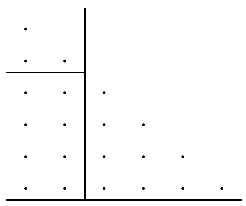
Example

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 \\ 2 & 4 & 5 & | & 1 & 3 \end{pmatrix} \qquad \tilde{\sigma} = \begin{pmatrix} 1 & 2 & | & 3 & 4 & 5 \\ 1 & 3 & | & 2 & 4 & 5 \end{pmatrix}$$

To determine the sign of $\tilde{\sigma}$, which will be of great importance in our future development, we need to know the value of $T_n - (T_r + T_{n-r})$ This is found by

$$\begin{aligned} T_n - (T_r + T_{n-r}) &= \frac{n(n+1)}{2} - \frac{r(r+1)}{2} - \frac{(n-r)(n-r+1)}{2} \\ &= \tfrac{1}{2}\big[(r+n-r)(n+1) - r(r+1) - (n-r)(n-r+1)\big] \\ &= \tfrac{1}{2}\big[r(n+1-r-1) - (n-r)(n+1-n+r-1)\big] \\ &= \tfrac{1}{2}[r(n-r) + (n-r)r] \\ &= r(n-r) \end{aligned}$$

This fact also has a geometrical demonstration:



$$T_2 + 2 \cdot 4 + T_4 = T_6.$$

We note also that for any permutation $\sigma \in \mathcal{S}_n$

$$\sum_{j=1}^{n} \sigma(j) = \sum_{j=1}^{n} j = T_n$$

so that we have

$$\begin{aligned} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tilde{\sigma}) &= (-1)^{\sum_{j=1}^{r} \sigma(j) - T_r}(-1)^{\sum_{j=r+1}^{n} \sigma(j) - T_{n-r}} \\ &= (-1)^{\sum_{j=1}^{n} \sigma(j) - (T_r + T_{n-r})} \\ &= (-1)^{T_n - (T_r + T_{n-r})} \\ &= (-1)^{r(n-r)}. \end{aligned}$$

Thus we see that

$$\operatorname{sgn}(\tilde{\sigma}) = (-1)^{r(n-r)}\operatorname{sgn}(\sigma).$$

This completes our study of increasing permutations.

## 4.4 Determinants

in this section we will discuss the computational aspects of Grassmann Algebra which are inextricably linked with the theory of determinants. Because we believe that determinants are important precisely because they are the coefficients of Grassmann Algebra, we will develop their theory as if the reader had never seen them before.

There is a definition of determinants (due to Weierstrass) that bases the theory on certain axioms. We will prove these axioms, (and point out the axioms in passing) but we will not develop the theory in this way.

The theory of determinants, since it involves permutations, cannot be made totally comfortable. The following treatment, while requiring close attention, is about as pleasant as possible. Readers who find the going rough in spots might profitably compare it with more traditional developments of the theory.

The principal problem in elementary computations in Grassmann Algebra involves the rearrangement of terms in a product, with consequent possible change in sign. Consider the product

$$v_{i_1} \wedge v_{i_2} \wedge \ldots \wedge v_{i_r}$$

where $i_1, i_2, \ldots i_r$ are integers between 1 and $m = \dim V$. As we will see, it is essential to rearrange the product into one in which the indices increase. For example, suppose $r = 4$ and $i_2 < i_4 < i_3 < i_1$. Since $v_i \wedge v_j = v_j \wedge v_i$ we have

$$\begin{aligned} v_{i_1} \wedge v_{i_2} \wedge v_{i_3} \wedge v_{i_r} &= -v_{i_2} \wedge v_{i_1} \wedge v_{i_3} \wedge v_{i_4} = v_{i_2} \wedge v_{i_3} \wedge v_{i_1} \wedge v_{i_4} \\ &= -v_{i_2} \wedge v_{i_3} \wedge v_{i_4} \wedge v_{i_1} = v_{i_2} \wedge v_{i_4} \wedge v_{i_3} \wedge v_{i_1} \end{aligned}$$

The process corresponds exactly to the interchange of adjacent elements turning 1, 2, 3, 4 into 2, 4, 3, 1, which determines the sign of the permutation

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

Hence,

$$v_{i_1} \wedge v_{i_2} \wedge v_{i_3} \wedge v_{i_r} = \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} v_{i_2} \wedge v_{i_4} \wedge v_{i_3} \wedge v_{i_1} .$$

Exactly the same reasoning establishes that, in general,

$$v_{i_1} \wedge \ldots \wedge v_{i_r} = \operatorname{sgn}(\pi) v_{i_{\pi(1)}} \wedge \ldots \wedge v_{i_{\pi(r)}} .$$

In particular if $i_l = l$ we have, for $\pi \in \mathcal{S}_r$,

$$v_1 \wedge \ldots \wedge v_r = \operatorname{sgn}(\pi) v_{\pi(1)} \wedge \ldots \wedge v_{\pi(r)} .$$

Since $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1}) = 1/\operatorname{sgn}(\pi)$, these equations may also be written

$$\begin{aligned} v_{i_{\pi(1)}} \wedge \ldots \wedge v_{i_{\pi(r)}} &= \operatorname{sgn}(\pi) v_{i_1} \wedge \ldots \wedge v_{i_r} \\ v_{\pi(1)} \wedge \ldots \wedge v_{\pi(r)} &= \operatorname{sgn}(\pi) v_1 \wedge \ldots \wedge v_r . \end{aligned}$$

As Axiom 4b of section 3.1 states, a basis of $\Lambda(V)$ is given by

$$e_\sigma = e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \ldots \wedge e_{\sigma(r)} \qquad \sigma \in \mathcal{S}_{n,r}.$$

This suggests that we show interest in elements of the form

$$v_\sigma = v_{\sigma(1)} \wedge v_{\sigma(2)} \wedge \ldots \wedge v_{\sigma(r)} \in \mathcal{S}_{n,r} \qquad \sigma \in \mathcal{S}_{n,r}.$$

where the $v_i \in V$ but $v_1, v_2, \ldots$ do not necessarily form a basis. As an example, let us consider

$$\begin{aligned} w_1 &= 2v_1 + 4v_2 - v_3 \\ w_2 &= v_1 - v_2 + 2v_3 \end{aligned}$$

where $w_i, v_j \in \mathcal{S}_{n,r}$, and we make no explicit assumptions about he dimension of the space $V$. Then

$$\begin{aligned} w_1 \wedge w_2 &= 2v_1 \wedge v_1 - 2v_1 \wedge v_2 + 4v_1 \wedge v_3 \\ &+ 4v_2 \wedge v_1 - 4v_2 \wedge v_2 + 8v_2 \wedge v_3 \\ &- 1v_3 \wedge v_1 - 1v_3 \wedge v_2 + 2v_3 \wedge v_3 \end{aligned}$$

Since $v_i \wedge v_i = 0$ and $v_j \wedge v_i = -v_i \wedge v_j$, we have

$$\begin{aligned} w_1 \wedge w_2 &= (-2-4)v_1 \wedge v_2 + (4+1)v_1 \wedge v_3 + (8-1)v_2 \wedge v_3 \\ &= -6v_1 \wedge v_2 + 5v_1 \wedge v_3 + 7v_2 \wedge v_3 \end{aligned}$$

The problem is how to come up with these coefficients without writing out all the intermediate steps. If we arrange the original coefficients as columns in a matrix

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

then the coefficient of $v_1 \wedge v_3$ is formed from the $1^{st}$ and $3^{rd}$ row, that is, from the square submatrix

$$\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

by forming $2 \cdot 2 - (-1) \cdot 1 = 5$, and similarly for $v_1 \wedge v_2$ and $v_2 \wedge v_3$. The important thing is that for each square matrix the same arithmetic process

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \to ad - bc$$

solves the coefficient problem. Clearly, the next problem is to increase the number of factors in the product and clarify the formation of the coefficients from the square submatrices. (The clever reader will have deduced that we are sneaking up on determinants).

To this end, suppose that (summation convention now in effect)

$$w_j = \alpha_j^i v_i \qquad j = 1, \dots, r \quad i = 1, \dots, m.$$

(We are still not making any assumptions about the dimension of the space or whether $v_1, \dots, v_m$ is a basis.) We now form

$$w_1 \wedge \dots \wedge v_r = \alpha_1^{i_1} \alpha_2^{i_2} \dots \alpha_r^{i_r} \, v_{i_1} \wedge \dots \wedge v_{i_1} \, .$$

It will now be necessary to rearrange the terms on the right hand side from $r$ simple sums into a sum whose form is much better adapted to further computation. The methodology here is quite important wand will recur in a number of critical places.

We must rewrite the last sum in the following way. We select an element $\sigma \in \mathcal{S}_{m,r}$ and group together those sets of indices which are rearrangements of $\sigma(1), \dots, \sigma(r)$. We then sum over all $\sigma \in \mathcal{S}_{m,r}$. In this way, we get all possible sets of values of $i_1, \dots, i_r$.

For a fixed $\sigma \in \mathcal{S}_{m,r}$ the terms in the previous sum whose indices are rearrangements of $\sigma(1), \dots, \sigma(r)$ can be rewritten using a $\pi \in \mathcal{S}_r$. (This is actually the key point in the derivation.) Thus $i_1, i_2, \dots, i_r$ when rearranged in increasing order become $i_{\pi(1)}, i_{\pi(2)}, \dots, i_{\pi(r)}$ which then coincide with the increasing $\sigma(1), \sigma(2), \dots, \sigma(r)$ where $i_{\pi(j)} = \sigma(j)$ and $i_k = \sigma(\pi^{-1}(k))$. We then have, using our previous knowledge of rearrangements,

$$\begin{aligned} v_{i_1} \wedge v_{i_2} \wedge \dots v_{i_r} \wedge &= \operatorname{sgn}(\pi) \, v_{i_{\pi(1)}} \wedge v_{i_{\pi(2)}} \wedge \dots \wedge v_{i_{\pi(r)}} \\ &= \operatorname{sgn}(\pi) \, v_{\sigma(1)} \wedge v_{\sigma(2)} \wedge \dots \wedge v_{\sigma(r)} \end{aligned}$$

All the terms which are rearrangements of $\sigma(1), \dots, \sigma(r)$ then sum to

$$\sum_{\pi \in \mathcal{S}_r} \alpha_1^{\sigma(\pi^{-1}(1))} \alpha_2^{\sigma(\pi^{-1}(2))} \dots \alpha_r^{\sigma(\pi^{-1}(r))} \operatorname{sgn}(\pi) \, v_{\sigma(1)} \wedge v_{\sigma(2)} \wedge \dots \wedge v_{\sigma(r)}$$

and the final sum will be

$$\begin{aligned} w_1 \wedge \dots \wedge w_r &= \alpha_1^{i_1}, \dots \alpha_1^{i_r} \, v_{i_1} \wedge \dots \wedge v_{i_r} \\ &= \sum_{\sigma \in \mathcal{S}_{m,r}} \sum_{\pi \in \mathcal{S}_r} \alpha_1^{\sigma(\pi^{-1}(1))} \dots \alpha_r^{\sigma(\pi^{-1}(r))} \operatorname{sgn}(\pi) \, v_{\sigma(1)} \wedge \dots \wedge v_{\sigma(r)} \end{aligned}$$

We will soon introduce the concept of determinant to conveniently express the inner sum. This method of breaking up a multiple sum $i_1, \dots, i_r$ into a sum over $\mathcal{S}_{m,r}$ and $\mathcal{S}_r$ will be called *resolving a sum by permutations.*

If we now write the coefficients in an $m \times r$ matrix

$$\begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \dots & \alpha_r^1 \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_r^2 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \alpha_1^m & \alpha_2^m & \dots & \alpha_r^m \end{pmatrix}$$

then the coefficient of $v_{\sigma(1)}\wedge\ldots\wedge v_{\sigma(r)}, \quad \sigma\in\mathcal{S}_{m,r}$ is associated with the square matrix

$$\begin{pmatrix} \alpha_1^{\sigma(1)} & \alpha_2^{\sigma(1)} & \ldots & \alpha_r^{\sigma(1)} \\ \alpha_1^{\sigma(2)} & \alpha_2^{\sigma(2)} & \ldots & \alpha_r^{\sigma(2)} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^{\sigma(r)} & \alpha_2^{\sigma(r)} & \ldots & \alpha_r^{\sigma(r)} \end{pmatrix}$$

in the sense that the coefficient associated with it,

$$\sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,\alpha_1^{\sigma(\pi^{-1}(1))}\ldots\alpha_r^{\sigma(\pi^{-1}(r))},$$

is a function of the entries of this particular square matrix. Suppose we now rearrange the $\alpha$'s so that, instead of coming in the order $1, 2, \ldots r$ they come in the order $\pi(1), \pi(2), \ldots, \pi(r)$ Since the $\alpha$'s are mere scalars, this will not effect the value and we will have

$$\sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,\alpha_{\pi(1)}^{\sigma(\pi^{-1}\pi(1))}\ldots\alpha_{\pi(r)}^{\sigma(\pi^{-1}\pi(r))} = \sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,\alpha_{\pi(1)}^{\sigma(1)}\ldots\alpha_{\pi(r)}^{\sigma(r)}.$$

We now observe that the $\sigma(i)$ are functioning as mere labels in to distinguish the rows of the coefficient matrix, and we can specialize the last expression without losing generality. Since the last expression *determines* the value of the coefficient in the Grassmann Algebra as a function of the coefficients of the vectors, we define

**Def** The determinant of the square matrix

$$\begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \ldots & \alpha_r^1 \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_r^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^r & \alpha_2^r & \ldots & \alpha_r^r \end{pmatrix}$$

is

$$\sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,\alpha_{\pi(1)}^1\ldots\alpha_{\pi(r)}^r.$$

We will write

$$\det(\alpha_j^i) = \sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,\alpha_{\pi(1)}^1\ldots\alpha_{\pi(r)}^r.$$

The more general formula with the $\sigma$'s results from this by mere substitution:

$$\det(\alpha_j^{\sigma(i)}) = \sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,\alpha_{\pi(1)}^{\sigma(1)}\ldots\alpha_{\pi(r)}^{\sigma(r)}$$

and the final formula for the Grassmann product with our new notation for the coefficients becomes

$$w_1 \wedge \ldots \wedge w_r = \sum_{\sigma \in \mathcal{S}_{m,r}} \det(\alpha_j^{\sigma(i)})\, v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(r)}, \qquad 1 \le i,j \le r.$$

We can now begin to exploit these formulas to derive properties of the determinant function from properties of the Grassmann Algebra. Let $e_1, e_2, \ldots, e_n$ be a basis for the vector space $V$. Axiom 4b tells us that $e_1 \wedge \ldots \wedge e_n$ is a basis of $\Lambda^n(V)$ (and hence not 0). Let

$$w_j = \alpha_j^i\, e_i \qquad 1 \le i,j \le r.$$

Since $\mathcal{S}_{n,n}$ consists solely of the identity, we have

$$w_1 \wedge \ldots \wedge w_n = \det(\alpha_j^i)\, e_1 \wedge \ldots \wedge e_n\,.$$

This formula allows us to derive the more elementary properties of determinants. For example

**Theorem 1** The determinant is a multilinear function of its columns.

**Proof** Let $u = \beta^i e_i$, $v = \sigma^i e_i$ and $w_j = \alpha_j^i\, e_i,\ j = 2, \ldots, n$. Then

$$\begin{aligned}
\mu\,(u \wedge w_2 \wedge \ldots \wedge w_n) &= \mu \det \begin{pmatrix} \beta^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \ldots & \ldots & \ldots & \ldots \\ \beta^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix} e_1 \wedge \ldots \wedge e_n \\
\nu\,(v \wedge w_2 \wedge \ldots \wedge w_n) &= \nu \det \begin{pmatrix} \gamma^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \ldots & \ldots & \ldots & \ldots \\ \gamma^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix} e_1 \wedge \ldots \wedge e_n \\
(\mu u + \nu v) \wedge w_2 \wedge \ldots \wedge w_n &= \det \begin{pmatrix} \mu\beta^1 + \nu\gamma^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \ldots & \ldots & \ldots & \ldots \\ \mu\beta^n + \nu\gamma^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix} e_1 \wedge \ldots \wedge e_n
\end{aligned}$$

Since the sum of the first two expressions on the left is the third, and since $e_1 \wedge \ldots \wedge e_n$ is a basis for $\Lambda^n$, we must have

$$\begin{aligned}
\mu \det \begin{pmatrix} \beta^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \ldots & \ldots & \ldots & \ldots \\ \beta^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix} \quad &+ \quad \nu \det \begin{pmatrix} \gamma^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \ldots & \ldots & \ldots & \ldots \\ \gamma^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix} \\
&= \det \begin{pmatrix} \mu\beta^1 + \nu\gamma^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \ldots & \ldots & \ldots & \ldots \\ \mu\beta^n + \nu\gamma^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix}
\end{aligned}$$

which shows that the determinant is linear in the first column. Similarly, it is linear in all the other columns.

**Theorem 2a** If a determinant has two identical columns then its value is zero.

**Proof** Let $w_j = \alpha^i_j e_i$. If $w_i = w_j$ then the $i^{th}$ and the $j^{th}$ columns of the determinant will be be identical. Also $w_1 \wedge w_2 \wedge w_3 \wedge \ldots \wedge w_n = 0$ and we have

$$0 = w_1 \wedge w_2 \wedge w_3 \wedge \ldots \wedge w_n = \det \begin{pmatrix} \alpha^1_1 & \alpha^1_2 & \alpha^1_3 & \ldots & \alpha^1_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha^n_1 & \alpha^n_2 & \alpha^n_3 & \ldots & \alpha^n_n \end{pmatrix} e_1 \wedge \ldots \wedge e_n$$

and since $e_1 \wedge \ldots \wedge e_n$ is a basis for $\Lambda^n(V)$ we must have

$$\det \begin{pmatrix} \alpha^1_2 & \alpha^1_1 & \alpha^1_3 & \ldots & \alpha^1_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha^n_2 & \alpha^n_1 & \alpha^n_3 & \ldots & \alpha^n_n \end{pmatrix} = 0$$

**Theorem 2b** Interchanging two adjacent columns of a determinant alters its sign. (We express this by saying that a determinant is an alternating function of its columns.)

**Proof** Let $w_j = \alpha^i_j e_i$. Then

$$\begin{aligned} w_1 \wedge w_2 \wedge w_3 \wedge \ldots \wedge w_n &= \det \begin{pmatrix} \alpha^1_1 & \alpha^1_2 & \alpha^1_3 & \ldots & \alpha^1_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha^n_1 & \alpha^n_2 & \alpha^n_3 & \ldots & \alpha^n_n \end{pmatrix} e_1 \wedge \ldots \wedge e_n \\ w_2 \wedge w_1 \wedge w_3 \wedge \ldots \wedge w_n &= \det \begin{pmatrix} \alpha^1_2 & \alpha^1_1 & \alpha^1_3 & \ldots & \alpha^1_n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha^n_2 & \alpha^n_1 & \alpha^n_3 & \ldots & \alpha^n_n \end{pmatrix} e_1 \wedge \ldots \wedge e_n \end{aligned}$$

Since $w_1 \wedge w_2 = -w_2 \wedge w_1$, the two determinants, with the first two columns switched, are negatives of one another. The result is clearly identical for interchanging any two adjacent columns.

**Corollary**

$$\det \begin{pmatrix} \alpha^1_{\sigma(1)} & \ldots & \alpha^1_{\sigma(n)} \\ \ldots & \ldots & \ldots \\ \alpha^n_{\sigma(1)} & \ldots & \alpha^n_{\sigma(n)} \end{pmatrix} = \mathrm{sgn}(\sigma) \det \begin{pmatrix} \alpha^1_1 & \ldots & \alpha^1_n \\ \ldots & \ldots & \ldots \\ \alpha^n_1 & \ldots & \alpha^n_n \end{pmatrix}$$

**Proof** The sign of the determinant on the left is related to that on the right according to the number of adjacent interchanges necessary to return $\sigma(1), \ldots, \sigma(n)$ to $1, \ldots, n$. But this is also a way to determine $\mathrm{sgn}(\sigma)$.

**Corollary** If two columns of a determinant are equal, the determinant is 0.

**Theorem 3** Let $(\alpha^i_j)$ be the identity matrix; $\alpha^i_j = 0$ for $i \neq j$ and $\alpha^i_j = 1$ for $i = j$. Then $\det(\alpha^i_j) = 1$.

**Proof** It is clear that $e_j = \alpha^i_j e_i$. Thus

$$e_1 \wedge \ldots \wedge e_n = \det(\alpha^i_j)\, e_1 \wedge \ldots \wedge e_n$$

by the basic formula, and this gives $\det(\alpha^i_j) = 1$.

Theorems 1,2b,3 are the basic axioms of Weierstrass for the determinant function on square matrices. From these three basic theorems the entire theory of determinants can be developed provided the characteristic of the Field is not 2. It is preferable to substitute Theorems 1,2a,3 for Weierstrass's Theorems 1,2b,3 because, if the characteristic of the Field is 2 then Theorem 2b can be derived from Theorem 2a but not vice versa. If the characteristic is not 2 then Theorem 2a and 2b can each be derived from the other. We leave it as an instructive exercise for the reader to derive Theorem 2b from Theorem 2a without using Grassmann products.

We will not follow the Weierstrass procedure however, because we wish to exploit the associative law for Grassmann products which makes the Laplace expansion (next section) much easier than deriving it from the Weierstrass axioms. Similarly, we did not derive Theorem 2b from Theorem 2a because our main goal is to illustrate that determinants are the coefficients in Grassmann algebra calculations and so we use Grassmann techniques where possible.

Next we prove one of the most important theorems of determinant theory. This is relatively easy for us to do using the properties of the Grassmann product. In fact, the ease and naturalness of this proof is an example of how productive the use of the Grassmann product can be. Without it, this theorem requires some sort of artifice, and this reflects the essential nature of the Grassmann product; without it one must resort to tricky procedures.

**Theorem 3** Let $A$ and $B$ be square matrices. Then $\det(AB) = \det(A)\det(B)$

**Proof** Let $A = (\alpha^i_j)$, $B = (\beta^i_j)$ and $C = (\gamma^i_j)$ where $1 \leq i, j \leq n$. Then $\gamma^i_j = \alpha^i_k \beta^k_j$ since this is the definition of matrix multiplication. Let $V$ be an n–dimensional vector space with basis $e_1, \ldots, e_n$. Let $v_k = \alpha^i_k e_i$ and $w_j = \beta^k_j v_k$. Then

$$\begin{aligned} w_1 \wedge \ldots \wedge w_n &= \det(\beta^k_j)\, v_1 \wedge \ldots \wedge v_n \\ v_1 \wedge \ldots \wedge v_n &= \det(\alpha^i_k)\, e_1 \wedge \ldots \wedge e_n \end{aligned}$$

so

$$\begin{aligned} w_1 \wedge \ldots \wedge w_n &= \det(\beta^k_j)\, v_1 \wedge \ldots \wedge v_n \\ &= \det(\beta^k_j)\det(\alpha^i_k)\, e_1 \wedge \ldots \wedge e_n \end{aligned}$$

On the other hand

$$w_j = \beta^k_j v_k = \beta^k_j(\alpha^i_k e_i) = (\beta^k_j \alpha^i_k) e_i = \gamma^i_j e_i$$

so

$$w_1 \wedge \ldots \wedge w_n = \det(\gamma^i_j)\, e_1 \wedge \ldots \wedge e_n$$

Comparing the two expressions for $w_1 \wedge \ldots \wedge w_n$, we see that $\det(\gamma^i_j) = \det(\alpha^i_k)\det(\beta^k_j)$.

Next we wish to discover how to compute a determinant by using its subdeterminants. This is done via the associative law for products. We will need to know the value of

$$(e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)}) \wedge (e_{\rho(1)} \wedge \ldots \wedge e_{\rho(n-r)})$$

for $\sigma \in \mathcal{S}_{n,r}$ and $\rho \in \mathcal{S}_{n,n-r}$. This will be 0 unless the $\rho(1), \ldots, \rho(n-r)$ are *all* distinct from *all* the $\sigma(1), \ldots, \sigma(n-r)$. Thus the $\rho(1), \ldots, \rho(n-r)$ must be the same numbers as the $\sigma(r+1), \ldots, \sigma(n)$. Since both are increasing sequences, $\sigma(r+j) = \rho(j)$. Similarly, $\rho(n-r+j) = \sigma(j)$. But then $\rho = \tilde{\sigma}$, the reverse of $\sigma$, and the non-zero elements of this form are

$$e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)} \wedge e_{\sigma(r+1)} \wedge \ldots \wedge e_{\sigma(n)} = e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(n)}$$

With this in mind, if $v_j = \alpha^i_j e_i$ we have, for $\tau \in \mathcal{S}_{n,r}$

$$\begin{aligned}
\det(\alpha^i_j) e_1 \wedge \ldots \wedge e_n &= \\
&= v_1 \wedge \ldots \wedge v_n \\
&= \operatorname{sgn}(\tau)\, v_{\tau(1)} \wedge \ldots \wedge v_{\tau(n)} \\
&= \operatorname{sgn}(\tau)\, (v_{\tau(1)} \wedge \ldots \wedge v_{\tau(r)}) \wedge (v_{\tau(r+1)} \wedge \ldots \wedge v_{\tau(n)}) \\
&= \operatorname{sgn}(\tau) \Big( \sum_{\sigma \in \mathcal{S}_{n,r}} \det(\alpha^{\sigma(i)}_{\tau(j)}) e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)} \Big) \Big( \sum_{\rho \in \mathcal{S}_{n,n-r}} \det(\alpha^{\rho(k)}_{\tau(l)}) e_{\rho(1)} \wedge \ldots \wedge e_{\rho(n-r)} \Big) \\
&= \operatorname{sgn}(\tau) \sum_{\substack{\sigma \in \mathcal{S}_{n,r} \\ \rho \in \mathcal{S}_{n,n-r}}} \det(\alpha^{\sigma(i)}_{\tau(j)}) \det(\alpha^{\rho(k)}_{\tau(l)})\, e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)} \wedge e_{\rho(1)} \wedge \ldots \wedge e_{\rho(n-r)} \\
&= \operatorname{sgn}(\tau) \sum_{\sigma \in \mathcal{S}_{n,r}} \operatorname{sgn}(\sigma) \det(\alpha^{\sigma(i)}_{\tau(j)}) \det(\alpha^{\sigma(r+k)}_{\tau(l)})\, e_1 \wedge \ldots \wedge e_n
\end{aligned}$$

where $1 \le i, j \le r$ and $1 \le k, l \le n-r$.

This may be simplified slightly since

$$\sum_{i=1}^{r} \sigma(i) - T_r + \sum_{j=1}^{r} \tau(j) - T_r = \sum_{i=1}^{r} \big(\sigma(i) + \tau(j)\big) - 2T_r$$

so that, using the formula for $\operatorname{sgn}(\sigma)$ when $\sigma \in \mathcal{S}_{n,r}$

$$\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = (-1)^{\sum_{i=1}^{r}\left(\sigma(i)+\tau(j)\right)-2T_r} = (-1)^{\sum_{i=1}^{r}\left(\sigma(i)+\tau(j)\right)}$$

and thus

$$\det(\alpha^i_j) = \sum_{\sigma \in \mathcal{S}_{n,r}} (-1)^{\sum_{i=1}^{r}\left(\sigma(i)+\tau(j)\right)} \det \begin{pmatrix} \alpha^{\sigma(1)}_{\tau(1)} & \ldots & \alpha^{\sigma(1)}_{\tau(r)} \\ \ldots & \ldots & \ldots \\ \alpha^{\sigma(r)}_{\tau(1)} & \ldots & \alpha^{\sigma(r)}_{\tau(r)} \end{pmatrix} \det \begin{pmatrix} \alpha^{\sigma(r+1)}_{\tau(r+1)} & \ldots & \alpha^{\sigma(r+1)}_{\tau(n)} \\ \ldots & \ldots & \ldots \\ \alpha^{\sigma(n)}_{\tau(r+1)} & \ldots & \alpha^{\sigma(n)}_{\tau(n)} \end{pmatrix}$$

This is called Laplace's expansion by complementary minors. For later purposes, we will need to express all this with less writing, so we introduce some notation.

First we write

**Def** $$\begin{vmatrix} \alpha_1^1 & \dots & \alpha_1^r \\ \dots & \dots & \dots \\ \alpha_r^1 & \dots & \alpha_r^r \end{vmatrix} = \det \begin{pmatrix} \alpha_1^1 & \dots & \alpha_1^r \\ \dots & \dots & \dots \\ \alpha_r^1 & \dots & \alpha_r^r \end{pmatrix}$$

Next we write, for a not necessarily square matrix $(\alpha_j^i)$,

**Def** $$\alpha_{\tau(j_1 \dots j_l)}^{\sigma(i_1 \dots i_l)} = \begin{vmatrix} \alpha_{\tau(j_1)}^{\sigma(i_1)} & \dots & \alpha_{\tau(j_l)}^{\sigma(i_1)} \\ \dots & \dots & \dots \\ \alpha_{\tau(j_l)}^{\sigma(i_1)} & \dots & \alpha_{\tau(j_l)}^{\sigma(i_l)} \end{vmatrix}$$

where if $\sigma$ and $\tau$ are the identity we abbreviate

$$\alpha_{j_1 \dots j_l}^{i_1 \dots i_l} = \begin{vmatrix} \alpha_{j_1}^{i_1} & \dots & \alpha_{j_l}^{i_1} \\ \dots & \dots & \dots \\ \alpha_{j_1}^{i_l} & \dots & \alpha_{j_l}^{i_l} \end{vmatrix}$$

and if $i_1, \dots, i_l$ and $j_1, \dots, j_l$ are just $1, \dots, l$ (and this is the common situation) we abbreviate with

$$\alpha_\tau^\sigma = \begin{vmatrix} \alpha_{\tau(1)}^{\sigma(1)} & \dots & \alpha_{\tau(l)}^{\sigma(1)} \\ \dots & \dots & \dots \\ \alpha_{\tau(l)}^{\sigma(1)} & \dots & \alpha_{\tau(l)}^{\sigma(l)} \end{vmatrix}$$

With these abbreviations in mind and recalling that $\sigma(r+k) = \tilde{\sigma}(k)$ for $\sigma \in \mathcal{S}_{n,r}$ and $\tilde{\sigma}$ the reverse of $\sigma$, we have Laplace's expansion written in the relatively benign forms

$$\begin{aligned} \det(\alpha_j^i) &= \sum_{\sigma \in \mathcal{S}_{n,r}} (-1)^{\sum_{i=1}^r \left(\sigma(i)+\tau(i)\right)} \alpha_{\tau(1,\dots,r)}^{\sigma(1,\dots,r)} \alpha_{\tau(r+1,\dots,n)}^{\sigma(r+1,\dots,n)} \\ &= \sum_{\sigma \in \mathcal{S}_{n,r}} (-1)^{\sum_{i=1}^r \left(\sigma(i)+\tau(i)\right)} \alpha_\tau^\sigma \alpha_{\tilde{\tau}}^{\tilde{\sigma}} \\ &= \sum_{\sigma \in \mathcal{S}_{n,r}} \operatorname{sgn}(\sigma\tau) \alpha_\tau^\sigma \alpha_{\tilde{\tau}}^{\tilde{\sigma}} \end{aligned}$$

Note that the $\tau$ here is some *fixed* member of $\mathcal{S}_{n,r}$ which the expander chooses at his own convenience. It amounts to selecting a set of r columns. For example in the expansion of

$$\begin{vmatrix} \alpha_1^1 & \alpha_2^1 & \alpha_3^1 & \alpha_4^1 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \alpha_4^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 & \alpha_4^3 \\ \alpha_1^4 & \alpha_2^4 & \alpha_3^4 & \alpha_4^4 \end{vmatrix}$$

let us select

$$\tau = \left( \begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{array} \right)$$

which amounts to selecting the first and third column. Using the formula and writing each $\sigma \in \mathcal{S}_{n,r}$ below the corresponding term, we have

$$
\begin{aligned}
\det(\alpha^i_j) &= \\
&= (-1)^{1+2+1+3} \begin{vmatrix} \alpha^1_1 & \alpha^1_3 \\ \alpha^2_1 & \alpha^2_3 \end{vmatrix} \begin{vmatrix} \alpha^3_2 & \alpha^3_4 \\ \alpha^4_2 & \alpha^4_4 \end{vmatrix} \quad + \quad (-1)^{1+3+1+3} \begin{vmatrix} \alpha^1_1 & \alpha^1_3 \\ \alpha^3_1 & \alpha^3_3 \end{vmatrix} \begin{vmatrix} \alpha^2_2 & \alpha^2_4 \\ \alpha^4_2 & \alpha^4_4 \end{vmatrix} \\
&\qquad \sigma = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{array}\right) \qquad\qquad \sigma = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{array}\right) \\
&+ (-1)^{1+4+1+3} \begin{vmatrix} \alpha^1_1 & \alpha^1_3 \\ \alpha^4_1 & \alpha^4_3 \end{vmatrix} \begin{vmatrix} \alpha^2_2 & \alpha^2_4 \\ \alpha^3_2 & \alpha^3_4 \end{vmatrix} \quad + \quad (-1)^{2+3+1+3} \begin{vmatrix} \alpha^2_1 & \alpha^2_3 \\ \alpha^3_1 & \alpha^3_3 \end{vmatrix} \begin{vmatrix} \alpha^1_2 & \alpha^1_4 \\ \alpha^4_2 & \alpha^4_4 \end{vmatrix} \\
&\qquad \sigma = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{array}\right) \qquad\qquad \sigma = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{array}\right) \\
&+ (-1)^{2+4+1+3} \begin{vmatrix} \alpha^2_1 & \alpha^2_3 \\ \alpha^4_1 & \alpha^4_3 \end{vmatrix} \begin{vmatrix} \alpha^1_2 & \alpha^1_4 \\ \alpha^3_2 & \alpha^3_4 \end{vmatrix} \quad + \quad (-1)^{3+4+1+3} \begin{vmatrix} \alpha^3_1 & \alpha^3_3 \\ \alpha^4_1 & \alpha^4_3 \end{vmatrix} \begin{vmatrix} \alpha^1_2 & \alpha^1_4 \\ \alpha^2_2 & \alpha^2_4 \end{vmatrix} \\
&\qquad \sigma = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{array}\right) \qquad\qquad \sigma = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{array}\right)
\end{aligned}
$$

An important special case of the foregoing occurs when $r = 1$ which means that only a single column is selected. Then

$$
\tau = \left(\begin{array}{c|ccccccc} 1 & 2 & \dots & j-1 & j & j+1 & \dots & n \\ j & 2 & \dots & j-1 & j+1 & j+2 & \dots & n \end{array}\right)
$$

and the $\sigma$ have the same form. The formula then reduces to

$$
\det(\alpha^i_j) = \sum_{i=1}^{n} (-1)^{i+j} \alpha^i_j \alpha^{1\dots i-1, i+1 \dots n}_{1 \dots j-1, j+1 \dots n}
$$

where the second factor is the determinant of the matrix obtained by eliminating the row and column in which $\alpha^i_j$ lies. This is the familiar expansion of a determinant by the cofactors of a column.

Naturally, since $\det A^\top = \det A$, we may do the expansion by selecting a set of *rows* instead of a set of *columns*.

## 4.5 Extending a Linear Transformation to the Grassmann Algebra and the Cauchy–Binet Theorem

Our next effort will be to extend a linear transformation $T : V \to W$ so that it maps $\Lambda^r(V)$ to $\Lambda^r(W)$. Consider the transformation $f : V \times \ldots \times V \to W$ given by

$$f(v_1, v_2, \ldots, v_r) = Tv_1 \wedge Tv_2 \wedge \ldots \wedge Tv_r$$

Clearly $f$ is multilinear and alternating, so that by Axiom 4b there is a map $\tilde{T} : \Lambda^r(V) \to \Lambda^r(W)$ which makes the following diagram commute.

$$\begin{array}{ccc} V_1 \times \ldots \times V_n & \xrightarrow{f} & \Lambda^r(W) \\ {\scriptstyle\wedge}\searrow & & \nearrow {\scriptstyle\tilde{T}} \\ & \Lambda^r(V) & \end{array}$$

By construction $\tilde{T}(v_1 \wedge \ldots \wedge v_r) = Tv_1 \wedge \ldots \wedge Tv_r$. We will omit the tilde in the future, writing

$$T(v_1 \wedge \ldots \wedge v_r) = Tv_1 \wedge \ldots \wedge Tv_r$$

and consider that we have extended $T$ to $\Lambda^r(V)$. We have shown that a linear transformation $T : V \to W$ may be lifted to a transformation (also called T) from $T : \Lambda^r(V) \to \Lambda^r(W)$. For the case $r = 0$, which is not covered by the previous conditions, we make the extension of $T$ equal to the identity, which is easily seen to be consistent with our requirement that $\alpha \wedge v = \alpha v$.

We ask, what is the matrix of $T$ on $\Lambda^r(V)$? To this end, let $e_1, \ldots, e_n$ be a basis for $V$ and $f_1, \ldots, f_m$ be a basis for $W$, and

$$Te_j = \alpha^i_j f_i \qquad \text{for} j = 1, \ldots, n$$

so that the $m \times n$ matrix $(\alpha^i_j)$ is the matrix of $T$ in the bases $e_j$ of $V$ and $f_i$ of $W$.

Since $e_\sigma = e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)}$ $\sigma \in \mathcal{S}_{n,r}$ form a basis of $\Lambda^r(V)$, we have

$$\begin{aligned}
Te_\sigma &= Te_{\sigma(1)} \wedge \ldots \wedge Te_{\sigma(r)} \\
&= (\alpha^{i_1}_{\sigma(1)} f_{i_1}) \wedge \ldots \wedge (\alpha^{i_1}_{\sigma(r)} f_{i_r}) \\
&= \alpha^{i_1}_{\sigma(1)} \ldots \alpha^{i_r}_{\sigma(r)} f_{i_1} \wedge \ldots \wedge f_{i_r} \\
&= \sum_{\rho \in \mathcal{S}_{m,r}} \Big( \sum_{\pi \in \mathcal{S}_r} \alpha^{\rho(\pi^{-1}(1)) \ldots \rho(\pi^{-1}(r))}_{\sigma(1) \,\ldots\ldots\ldots\, \sigma(r)} \Big) f_{\rho(1)} \wedge \ldots \wedge f_{\rho(r)} \\
&= \sum_{\rho \in \mathcal{S}_{m,r}} \det \begin{pmatrix} \alpha^{\rho(1)}_{\sigma(1)} & \ldots & \alpha^{\rho(1)}_{\sigma(r)} \\ \ldots & \ldots & \ldots \\ \alpha^{\rho(r)}_{\sigma(1)} & \ldots & \alpha^{\rho(r)}_{\sigma(r)} \end{pmatrix} f_\rho \\
&= \sum_{\rho \in \mathcal{S}_{m,r}} \alpha^{\rho(1 \ldots r)}_{\sigma(1 \ldots r)} f_\rho \\
&= \sum_{\rho \in \mathcal{S}_{m,r}} \alpha^\rho_\sigma \, f_\rho
\end{aligned}$$

where we are using the abbreviations for the subdeterminants of $(\alpha^i_j)$ introduced in section 3.3. The fourth equation is derived from the third equation by the method of resolving a sum by permutations also introduced in section 3.3.

Thus we see that the matrix coefficients for the extension of $T$ to $\Lambda^r(V)$ are the subdeterminants of order $r$ of the matrix of $T$.

It now seems reasonable in view of the form of the last equation to use the summation convention for summing over increasing permutations, so the last equation, using this convention, can be rewritten as

$$Te_\sigma = \alpha^\rho_\sigma\, f_\rho\,.$$

For consistency, we remark that the former type of summation convention

$$Te_j = \alpha^i_j\, f_i$$

may be looked on as a special case of the more general indexing by increasing permutations, in which we take $i$ to stand for the $\sigma \in \mathcal{S}_{1,m}$ given by

$$\left(\begin{array}{c|cccccc} 1 & 2 & \dots & i & i+1 & \dots & m \\ i & 1 & \dots & i-1 & i+1 & \dots & m \end{array}\right).$$

This interpretation of subdeterminants as elements in the matrix of the linear transformation $T\colon \Lambda^r(V) \to \Lambda^r(W)$ may be used to prove the Cauchy-Binet theorem relating the determinants of a product to the subdeterminants of each of the factors. Indeed, Let $S\colon U \to V$ and $T\colon V \to W$. Let $g_1, \dots g_p$ be a basis for $U$, $e_1, \dots e_n$ a basis for $V$, and $f_1, \dots f_m$ a basis for $W$. We form the matrices for $S$ and $T$:

$$Sg_k = \beta^i_k\, e_i \quad \text{and} \quad Te_i = \alpha^j_i\, f_j$$

If we now set $(\gamma^k_j) = (\alpha^j_i)(\beta^i_k)$ then $(\gamma^k_j)$ will be the matrix of $TS$ in the bases $g_1, \dots g_p$ of $V$ and $f_1, \dots f_m$ of $W$. Going over to the spaces $\Lambda^r(U)$, $\Lambda^r(V)$ and $\Lambda^r(W)$, the above analysis shows that (summation convention active)

$$\begin{array}{rcll} (TS)g_\sigma &=& \gamma^\rho_\sigma f_\rho & \text{for} \quad \rho \in \mathcal{S}_{m,r}, \quad \sigma \in \mathcal{S}_{p,r} \\ Sg_\sigma &=& \beta^\tau_\sigma e_\tau & \text{for} \quad \tau \in \mathcal{S}_{n,r} \\ Te_\tau &=& \alpha^\rho_\tau f_\rho & \end{array}$$

so that

$$\begin{array}{rcl} \gamma^\rho_\sigma f_\rho &=& (TS)g_\sigma = T(Sg_\sigma) = T(\beta^\tau_\sigma e_\tau) = \beta^\tau_\sigma T e_\tau \\ &=& \beta^\tau_\sigma \alpha^\rho_\tau f_\rho = (\alpha^\rho_\tau \beta^\tau_\sigma) f_\rho \end{array}$$

and, since $f_\rho$ is a basis of $\Lambda^r(W)$,

$$\gamma^\rho_\sigma = \alpha^\rho_\tau \beta^\tau_\sigma$$

This is the Cauchy Binet theorem: The minors of the product of two matrices are the sums of products of the minors of the matrices.

There are certain rules useful for the computation of determinants which we state below. These are simply restatements of theorems and a corollary from the previous section.

**Type I** Multiplying a single column of a square matrix by a scalar results in multiplying the determinant of the matrix by the same scalar.

**Type II** Interchanging the columns of a matrix changes the sign of the determinant.

**Type III** Adding a multiple of a column to another column does not effect the value of a determinant.

We are restating these rules here because we now want to prove that these rules are valid also for rows. This is most easily verified by showing that the determinant of a square matrix is equal to the determinant of its transpose.

**Def** If a matrix $A = (\alpha^i_j)$, then the transpose of $A$, denoted by $A^\top$ is given by

$$\det A^\top = (\beta^j_i) \qquad \text{where} \quad \beta^j_i = \alpha^i_j$$

We now have the extremely important theorem

**Theorem** $\det(A^\top) = \det(A)$

I do not know of a proof of this theorem in the present context which uses the properties of Grassmann products in an intelligent manner. The following proof, while certainly adequate, uses the properties of permutations in an inelegant way. To clarify the proof, we will look at an example first. Let $r = 4$ and

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \qquad \text{so that} \quad \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

Then a typical term of the determinant calculation will be

$$\text{sgn}(\pi)\, \beta^1_{\pi(1)} \beta^2_{\pi(2)} \beta^3_{\pi(3)} \beta^4_{\pi(4)} = \text{sgn}(\pi)\, \beta^1_2 \beta^2_4 \beta^3_1 \beta^4_3$$

and we can rearrange the terms so that they come out

$$\text{sgn}(\pi)\, \beta^3_1 \beta^1_2 \beta^4_3 \beta^2_4 = \text{sgn}(\pi)\, \beta^{\pi^{-1}(1)}_1 \beta^{\pi^{-1}(2)}_2 \beta^{\pi^{-1}(3)}_3 \beta^{\pi-1(4)}_4$$

This is done in the general case in the following proof of the theorem.

**Proof of the Theorem** By definition

$$\det(A^\top) = \sum_{\pi \in \mathcal{S}_r} \text{sgn}(\pi)\, \beta^1_{\pi(1)} \ldots \beta^r_{\pi(r)}$$

Rearranging the elements in the product as we did in the example we have

$$\begin{aligned} \det(A^\top) &= \sum_{\pi \in \mathcal{S}_r} \text{sgn}(\pi)\, \beta^{\pi^{-1}(1)}_1 \ldots \beta^{\pi^{-1}(r)}_r \\ &= \sum_{\pi \in \mathcal{S}_r} \text{sgn}(\pi^{-1})\, \beta^{\pi^{-1}(1)}_1 \ldots \beta^{\pi^{-1}(r)}_r && \text{since } \text{sgn}(\pi^{-1}) = \text{sgn}(\pi) \\ &= \sum_{\tau \in \mathcal{S}_r} \text{sgn}(\tau)\, \beta^{\tau(1)}_1 \ldots \beta^{\tau(r)}_r && \text{where } \tau = \pi^{-1} \end{aligned}$$

because $\tau = \pi^{-1}$ runs through $\mathcal{S}_r$ when $\pi$ does. But then

$$\begin{aligned}\det(A^\top) &= \sum_{\tau\in\mathcal{S}_r} \operatorname{sgn}(\tau)\, \beta_1^{\tau(1)} \dots \beta_r^{\tau(r)} \\ &= \sum_{\tau\in\mathcal{S}_r} \operatorname{sgn}(\tau)\, \alpha_{\tau(1)}^1 \dots \alpha_{\tau(r)}^r \\ &= \det(A).\end{aligned}$$

It should now be obvious that the above three rules of computation are valid for rows as well as columns, because the rows of $A$ are the columns of $A^\top$.

## 4.6 The Equivalence of Axioms 4a and 4b

In this section we are going to prove the equivalence of the three natural ways of guaranteeing that the Grassmann Algebra is as large as possible consistent with bilinearity and anticommutativity. This is not difficult, but also not very interesting, so the reader might want to consider carefully reading over the material until the proofs start and reserving the perusal of the proofs to a second reading.

For convenience of reference we repeat here Axioms 4a and 4b and in addition Axiom 4a* which is a variant of Axiom 4a and will be included in the equivalence discussion.

**Axiom 4a** Let $G(v_1, \ldots, v_r)$ be an alternating multilinear function from $V \times \ldots \times V \to W$. Let $\Phi$ be the map $\Phi\colon V \times \ldots \times V \to \Lambda^r(V)$ given by

$$\Phi(v_1, \ldots, v_r) \to v_1 \wedge \ldots \wedge v_r$$

(which is alternating and multilinear by Axioms 2 and 3.) Then there exists a unique map $\tilde{G}\colon \Lambda^r(V) \to W$ so that the following diagram commutes:

$$\begin{array}{ccc} V_1 \times \ldots \times V_n & \xrightarrow{G} & W \\ {}_{\Phi}\searrow & & \nearrow_{\tilde{G}} \\ & \Lambda^r(V) & \end{array}$$

The commutativity of the the diagram says that $G = \tilde{G} \circ \Phi$.

We next present Axiom 4a* which differs from Axiom 4a only by having the range of the multilinear function be the field over which the vectors spaces are defined instead of the vector space $W$. This looks weaker than Axiom 4a but is really almost trivially equivalent.

**Axiom 4a*** Let $G(v_1, \ldots, v_r)$ be an alternating multilinear function from $V \times \ldots \times V \to$ Field. Let $\Phi$ be the map $\Phi\colon V \times \ldots \times V \to \Lambda^r(V)$ given by

$$\Phi(v_1, \ldots, v_r) \to v_1 \wedge \ldots \wedge v_r$$

(which is alternating and multilinear by Axioms 2 and 3.) Then there exists a unique map $\tilde{G}\colon \Lambda^r(V) \to$ Field so that the following diagram commutes:

$$\begin{array}{ccc} V_1 \times \ldots \times V_n & \xrightarrow{G} & \text{Field} \\ {}_{\Phi}\searrow & & \nearrow_{\tilde{G}} \\ & \Lambda^r(V) & \end{array}$$

The commutativity of the the diagram says that $G = \tilde{G} \circ \Phi$.

Finally, we present again Axiom 4b

**Axiom 4b** The set of all products

$$e_{i_1} \wedge \ldots \wedge e_{i_r} \qquad \begin{cases} 1 \le r \le n \\ i_1 < i_2 \ldots < i_r \end{cases}$$

is linearly independent, where $e_1, \dots, e_n$ is a basis for $V$.

We now begin the proof of the equivalence of the three Axioms. The scheme will be Axiom 4b $\Longrightarrow$Axiom 4a* $\Longrightarrow$Axiom4a $\Longrightarrow$Axiom4b. Persons without interest in these details may now safely skip to the next section, as the material worked through in these proofs in not of great interest for applications.

**Proof that Axiom 4b $\Longrightarrow$Axiom 4a***

Let $F(v_1, \dots . v_r)$ be an alternating multilinear function from $V$ to the Field. It should be clear from previous work that

$$F(v_{\pi(1)}, \dots, v_{\pi(r)}) = \text{sgn}(\pi) F(v_1, \dots, v_r) \qquad \text{for } \pi \in \mathcal{S}_r$$

(This is so because $\text{sgn}(\pi)$ is $-1$ raised to the power equal to the number of interchanges necessary to restore $\pi(1), \dots, \pi(r)$ to $1, \dots, r$.) We now define $\tilde{F}\colon \Lambda^r(V) \to$ Field by defining it on a *basis* of $\Lambda^r(V)$. By Axiom 4b such a basis is $e_\sigma = e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(r)}, \quad \sigma \in \mathcal{S}_{n,r}$. Define $\tilde{F}$ by

$$\tilde{F}(e_\sigma) = \tilde{F}(e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(r)}) = F(e_{\sigma(1)}, \dots, e_{\sigma(r)})$$

Since $\tilde{F}$ is defined on the basis elements, it is uniquely defined on all of $\Lambda^r(V)$. It remains to show that for any $v_1, \dots, v_r \in V$, we have $F(v_1, \dots, v_r) = \tilde{F}(v_1 \wedge \dots \wedge v_r)$. Let $v_j = \alpha^i_j\, e_i$. Then

$$\begin{aligned} F(v_1, \dots, v_r) &= F(\alpha^{i_1}_1 e_{i_1}, \dots, \alpha^{i_r}_r e_{i_r}) \\ &= \alpha^{i_1}_1, \dots, \alpha^{i_r}_r F(e_{i_1}, \dots, e_{i_r}) \end{aligned}$$

We now resolve the sum by permutations as in Section 3.3

$$F(v_1, \dots, v_r) = \sum_{\sigma \in \mathcal{S}_{n,r}} \Big( \sum_{\pi \in \mathcal{S}_r} \alpha_1^{\sigma(\pi^{-1}(1))}, \dots, \alpha_r^{\sigma(\pi^{-1}(r))} \text{sgn}(\pi) \Big) F(e_{\sigma(1)}, \dots, e_{\sigma(r)})$$

where we have rearranged the arguments of $F(e_{i_1}, \dots, e_{i_r})$ into $\text{sgn}(\pi) F(e_{i_{\pi(1)}}, \dots, e_{r_{\pi(r)}}) = \text{sgn}(\pi) F(e_{\sigma(1)}, \dots, e_{\sigma(r)})$ so that the indices of the arguments increase. We now replace $F(e_{\sigma(1)}, \dots, e_{\sigma(r)})$ by $\tilde{F}(e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(r)})$ and, reversing the process, we get

$$\begin{aligned} F(v_1, \dots, v_r) &= \sum_{\sigma \in \mathcal{S}_{n,r}} \Big( \sum_{\pi \in \mathcal{S}_r} \alpha_1^{\sigma(\pi^{-1}(1))}, \dots, \alpha_r^{\sigma(\pi^{-1}(r))} \text{sgn}(\pi) \Big) \tilde{F}(e_{\sigma(1)} \wedge \dots \wedge e_{\sigma(r)}) \\ &= \alpha^{i_1}_1, \dots, \alpha^{i_r}_r \tilde{F}(e_{i_1} \wedge \dots \wedge e_{i_r}) \\ &= \tilde{F}(\alpha^{i_1}_1 e_{i_1} \wedge \dots \wedge \alpha^{i_r}_r e_{i_r}) \\ &= \tilde{F}(v_1 \wedge \dots \wedge v_r) \end{aligned}$$

as desired. Thus Axiom 4b implies Axiom 4a*.

**Proof that Axiom 4a* $\Longrightarrow$Axiom 4a**

Assume that we have a multilinear function $G: V \times \ldots \times V \to W$. Let $f_1, \ldots, f_m$ be a basis for $W$ and let $f^1, \ldots, f^m$ be the dual basis for $W$, which means that if $w \in W$ and $w = \beta^j f_j$ then $f^i(w) = \beta^i$. Each $f^i$ is a linear functional from $W$ to the Field, and $G^i = f^i \circ G$ for each $i$ is then a multilinear functional from $V, \ldots, V$ to the Field.

By Axiom 4a* there is a linear functional $\tilde{G}^i$ from $V \wedge \ldots \wedge V$ to the Field satisfying

$$\tilde{G}^i(v_1 \wedge \ldots \wedge v_r) = G^i(v_1, \ldots, v_r).$$

We may now reconstitute $\tilde{G}$ from the $\tilde{G}^i$ as follows. Set

$$\tilde{G}(v_1 \wedge \ldots \wedge v_r) = \tilde{G}^i(v_1 \wedge \ldots \wedge v_r) f_i.$$

$\tilde{G}$ is obviously a linear function from $V \wedge \ldots \wedge V$ to $W$. Then we have

$$\begin{aligned} G(v_1, \ldots, v_r) &= f^i\big(G(v_1, \ldots, v_r)\big) f_i \\ &= \big((f^i \circ G)(v_1, \ldots, v_r)\big) f_i \\ &= G^i(v_1, \ldots, v_r) f_i \\ &= \tilde{G}^i(v_1 \wedge \ldots \wedge v_r) f_i \\ &= \tilde{G}(v_1 \wedge \ldots \wedge v_r) \end{aligned}$$

as desired. We have shown that Axiom 4a* implies Axiom 4a.

**Proof that Axiom 4a $\Longrightarrow$Axiom 4b**

Before doing the proof itself, we will prove a lemma which has a certain interest of its own.

**Lemma** Given any $r$ linear functionals $f^i: V \to$ Field where $1 \le i \le r$ we can construct an alternating multilinear functional $F: V \times \ldots \times V \to$ Field and, given Axiom 4a, a linear functional $\tilde{F}: V \wedge \ldots \wedge V \to$ Field by

$$F(v_1, \ldots, v_r) = \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) f^1(v_{\pi(1)}) f^2(v_{\pi(2)}) \ldots f^r(v_{\pi(r)})$$

and then

$$\tilde{F}(v_1 \wedge \ldots \wedge v_r) = F(v_1, \ldots, v_r)$$

**Proof** It is clear that the $F$ given by the above formula is multilinear, so the only thing left to prove is that $F$ is alternating. Indeed, let $\sigma \in \mathcal{S}_r$. Then

$$\begin{aligned} F(v_{\sigma(1)}, \ldots, v_{\sigma(r)}) &= \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) f^1(v_{\pi(\sigma(1))}) f^2(v_{\pi(\sigma(2))}) \ldots f^r(v_{\pi(\sigma(r))}) \\ &= \operatorname{sgn}(\sigma) \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi\sigma) f^1(v_{\pi(\sigma(1))}) f^2(v_{\pi(\sigma(2))}) \ldots f^r(v_{\pi(\sigma(r))}) \\ &= \operatorname{sgn}(\sigma) \sum_{\rho \in \mathcal{S}_r} \operatorname{sgn}(\rho) f^1(v_{\rho(1)}) f^2(v_{\rho(2)}) \ldots f^r(v_{\rho(r)}) \\ &= \operatorname{sgn}(\sigma) F(v_1, \ldots, v_r) \end{aligned}$$

because $\pi\sigma$ runs through all the elements of $\mathcal{S}_r$ once when $\pi$ runs through all the elements of $\mathcal{S}_r$ once. This suffices to show that $F$ is alternating.

Now let $e_1, \ldots, e_r$ be a basis of $V$ and $e^1, \ldots, e^r$ its dual basis. We must show that the $e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)}, \quad \sigma \in \mathcal{S}_{n,r}$ are linearly independent in $\Lambda^r(V)$. (This is Axiom 4b). To this end, assume

$$\sum_{\sigma \in \mathcal{S}_{n,r}} a^\sigma e_\sigma = \sum_{\sigma \in \mathcal{S}_{n,r}} a^{\sigma(1)\ldots\sigma(r)} e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)} = 0$$

We must prove that all $a^\sigma$ are 0. We form the alternating linear functional (as in the Lemma)

$$F^\rho(v_1, \ldots, v_r) = \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) e^{\rho(1)}(v_{\pi(1)}) e^{\rho(2)}(v_{\pi(2)}) \ldots e^{\rho(r)}(v_{\pi(r)})$$

with some fixed $\rho \in \mathcal{S}_{n,r}$. By Axiom 4a, there exists a unique $\tilde{F}^\rho \colon \Lambda^r(V) \to \text{Field}$ satisfying

$$\tilde{F}^\rho(v_1 \wedge \ldots \wedge v_r) = F^\rho(v_1, \ldots, v_r)$$

for all $v_1, \ldots, v_r \in V$. Applying this to the supposed linear dependence, we have

$$\begin{aligned}
0 &= \tilde{F}^\rho\big( \sum_{\sigma \in \mathcal{S}_{n,r}} a^\sigma e_\sigma \big) \\
&= \sum_{\sigma \in \mathcal{S}_{n,r}} a^\sigma \tilde{F}^\rho(e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)}) \\
&= \sum_{\sigma \in \mathcal{S}_{n,r}} a^\sigma F^\rho(e_{\sigma(1)}, \ldots, e_{\sigma(r)}) \\
&= \sum_{\sigma \in \mathcal{S}_{n,r}} a^\sigma \Big( \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) e^{\rho(1)}(e_{\sigma(\pi(1))}) e^{\rho(2)}(e_{\sigma(\pi(2))}) \ldots e^{\rho(r)}(e_{\sigma(\pi(r))}) \Big) \\
&= a^\rho
\end{aligned}$$

because the interior sum in the next to the last equality will be 0 unless $\rho(k) = \sigma(\pi(k))$ for $k = 1, \ldots, r$ by the definition of the dual basis. This can only occur if $\pi =$ identity (so that $\sigma(\pi(k))$ will increase with k as $\rho$ does) and then $\sigma = \rho$ (because $\sigma, \rho \in \mathcal{S}_{n,r}$ are both determined by the first $r$ values). But in this single nonzero case, the interior sum is equal to 1. This completes the the proof that Axiom 4a implies Axiom 4b.

## 4.7 Products and Linear Dependence

In this section we are going to develop certain relationships between products in the Grassmann algebra, and indicate the connections between products and to linear algebra concepts of independence and span.

Let us fix some notations which, with minor later modifications, will be in effect for the remainder of the book. We define

**Def** An element of the Grassmann algebra $\Lambda(V)$ is simple if and only if it can be expressed as a product of vectors $v_1, \ldots, v_r$. When we discuss Grassmann algebra in general, we will use the upper case Latin letters $F, G, H$ to denote simple elements.

Of course, not all elements are simple. Let $e_1, e_2, e_3, e_4$ be a basis of $\mathbf{R^4}$. The the element

$$e_1 \wedge e_2 + e_3 \wedge e_4$$

is not simple.

Furthermore, in general elements of the Grassmann algebra will be denoted the upper case Latin letters $A, B, C, D$. These elements are, of course, sums of simple elements.

**Def** The degree of a simple element is the of vectors in the product:

$$\deg(v_1 \wedge v_2 \wedge \ldots \wedge v_r) = r$$

**Def** An element $A \in \Lambda(V)$ is homogeneous if all the terms (simple summands) of $A$ have the same degree $r$, in which case $A \in \Lambda(V)$.

The following theorem is basic to the application of Grassmann algebra

**Theorem** $v_1 \wedge v_2 \wedge \ldots \wedge v_r = 0 \Longleftrightarrow$ the set of vectors $\{v_1, \ldots, v_r\}$ is linearly dependent.

**Proof** $\Leftarrow$: Suppose, for example, that $v_r = \sum_{i=1}^{r-1} v_i$. Then

$$\begin{aligned} v_1 \wedge \ldots \wedge v_r &= v_1 \wedge \ldots \wedge v_{r-1} \wedge \Big(\sum_{i=1}^{r-1} \alpha^i v_i\Big) \\ &= \sum_{i=1}^{r-1} \alpha^i v_1 \wedge \ldots \wedge v_{r-1} \wedge v_i \\ &= 0 \end{aligned}$$

because the sum contains a repeated factor.

$\Rightarrow$: Suppose that the set $\{v_1, \ldots, v_r\}$ is linearly independent. The $v_1, \ldots, v_r$ may be extended to a basis $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$ of $V$. Axiom 4b then guarantees that the products

$$v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(r)}, \qquad \sigma \in \mathcal{S}_{n,r}$$

form a basis of $\Lambda^r(V)$. But setting $\sigma$ equal to the identity, we see that $v_1 \wedge \ldots \wedge v_r$ is a basis element, and hence cannot be 0.

We now note a trivial but important corollary.

**Corollary** Suppose $\dim(V) = n$ and $r < n$ and $\deg(F) = r$ where $F \neq 0$ is simple. Let $s$ be given with $s \leq n - r$. The there is a simple $G$ with $\deg(G) = s$ and $F \wedge G \neq 0$.

**Proof** Let $F = v_1 \wedge \ldots \wedge v_r \neq 0$. Then by the previous theorem $v_1, \ldots, v_r$ are linearly independent. Complete $v_1, \ldots, v_r$ to a basis with vectors $v_{r+1}, \ldots, v_n$. Then $v_1 \wedge \ldots \wedge v_n \neq 0$. Let $G = v_{r+1} \wedge \ldots \wedge v_{r+s}$. We have $F \wedge G = v_1 \wedge \ldots \wedge v_{r+s}$ and $r + s \leq r + n - r = n$ and thus $F \wedge G \neq 0$ since $v_1 \wedge v_n \neq 0$.

# Chapter 5

# Grassmann Algebra on the Space $V^*$

## 5.1 Introduction

In this chapter we develop Grassmann Algebra on the conjugate or dual space. Grassmann Algebra derives much of its power from the interaction of $\Lambda^r(V)$ with its dual which can be identified with $\Lambda^r(V^*)$. The fundamental interrelationship is through a certain determinant. We will call this result Grassmann's theorem, although from some points of view it is more like a definition than a theorem. It is

**Theorem** If $v_1, \dots, v_r \in V$ and $f^1, \dots, f^r \in V^*$ then the action of $f^1 \wedge \dots \wedge f^r \in \Lambda^r(V^*)$ on $v_1 \wedge \dots \wedge v_r \in \Lambda^r(V)$ is given by

$$(f^1 \wedge \dots \wedge f^r)(v_1 \wedge \dots \wedge v_r) = \det \begin{pmatrix} f^1(v_1) & f^1(v_2) & \dots & f^1(v_r) \\ f^2(v_1) & f^2(v_2) & \dots & f^2(v_r) \\ \dots & \dots & \dots & \dots \\ f^r(v_1) & f^r(v_2) & \dots & f^r(v_r) \end{pmatrix}$$

How do we prove this result? We will look at three ways of deriving this result in the next three sections. First, we may more or less define it to be true. Second, we may derive the result from previous results on tensor products. Thirdly we may derive it by specifying an action of the dual basis of $\Lambda^r(V^*)$ on a basis of $\Lambda^r(V)$.

Because of the way we write elements of $V^*$ there are some matrix trivialities to discuss. For $f^i \in V^*$ we write $f^i = \beta^i_j e^i, \quad i = 1, \dots, r$, where $\{e^j\}$ is the dual basis in $V^*$ of the basis $\{e_i\}$ of $V$. To maintain consistency with matrix notation, the coefficients of the $f^i$ are thought of as *rows* of the matrix

$$\begin{pmatrix} \beta^1_1 & \beta^1_2 & \dots & \beta^1_n \\ \beta^2_1 & \beta^2_2 & \dots & \beta^2_n \\ \dots & \dots & \dots & \dots \\ \beta^1_r & \beta^r_2 & \dots & \beta^r_n \end{pmatrix}$$

and the expression of $f^1 \wedge \dots \wedge f^r$ would then, if we repeated the analysis of section 3.3, come out as

$$\sum_{\sigma \in \mathcal{S}_{n,r}} \sum_{\pi \in \mathcal{S}_r} \beta^1_{\sigma(\pi^{-1}(1))} \cdots \beta^r_{\sigma(\pi^{-1}(r))} f^{\sigma(1)} \wedge \dots \wedge f^{\sigma(r)}$$

where the permutations are now acting on the *lower* indices. Fortunately, in view of the fact that $\det(A^\top) = \det(A)$, it is irrelevant whether the permutations act on upper or lower indices, so that all the determinental identities remain valid for $V^*$ as for $V$.

## 5.2 Grassmann's Theorem by Fiat

Nothing prevents us from taking Grassmann's theorem as simply a definition of the action of $\Lambda^r(V^*)$ on $\Lambda^r(V)$. However, it would be necessary to show that the action is well defined;

$$\text{if} \quad v_1 \wedge \ldots \wedge v_r = w_1 \wedge \ldots \wedge w^r \quad \text{then} \quad (f^1 \wedge \ldots \wedge f^r)(v_1 \wedge \ldots \wedge v_r) = (f^1 \wedge \ldots \wedge f^r)(w_1 \wedge \ldots \wedge w_r)$$

and similarly for $f^1 \wedge \ldots \wedge f^r = g^1 \wedge \ldots \wedge g^r$. This would be extremely tedious, and to reduce the tedium authors who use this method generally omit the verification.

A much better method is as follows. Define a function

$$(f^1 \wedge \ldots \wedge f^r)(v_1, \ldots, v_r) = \det(f^i(v_j))$$

and show, which is trivial, that it is an alternating bilinear function on $V \times \ldots \times V$. By Axiom 4a, this induces the map required by Grassmann's theorem. This still leaves the question of whether the map is well defined on $f^1 \wedge \ldots \wedge f^r$, but this becomes obvious if we remember that $V^{**} = V$, so that the roles of $f^i$ and $v_j$ may be interchanged. We may now regard Grassmann's theorem as proved, but for many a sense of unease will remain, in that no *derivation* of Grassmann's theorem has been provided. While this would not be a drawback in, for example, number theory, it is more uncomfortable in Linear Algebra or Differential Geometry. Therefore, in the next two sections we will present two methods for deriving Grassmann's theorem. It is worth noting that already we know that

$$\dim \Lambda^r(V^*) = \binom{n}{r} = \dim \Lambda^r(V)$$

and this is sufficient to guarantee that there is an isomorphism of $\Lambda^r(V^*)$ and $\Lambda^r(V)^*$.

## 5.3 Grassmann's Theorem by Tensor Products

The easiest derivation of Grassmann's theorem is by the use of tensor products, since we already know the action of $f^1 \otimes \dots \otimes f^r$ on $v_1 \otimes \dots \otimes v_r$ which is

$$(f^1 \otimes f^2 \otimes \dots \otimes f^r)(v_1 \otimes v_2 \otimes \dots \otimes v_r) = f^1(v_1) f^2(v_2) \dots f^r(v_r).$$

The drawback is that this method is not available if Grassmann products are approached without use of tensor products, as for example, would be done in an axiomatic treatment.

We recall that we left unspecified the multiplicative function $S(r!)$ in the definition

$$v_1 \wedge \dots \wedge v_r = \frac{S(r!)}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) v_{\pi(1)} \otimes \dots \otimes v_{\pi(r)} \tag{1}$$

At this point, as we will see, the most natural choice is to set

$$S(r!) = \sqrt{r}!$$

and thus

$$v_1 \wedge \dots \wedge v_r = \tfrac{1}{\sqrt{r!}} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) v_{\pi(1)} \otimes \dots \otimes v_{\pi(r)} \tag{2}$$

We will then have

$$\begin{aligned}
(f^1 \wedge \dots \quad \wedge f^r)(v_1 \wedge \dots \wedge v_r) & \\
= \quad & \Big(\tfrac{1}{\sqrt{r!}} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) f^{\pi(1)} \otimes \dots \otimes f^{\pi(r)}\Big)\Big(\tfrac{1}{\sqrt{r!}} \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(r)}\Big) \\
= \quad & \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) (f^{\pi(1)} \otimes \dots \otimes f^{\pi(r)})(v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(r)}) \\
= \quad & \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) f^{\pi(1)}(v_{\sigma(1)}) f^{\pi(2)}(v_{\sigma(2)}) \dots f^{\pi(r)}(v_{\sigma(r)}) \\
= \quad & \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) \det \begin{pmatrix} f^{\pi(1)}(v_1) & \dots & f^{\pi(1)}(v_r) \\ \dots & \dots & \dots \\ f^{\pi(r)}(v_1) & \dots & f^{\pi(r)}(v_r) \end{pmatrix} \\
= \quad & \tfrac{1}{r!}\, r! \det \begin{pmatrix} f^1(v_1) & \dots & f^1(v_r) \\ \dots & \dots & \dots \\ f^r(v_1) & \dots & f^r(v_r) \end{pmatrix} \\
= \quad & \det(f^i(v_j))
\end{aligned}$$

Notice here the critical role played by the factor $\frac{1}{\sqrt{r!}}$. Its presence in both definitions of Grassmann products in terms of tensor products contributes the critical factor $\frac{1}{r!}$ which cancels out near the end. To be more precise, if we wish to use a formula like (2) to define Grassmann products in terms of tensor products, and if we wish to use the *same* factor for both the space $V$ and the dual space $V^*$, then the choice of $S(r!)$ in formula (1) must be $S(r!) = \sqrt{r!}$

## 5.4 Grassmann's Theorem by Use of a Basis

To derive Grassmann's theorem in a plausible way without tensor products we may proceed by insisting that the relationship between basis and dual basis persist in $\Lambda^r(V)$ and $\Lambda^r(V^*)$ in analogy to the way it works in $\Lambda^1(V) = V$ and $\Lambda^1(V^*) = V^*$ Essentially, we are again defining Grassmann's theorem to be true, but we are doing so in a more plausible way than in section 5.2.

Let $e_\sigma, \quad \sigma \in \mathcal{S}_{n,r}$ be a basis for $\Lambda^r(V)$ and $e^\tau, \quad \tau \in \mathcal{S}_{n,r}$ be a basis for $\Lambda^1(V^*)$, where

$$e^\tau = e^{\tau(1)} \wedge \ldots \wedge e^{\tau(r)}$$

and $e^1, e^2, \ldots, e^n \in V^*$ is the dual basis to $e_1, e_2, \ldots, e_n \in V$. By definition of the dual basis we have

$$e^i(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

We regard the index $i$ to be a way of writing

$$i = \left( \begin{array}{c|cccccc} 1 & 2 & \ldots & i & i+1 & \ldots & n \\ i_1 & 1 & \ldots & i-1 & i+1 & \ldots & n \end{array} \right)$$

and similarly for $j$, so that a reasonable generalization becomes

$$e^\sigma(e_\tau) = \begin{cases} 1 & \text{if } \sigma = \tau \\ 0 & \text{if } \sigma = \tau \end{cases} \qquad \text{for} \quad \sigma, \tau \in \mathcal{S}_{n,r} \, .$$

This define the action of the basis element $e^\sigma \in \Lambda^r(V^*)$ on the basis element $e_\tau \in \Lambda^r(V)$, and hence defines $e^\sigma$ uniquely on all of $\Lambda^r(V)$. Once the action of the basis elements of $\Lambda^r(V^*)$ have been defined, the action of any element of $\Lambda^r(V^*)$ is uniquely defined by linearity.

With the action of $\Lambda^r(V^*)$ on $\Lambda^r(V)$ defined for basis elements, we may derive the general formulas as follows. Let $v_j = \alpha^i_j e_i$ and $f^l = \beta^l_k e^k$, where $j = 1, \ldots, n$ and $l = 1, \ldots, n$. Then

$$\begin{aligned}
(f^1 \wedge \ldots \wedge f^r)(v_1 \wedge \ldots \wedge v_r) &= \Big( \sum_{\sigma \in \mathcal{S}_{n,r}} \det(\beta^l_{\sigma(j)}) e^\sigma \Big) \Big( \sum_{\rho \in \mathcal{S}_{n,r}} \det(\alpha^{\rho(l)}_j) e_\rho \Big) \\
&= \sum_{\sigma, \rho \in \mathcal{S}_{n,r}} \det(\beta^l_{\sigma(j)}) \det(\alpha^{\rho(l)}_j) e^\sigma(e_\rho) \\
&= \sum_{\sigma \in \mathcal{S}_{n,r}} \det(\beta^l_{\sigma(j)}) \det(\alpha^{\sigma(l)}_j) \\
&= \det\big(\beta^l_i \alpha^i_j\big)
\end{aligned}$$

by the Cauchy–Binet theorem. We now note that

$$\begin{aligned}
\det\big(f^l(v_j)\big) &= \det\big((\beta^l_k e^k)(\alpha^i_j e_i)\big) \\
&= \det\big(\beta^l_k \alpha^i_j e^k(e_i)\big) \\
&= \det\big(\beta^l_i \alpha^i_j\big)
\end{aligned}$$

so that we have Grassmann's theorem

$$(f^1 \wedge \ldots \wedge f^r)(v_1 \wedge \ldots \wedge v_r) = \det\bigl(f^l(v_j)\bigr) .$$

## 5.5 The Space of Multilinear Functionals

We now wish to consider the space of $r$-multilinear functionals $F(x_1, \dots, x_r)$ on $V$. We first note that if $e_1, \dots, e_n$ is a basis for $V$ then , with $v_j = \alpha^i_j e_i$, we have

$$\begin{aligned} F(v_1, \dots, v_r) &= F(\alpha^{i_1} e_{i_1}, \dots, \alpha^{i_r} e_{i_r}) \\ &= \alpha^{i_1} \dots \alpha^{i_r} F(e_{i_1}, \dots, e_{i_r}) \,. \end{aligned}$$

Thus two things become obvious. First, a multilinear functional is completely determined by its values on $r$-tuples of basis elements for inputs. Second, if the values are specified on all $r$-tuples of basis elements as inputs then the above equation, using these values, will generate a multilinear functional.

Since the multilinear functionals are clearly a vector space, it is reasonable to seek a basis and determine the dimension. A basis is easily found; we set

$$F^{i_1 \dots i_r}(v_1, \dots, v_r) = \begin{cases} 1 & \text{if } v_j = e_{i_j}, \quad 1 \leq j \leq r \\ 0 & \text{for any other combination of basis-vector inputs.} \end{cases}$$

By the above formula, this determines a unique multilinear functional $F^{i_1 \dots i_r}$ and we may then write any multilinear functional $F$ in terms of the $F^{i_1 \dots i_r}$ by the following method. First, with $v_j = \alpha^i_j e_i$,

$$\begin{aligned} F^{i_1 \dots i_r}(v_1, \dots, v_r) &= \alpha^{j_1}_1 \dots \alpha^{j_r}_r F^{i_1 \dots i_r}(e_{j_1}, \dots, e_{j_r}) \\ &= \alpha^{i_1}_1 \dots \alpha^{i_r}_r \end{aligned}$$

since the term with subscripts $i_1, \dots, i_r$ is the only non-zero term. Then it follows that

$$\begin{aligned} F(v_1, \dots v_r) &= \alpha^{i_1}_1 \dots \alpha^{i_r}_r F(e_{i_1}, \dots, e_{i_r}) \\ &= F^{i_1 \dots i_r}(v_1, \dots, v_r) F(e_{i_1}, \dots, e_{i_r}) \,, \end{aligned}$$

from which we see that the $F^{i_1 \dots i_r}$ span the space of multilinear functionals. These are clearly linearly independent, for if we have a linear combination

$$\alpha_{i_1 \dots i_r} F^{i_1 \dots i_r} = 0$$

then, applying it to the arguments $e_{j_1}, \dots, e_{j_r}$, we will have

$$\alpha_{j_1 \dots j_r} = 0 \,.$$

We can now clearly see that the dimension of the space of multilinear functionals is $[\dim(V)]^r = n^r$.

Readers familiar with chapter 2 (tensor products) will have noticed a similarity between the space of multilinear functionals and $\bigotimes_{i=1}^r V^*$. Essentially, they are the same:

**Theorem** $\bigotimes_{i=1}^r V^*$ and the space of multilinear functionals are isomorphic.

**Proof** Consider the mapping $F : V^* \times \ldots \times V^* \to$ (space of multilinear functionals) given by

$$[F(f^1,\ldots,f^r)](v_1,\ldots,v_r) = f^1(v_1)f^2(v_2)\ldots f^r(v_r)\,.$$

$F$ is clearly multilinear. Hence by the fundamental theorem on tensor products, there is a map

$$\tilde{F} : V^* \otimes \ldots \otimes V^* \to \text{(space of multilinear functionals)}$$

so that

$$\tilde{F}(f^1 \otimes \ldots \otimes f^r) = F(f^1,\ldots,f^r)\,.$$

I claim $\tilde{F}$ is an isomorphism. For consider the image $F(e^{i_1} \otimes \ldots \otimes e^{i_r})$ of the element $e^{i_1} \otimes \ldots \otimes e^{i_r}$. We have

$$\tilde{F}(e^{i_1} \otimes \ldots \otimes e^{i_r})(v_1,\ldots,v_r) = e^{i_1}(v_1)e^{i_2}(v_2)\ldots e^{i_r}(v_r)\,.$$

This will be non-zero for precisely one set of basis vector inputs, namely $v_j = e_{i_j}$, and for that set of inputs it well be 1. Hence

$$\tilde{F}(e^{i_1} \otimes \ldots \otimes e^{i_r}) = F^{i_1\ldots i_r}$$

which we previously defined. Since these elements are a basis, and therefore generate the space of Linear functionals, $\tilde{F}$ is onto. But

$$\dim(\bigotimes_{i=1}^{r} V^*) = n^r = \dim(\text{space of multilinear functionals})$$

Hence $\tilde{F}$ is one-to-one, and thus an isomorphism

Since the Grassmann products $\Lambda^r(V^*)$ may be constructed inside $\bigotimes_{i=1}^r V^*$ as shown in chapter 2, it must, by the last theorem, be possible to construct a copy of $\Lambda^r(V^*)$ inside the space of $r$-multilinear functionals. Guided by the above isomorphism $\tilde{F}$, we have the correspondence

$$e^{i_1} \otimes \ldots \otimes e^{i_r} \overset{\tilde{F}}{\longleftrightarrow} F^{i_1\ldots i_r}\,.$$

But then, as shown in chapter 2, we may take

$$\begin{aligned}
e^{i_1} \wedge \ldots \wedge e^{i_r} &= \sqrt{r!}\,\Pi\,\mathrm{sgn}(\pi)\, e^{i_{\pi(1)}} \otimes \ldots \otimes e^{i_{\pi(r)}} \\
&= \sqrt{r!}\,\tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \mathrm{sgn}(\pi) e^{i_{\pi(1)}} \otimes \ldots \otimes e^{i_{\pi(r)}} \\
&\overset{\tilde{F}}{\longleftrightarrow} \sqrt{r!}\,\tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \mathrm{sgn}(\pi) F^{i_{\pi(1)}\ldots i_{\pi(r)}}
\end{aligned}$$

This suggests the introduction of an operator $\Pi$ on the space of multilinear functionals that duplicates the activity of $\Pi$ on $\bigotimes_{i=1}^{r} V^*$. We define

$$\Pi F(v_1, \dots, v_r) = \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) F(v_{\pi(1)}, \dots, v_{\pi(r)})$$

and note

$$\begin{aligned} \Pi F(v_{\sigma(1)}, \dots, v_{\sigma(r)}) &= \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) F(v_{\pi(\sigma(1))}, \dots, v_{\pi(\sigma(r))}) \\ &= \operatorname{sgn}(\sigma)\, \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma) F(v_{(\pi\sigma)(1)}, \dots, v_{(\pi\sigma)(r)}) \\ &= \operatorname{sgn}(\sigma)\, \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi\sigma) F(v_{(\pi\sigma)(1)}, \dots, v_{(\pi\sigma)(r)}) \\ &= \operatorname{sgn}(\sigma)\, \Pi F(v_1, \dots, v_r) \end{aligned}$$

since $\pi\sigma$ runs through all the permutations of $\mathcal{S}_r$ exactly once when $\pi$ runs through all the permutations of $\mathcal{S}_r$ exactly once.

Thus for any $F$, $\Pi\, F$ is an alternating multilinear functional. Moreover, all alternating multilinear functionals arise in this way since

$$\begin{aligned} \Pi\Pi F(v_1, \dots, v_r) &= \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi) \Pi F(v_{\pi(1)}, \dots, v_{\pi(r)}) \\ &= \tfrac{1}{r!} \sum_{\pi \in \mathcal{S}_r} \operatorname{sgn}(\pi)^2\, \Pi F(v_1, \dots, v_r) \\ &= \Pi F(v_1, \dots, v_r) \end{aligned}$$

The alternating multilinear functionals thus constitute a subspace, the range of $\Pi$, of the space of all multilinear functionals.

## 5.6 The Star Operator and Duality

The star operator sets up a duality between $\Lambda^p(V)$ and $\Lambda^{n-p}(V^*)$. It is one of Grassmann's most fundamental contributions because it algebraizes the principle of duality in projective geometry and elsewhere. It also expresses the dual tensor in differential geometry.

The star operator is not absolute; it requires some additional structure on $V$ to define. This can be done in a variety of ways, but the simplest seems to be to choose a basis for the 1–dimensional space $\Lambda^n(V^*)$ (where $n = \dim(V)$). Once this is chosen, the star operator is uniquely determined. If the basis of $\Lambda^n(V^*)$ is changed, the $*$ operator accumulates a constant factor but is otherwise unaffected. In the presence of a metric, more stability is possible, as we will discuss in the next chapter. Suppose we have selected a fixed basis element $\Omega^* \in \Lambda^n(V^*)$. Let $m \in \Lambda^{n-p}(V^*)$ be fixed and $l \in \Lambda^p(V^*)$. We have $l \wedge m \in \Lambda^n(V^*)$ which is 1–dimensional. Given our basis element $\Omega^*$ we have, for some element $f_m(l)$

$$l \wedge m = f_m(l)\Omega^*$$

where $f_m(l)$, as a function of $l \in \Lambda^p(V^*)$, is clearly a linear functional of $\Lambda^p(V^*)$ to the Field. Since the dual space of $\Lambda^p(V^*)$ can be identified with $\Lambda^p(V)$, we can find a element $v_m \in \Lambda^p(V)$ so that we have

$$\begin{aligned} f_m(l) &= \langle l, v_m\rangle \\ l \wedge m &= \langle l, v_m\rangle\Omega^* \end{aligned}$$

Now the map $m \mapsto f_m \mapsto v_m$ is clearly linear as a function from $\Lambda^{n-p}(V^*)$ to $\Lambda^p(V)$. We will show that this map is injective. Indeed, suppose that $m \mapsto 0$ for some $m$. Then $l \wedge m = 0$ for all $l$, so we know that $m = 0$ by section $(* * * * * * * * * **)$. Since $\dim \Lambda^p(V) = \binom{n}{p} = \dim \Lambda^{n-p}(V^*)$, the mapping $m \mapsto v_m$ is an isomorphism of $\Lambda^{n-p}(V^*)$ onto $\Lambda^p(V)$ and thus has an inverse $*\colon \Lambda^p(V) \to \Lambda^{n-p}(V^*)$ satisfying

$$l \wedge *v = \langle l, v\rangle\Omega^* \qquad \text{for} \quad l \in \Lambda^p(V^*), \quad v \in \Lambda^p(V)$$

which is the primary equation for the $*$ operator. Similarly, we can define an operator $*\colon \Lambda^p(V^*) \to \Lambda^{n-p}(V)$ defined by the dual equation

$$u \wedge *l = \langle l, u\rangle\Omega \qquad \text{for} \quad u \in \Lambda^p(V), \quad l \in \Lambda^p(V^*)$$

where $\Omega$ is a basis element of $\Lambda^n(V)$.

Either $*$ may be used in isolation with the selection of an appropriate $\Omega^*$ or $\Omega$. However, they will not interact properly unless we have the additional condition

$$\langle \Omega^*, \Omega\rangle = 1$$

*From now on we will always assume this condition is satisfied.*

Note that a choice of either $\Omega^*$ or $\Omega$ together with the above condition determines the other uniquely.

The $*$ operator depends on the choice of $\Omega^*$ (or $\Omega$). If a new choice is made, the $*$ operator will accumulate a constant factor.

Indeed, let us suppose that we make another choice $\tilde{\Omega}$ of $\Omega$, so that

$$\tilde{\Omega} = \kappa\Omega.$$

Then the condition $\langle \Omega^*, \Omega \rangle = 1$ will force

$$\tilde{\Omega}^* = \tfrac{1}{\kappa}\Omega^* .$$

Let $\tilde{*}$ be the corresponding $*$ operator for $\tilde{\Omega}^*$ and $\tilde{\Omega}$. Then

$$\begin{aligned} l \wedge \tilde{*}v &= \langle l, v\rangle \tilde{\Omega}^* \qquad \text{for} \quad l \in \Lambda^p(V*) \quad v \in \Lambda^p(V) \\ &= \langle l, v\rangle \tfrac{1}{\kappa}\Omega^* \\ &= \tfrac{1}{\kappa}\langle l, v\rangle \Omega^* \\ &= \tfrac{1}{\kappa}\, l \wedge *v \\ &= l \wedge (\tfrac{1}{\kappa} * v) \end{aligned}$$

which shows that

$$\tilde{*}v = \tfrac{1}{\kappa} * v, \qquad \text{for} \quad v \in \Lambda^p(V).$$

Similarly, for $l \in \Lambda^p(V^*)$ and $u \in \Lambda^p(V)$

$$\begin{aligned} u \wedge \tilde{*}l &= \langle l, u\rangle \tilde{\Omega} = \langle l, v\rangle \kappa\Omega = \kappa\langle l, v\rangle\Omega = \kappa\, u \wedge *l \\ &= u \wedge (\kappa * l) \end{aligned}$$

so that

$$\tilde{*}l = \kappa(*l) \qquad \text{for} \quad l \in \Lambda^p(V^*).$$

Now suppose we are given a basis $e_1, \ldots, e_n$ of $V$ and we set $\Omega = e_1 \wedge \ldots \wedge e_n$. Let $e^1, \ldots, e^n$ be the dual basis and set $\Omega^* = e^1 \wedge \ldots \wedge e^n$. Then

$$\langle \Omega^*, \Omega\rangle = \langle e^1 \wedge \ldots \wedge e^n, e_1 \wedge \ldots \wedge e_n\rangle = \det\big((e^i, e_j)\big) = 1$$

by Grassmann's theorem, so $\Omega^*$ and $\Omega$ satisfy the required condition $\langle \Omega^*, \Omega\rangle = 1$.

We now wish a formula for $*e_\sigma$ where $e_\sigma \in \mathcal{S}_{n,p}$ is a basis element of $\Lambda^p(V)$. We recall that for basis elements $e^\pi$, $\pi \in \mathcal{S}_{n,p}$ of $\Lambda^p(V^*)$ and $e^\rho$, $\rho \in \mathcal{S}_{n,n-p}$ of $\Lambda^{n-p}(V^*)$ we have

$$e^\pi \wedge e^\rho = \begin{cases} 0 & \text{for } e^\rho \neq e^{\tilde{\pi}} \\ \operatorname{sgn}(\pi)\Omega^* & \text{for } e^\rho = e^{\tilde{\pi}} \end{cases}$$

where $\tilde{\pi}$ is the reverse of $\pi$ (see section 3.2). Let (summation convention!) $*e_\sigma = \alpha_\rho e^\rho$. Then

$$\begin{aligned} \langle e^\pi, e_\sigma\rangle\Omega^* &= e^\pi \wedge *e_\sigma \qquad \pi, \sigma \in \mathcal{S}_{n,p} \\ &= e^\pi \wedge \alpha_\rho e^\rho \qquad \rho \in \mathcal{S}_{n,n-p} \\ \delta^\pi_\sigma\, \Omega^* &= \operatorname{sgn}(\pi)\alpha_{\tilde{\pi}}\Omega^* \end{aligned}$$

since the only value of $\rho$ in the second line to contribute a non zero value is $\rho = \tilde{\pi}$. From the last equation we see that $\alpha_{\tilde{\pi}} = 0$ when $\pi \neq \sigma$ and $\alpha_{\tilde{\pi}} = \text{sgn}(\pi)$ when $\pi = \sigma$. Thus, since $\tilde{\pi}$ runs through $\mathcal{S}_{n,n-p}$ when $\pi$ runs through $\mathcal{S}_{n,p}$ we have

$$*e_\sigma = \alpha_\rho e^\rho = \alpha_{\tilde{\pi}} e^{\tilde{\pi}} = \text{sgn}(\sigma) e^{\tilde{\sigma}}$$

which is the desired formula. Similarly

$$*e^\sigma = \text{sgn}(\sigma) e_{\tilde{\sigma}}\,.$$

Now we want to compute $**v$. Recall from our work on permutations in section 3.2 that if $T_n = \frac{n(n+1)}{2}$ is the $n^{th}$ triangular number then

$$T_p + T_{n-p} + p(n-p) = T_n$$

and

$$\text{sgn}(\pi) = (-1)^{\sum_{k=1}^{p} \pi(k) - T_p} \qquad \text{for} \quad \pi \in \mathcal{S}_{n,p}$$

and thus

$$\begin{aligned}
\text{sgn}(\pi)\text{sgn}(\tilde{\pi}) &= (-1)^{\sum_{k=1}^{p} \pi(k) - T_p} \times (-1)^{\sum_{l=1}^{n-p} \tilde{\pi}(l) - T_{n-p}} \\
&= (-1)^{\sum_{k=1}^{n} \pi(k) - T_p - T_{n-p}} \\
&= (-1)^{T_n - T_p - T_{n-p}} \\
&= (-1)^{p(n-p)}
\end{aligned}$$

From this we easily derive

**Theorem** The formulas for $**$ are

$$\begin{aligned}
**v &= (-1)^{p(n-p)} v \qquad \text{for} \quad v \in \Lambda^p(V) \\
**l &= (-1)^{p(n-p)} v \qquad \text{for} \quad l \in \Lambda^p(V^*)
\end{aligned}$$

**Proof** It suffices to prove the formula for $v$, the proof for $l$ being identical. It also suffices to prove the formula for basis elements $e_\pi$, $\pi \in \mathcal{S}_{n,p}$ since it will then follow for general $v$ by linearity.

$$\begin{aligned}
**e_\pi &= *\,\text{sgn}(\pi) e^{\tilde{\pi}} \\
&= \text{sgn}(\pi) * e^{\tilde{\pi}} \\
&= \text{sgn}(\pi)\text{sgn}(\tilde{\pi}) e_{\tilde{\tilde{\pi}}} \\
&= (-1)^{p(n-p)} e_\pi
\end{aligned}$$

It would be preferable to derive the formula $**v = (-1)^{p(n-p)}$ without the use of a basis, but I have not been able to find a way to do this. However, from the definitions the following related formulas are easily derivable:

$$\begin{aligned}
\langle *v, *m \rangle \Omega^* &= *v \wedge **m = (-1)^{p(n-p)} **m \wedge *v = (-1)^{p(n-p)} \langle **m, v \rangle \Omega^* \\
\langle *v, *m \rangle \Omega &= *m \wedge **v = (-1)^{p(n-p)} **v \wedge *m = (-1)^{p(n-p)} \langle m, **v \rangle \Omega
\end{aligned}$$

and thus

$$\langle *v, *m\rangle = (-1)^{p(n-p)}\langle * * m, v\rangle = (-1)^{p(n-p)}\langle m, * * v\rangle$$

**Corollary** For $v \in \Lambda^p(V)$ and $l \in \Lambda^p(V^*)$ we have

$$\langle *v, *l\rangle = \langle l, v\rangle$$

**Proof** From the above formula and the Theorem

$$\langle *v, *m\rangle = \langle m, (-1)^{p(n-p)} * *v\rangle = \langle m, v\rangle$$

In the foregoing discussion we have tacitly assumed that $1 \le p \le n-1$. We now wish to complete the discussion for the cases $p = 0$ and $p = n$. Recall that $\Lambda^0(V)$ is the Field of scalars, and that for $\alpha \in \Lambda^0(V)$, $A \in \Lambda^p(V)$, $0 \le p \le n$ we defined $\alpha \wedge A = A \wedge \alpha = \alpha A$. Also, $\langle \lambda, \alpha\rangle = \lambda\alpha$ for $\lambda \in \Lambda^0(V^*)$ by definition. Thus we can rewrite the basic equation

$$l \wedge *u = \langle l, u\rangle\Omega^* \qquad \text{for} \quad l \in \Lambda^p(V^*), \quad u \in \Lambda^p(V)$$

as

$$\begin{aligned} \lambda \wedge *\alpha &= \langle \lambda, \alpha\rangle\Omega^* \qquad \text{for} \quad \lambda \in \Lambda^0(V^*), \quad u \in \Lambda^0(V) \\ \lambda * \alpha &= \lambda\,\alpha\Omega^* \end{aligned}$$

which gives us that

$$*\alpha = \alpha\Omega^*$$

so

$$*1 = \Omega^*$$

and then

$$1 = (-1)^{0\cdot(n-0)} \cdot 1 = * * 1 = *\Omega^* \, .$$

Similarly, when 1 is regarded as a member of $\Lambda^0(V^*)$, we derive

$$*1 = \Omega \qquad \text{and} \qquad 1 = *\Omega.$$

Notice that $*1$ depends on which $*$ is being used.

The foregoing formula for $*$ on the basis elements, $*e_\sigma = \mathrm{sgn}(\sigma)e^{\tilde{\sigma}}$ is actually valid for a wider class of products than just the basis elements $e_\sigma$. In fact, we have for any $\sigma \in \mathcal{S}_n$

$$*(e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(p)} = \mathrm{sgn}(\sigma)e^{\sigma(p+1)} \wedge \ldots \wedge e^{\sigma(n)}$$

This formula may be derived by examining interchanges of elements, but the following technique is much more interesting and may have applications in other contexts.

The general concept is perhaps clearer if illustrated by an example. We take $n = 7$ and $p = 4$ and

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 2 & 6 & 3 & 7 & 1 & 5 \end{pmatrix}$$

Since $p = 4$ we will now arrange the first four elements in $\sigma$ in increasing order, and then the last three elements also in increasing order to get

$$\sigma = \begin{pmatrix} 2 & 4 & 1 & 3 & 6 & 7 & 5 \\ 2 & 3 & 4 & 6 & 1 & 5 & 7 \end{pmatrix}$$

Now the central point is the rearrangement of the top level of the permutation $\sigma$ can be accomplished by another permutation $\pi$ which clearly is (compare top line of last equation for $\sigma$ with outputs for $\pi$)

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 1 & 3 & 6 & 7 & 5 \end{pmatrix}$$

We then have

$$\tau = \sigma\pi = \left(\begin{array}{cccc|ccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 6 & 1 & 5 & 7 \end{array}\right) \in \mathcal{S}_{7,4}$$

because the first four and last three elements were arranged in increasing order, as required to be in $\mathcal{S}_{7,4}$. It is worth noting that we could have formed $\tau$ immediately from $\sigma$ by rearrangement of the first four and last three elements in increasing order, and then found $\pi$ as

$$\pi = \sigma^{-1}\tau.$$

The permutation $\pi$ has the interesting property that it exchanges the first four elements among themselves and last three elements among themselves. To see this, note that for $1 \le j \le 4$ we have $\sigma(\pi(j)) = \tau(j) \in \{\sigma(1), \sigma(2), \sigma(3), \sigma(4)\}$ so we must have $\pi(j) \in \{1, 2, 3, 4\}$, and similarly for $5 \le j \le 7$. Thus $\pi$ can be written as a product $\pi = \pi_1\pi_2$ where

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 1 & 3 & 5 & 6 & 7 \end{pmatrix} \qquad \text{and} \qquad \pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 3 & 4 & 6 & 7 & 5 \end{pmatrix}$$

Now all this reasoning is perfectly general; we can for any $\sigma \in \mathcal{S}_n$ and any $p$ with $1 \le p \le n$ a permutation $\pi \in \mathcal{S}_n$ and a $\tau \in \mathcal{S}_n$ and any $p$ with $1 \le p \le n$ a permutation $\pi \in \mathcal{S}_{n,p}$ so that

$$\tau = \sigma\pi$$

and $\pi$ has the property that

$$\pi = \pi_1\pi_2$$

where for $p+1 \le j \le n$ we have $\pi_1(j) = j$ and for $1 \le k \le p$ we have $\pi_2(k) = k$. We then have, since $\tau \in \mathcal{S}_{n,p}$,

$$*e_\tau = \text{sgn}(\tau)e^{\tilde{\tau}}$$

or, more explicitly,

$$*(e_{\tau(1)} \wedge \ldots \wedge e_{\tau(p)}) = \operatorname{sgn}(\tau) e^{\tau(p+1)} \wedge \ldots \wedge e^{\tau(n)}$$

We now replace $\tau$ by $\sigma\pi$ and using the elementary properties of permutations as they relate to products of vectors we have

$$\begin{aligned}
*(e_{\sigma\pi(1)} \wedge \ldots \wedge e_{\sigma\pi(p)}) &= \operatorname{sgn}(\sigma\pi) e^{\sigma\pi(p+1)} \wedge \ldots \wedge e^{\sigma\pi(n)} \\
\operatorname{sgn}(\pi_1) * (e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(p)}) &= \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\operatorname{sgn}(\pi_2) e^{\sigma(p+1)} \wedge \ldots \wedge e^{\sigma(n)} \\
(e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(p)}) &= \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\operatorname{sgn}(\pi_2)\operatorname{sgn}(\pi_1^{-1}) e^{\sigma(p+1)} \wedge \ldots \wedge e^{\sigma(n)} \\
&= \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)\operatorname{sgn}(\pi_2)\operatorname{sgn}(\pi_1) e^{\sigma(p+1)} \wedge \ldots \wedge e^{\sigma(n)} \\
&= \operatorname{sgn}(\sigma)[\operatorname{sgn}(\pi)]^2 e^{\sigma(p+1)} \wedge \ldots \wedge e^{\sigma(n)} \\
&= \operatorname{sgn}(\sigma) e^{\sigma(p+1)} \wedge \ldots \wedge e^{\sigma(n)}
\end{aligned}$$

as desired.

It is worth noting here that all we require of $e_1, \ldots, e_n$ is that it be a basis, that is that it be a linearly independent set. Thus the formula will work for any permutation and any set of linearly independent vectors. Of course, the catch is one must first find the dual set of elements in the dual space $V^*$.

## 5.7 The $\delta$ systems and $\epsilon$ systems

For computations of various quantities in coordinates it is useful to define certain quantities which act more like what we are accustomed to in linear algebra. The quantities also make it easier to interact the Grassmann product with other types of quantities. On the other hand, these quantities can be used to obscure the role of increasing permutations on Grassmann algebra, with the result that the theory looks much more complicated than it really is.

Let $V$ be an $n$-dimensional vector space, $e_1, \dots e_n$ a basis of $V$ and $e^1, \dots e^n$ the dual basis of $V^*$. To maintain symmetry we will express the value of a linear functional $f \in V^*$ on an element $v \in V$ by $\langle f, v\rangle$ rather than $f(v)$ and similarly for elements of $\Lambda^r(V^*)$ and $\Lambda^r(V)$. We now define

**Def** $$\delta^{i_1,\dots,i_r}_{j_1,\dots,j_r} = \langle e^{i_1} \wedge \dots \wedge e^{i_r}, e_{j_1} \wedge \dots \wedge e_{j_r}\rangle.$$

As a special case we have

$$\delta^{i_1}_{j_1} = \langle e^{i_1}, e_{j_1}\rangle$$

This is the ordinary Kronecker delta. The above represents a generalization. Next, by Grassmann's theorem we have

$$\begin{aligned}
\delta^{i_1,\dots,i_r}_{j_1,\dots,j_r} &= \det \begin{pmatrix} \langle e^1, e_1\rangle & \dots & \langle e^1, e_r\rangle \\ \dots & \dots & \dots \\ \langle e^r, e_1\rangle & \dots & \langle e^r, e_r\rangle \end{pmatrix} \\
&= \det \begin{pmatrix} \delta^1_1 & \dots & \delta^1_r \\ \dots & \dots & \dots \\ \delta^r_1 & \dots & \delta^r_r \end{pmatrix}
\end{aligned}$$

This last equation is used as a definition of $\delta^{i_1,\dots,i_r}_{j_1,\dots,j_r}$ when this quantity is required and no Grassmann algebra is available.

These equations allow us to discuss $\delta^{i_1,\dots,i_r}_{j_1,\dots,j_r}$ in more detail. First we notice that if $j_k$ is not among $i_1, \dots, i_r$ then $\delta^{i_l}_{j_k} = 0$ for all $l = 1, \dots, r$ so the determinant has the value 0. From this we see that

$$\delta^{i_1,\dots,i_r}_{j_1,\dots,j_r} = 0 \qquad if \quad \{i_1, \dots i_r\} \neq \{j_1, \dots j_r\} \quad \text{as sets.}$$

We also notice that $\delta^{i_1,\dots,i_r}_{j_1,\dots,j_r} = 0$ if $i_1 \dots i_r$ are not all distinct (since the determinant would then have a repeated column) and similarly for the indices $j_1 \dots j_r$.

Next we have for $\pi \in \mathcal{S}_r$

$$\begin{aligned}
\delta^{i_{\pi(1)}\cdots i_{\pi(r)}}_{j_1 \quad \dots \quad j_r} &= \det \begin{pmatrix} \langle e^{i_{\pi(1)}}, e_{j_1}\rangle & \dots & \langle e^{i_{\pi(1)}}, e_{j_r}\rangle \\ \dots & \dots & \dots \\ \langle e^{i_{\pi(r)}}, e_{j_1}\rangle & \dots & \langle e^{i_{\pi(r)}}, e_{j_r}\rangle \end{pmatrix} \\
&= \operatorname{sgn}(\pi) \det \begin{pmatrix} \langle e^{i_1}, e_{j_1}\rangle & \dots & \langle e^{i_1}, e_{j_r}\rangle \\ \dots & \dots & \dots \\ \langle e^{i_r}, e_{j_1}\rangle & \dots & \langle e^{i_r}, e_{j_r}\rangle \end{pmatrix} \\
&= \operatorname{sgn}(\pi) \delta^{i_1 \dots i_r}_{j_1 \dots j_r}
\end{aligned}$$

Setting $i_k = j_k = k$ for $k = 1, \ldots, r$ gives

$$\delta^{i_{\pi(1)} \ldots i_{\pi(r)}}_{j_1 \;\ldots\; j_r} = \mathrm{sgn}(\pi)\delta^{1\ldots r}_{1\ldots r} = \mathrm{sgn}(\pi).$$

Here are some examples:

**Examples:** $$\begin{array}{rcccccc} 1 & = & \delta^{137}_{137} & = & -\delta^{173}_{137} & = & \delta^{713}_{137} \\ 0 & = & \delta^{136}_{137} \end{array}$$

We now derive some computational formulas useful from time to time. The reader may wish to skim this material until she has reason to want to understand the proofs, which are straightforward but not very interesting.

We wish to contract $\delta^{i_1,\ldots,i_r}_{j_1,\ldots,j_r}$ on the last two indices, which means that we make the last upper and lower indices equal and then sum from 1 to $n$. To do this, we must first expand by the last column the determinant

$$\delta^{i_1\ldots i_{r-1} i_r}_{j_1\cdots j_{r-1} j_r} = \begin{vmatrix} \delta^{i_1}_{j_1} & \ldots & \delta^{i_1}_{j_{r-1}} & \delta^{i_1}_{j_r} \\ \delta^{i_2}_{j_1} & \ldots & \delta^{i_2}_{j_{r-1}} & \delta^{i_2}_{j_r} \\ \ldots & \ldots & \ldots & \ldots \\ \delta^{i_r}_{j_1} & \ldots & \delta^{i_r}_{j_{r-1}} & \delta^{i_r}_{j_r} \end{vmatrix}$$

which gives

$$(-1)^r \delta^{i_1}_{j_r}\delta^{i_2\ldots i_r}_{j_1\cdots j_{r-1}} + (-1)^{r+1}\delta^{i_2}_{j_r}\delta^{i_1 i_3\ldots i_r}_{j_1\cdots j_{r-1}} + (-1)^{r+2}\delta^{i_3}_{j_r}\delta^{i_1 i_2 i_4\ldots i_r}_{j_1\;\cdots\; j_{r-1}} + \ldots + + (-1)^{r+r}\delta^{i_r}_{j_r}\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}}$$

We now identify $i_r$ and $j_r$ and sum over the repeated index

$$\begin{aligned}
\delta^{i_1\ldots i_{r-1} i_r}_{j_1\cdots j_{r-1} i_r} &= (-1)^r \delta^{i_1}_{i_r}\delta^{i_2\ldots i_r}_{j_1\cdots j_{r-1}} + (-1)^{r+1}\delta^{i_2}_{i_r}\delta^{i_1 i_3\ldots i_r}_{j_1\cdots j_{r-1}} + (-1)^{r+2}\delta^{i_3}_{i_r}\delta^{i_1 i_2 i_4\ldots i_r}_{j_1\;\cdots\; j_{r-1}} + \\
&\quad + \ldots + (-1)^{r+r}\delta^{i_r}_{i_r}\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}} \\
&= (-1)^r \delta^{i_2\ldots i_{r-1} i_1}_{j_1\cdots j_{r-1}} + (-1)^{r+1}\delta^{i_1 i_3\ldots i_{r-1} i_2}_{j_1 \;\;\ldots\;\; j_{r-1}} + (-1)^{r+2}\delta^{i_1 i_2 i_4\ldots i_{r-1} i_3}_{j_1 \;\;\ldots\;\; j_{r-1}} + \\
&\quad + \ldots + n\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}} \\
&= (-1)^r(-1)^{r-1}\delta^{i_1 i_2\ldots i_{r-1}}_{j_1\;\ldots\; j_{r-1}} + (-1)^{r+1}(-1)^{r-2}\delta^{i_1 i_2 i_3\ldots i_{r-1}}_{j_1\;\ldots\; j_{r-1}} + \\
&\quad + (-1)^{r+2}(-1)^{r-3}\delta^{i_1 i_2 i_3 i_4\ldots i_{r-1}}_{j_1 \;\;\ldots\;\; j_{r-1}} + \ldots + n\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}} \\
&= (r-1)(-1)\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}} + n\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}} \\
&= (n-r+1)\delta^{i_1\ldots i_{r-1}}_{j_1\cdots j_{r-1}}\,.
\end{aligned}$$

We can now repeat this process;

$$\begin{aligned}
\delta^{i_1\ldots i_s i_{s+1}\ldots i_r}_{j_1\ldots j_s i_{s+1}\ldots i_r} &= (n-r+1)\delta^{i_1\ldots i_s i_{s+1}\ldots i_{r-1}}_{j_1\ldots j_s i_{s+1}\ldots i_{r-1}} \\
&= (n-r+1)(n-r+2)\delta^{i_1\ldots i_s i_{s+1}\ldots i_{r-2}}_{j_1\ldots j_s i_{s+1}\ldots i_{r-2}} \\
&= \ldots \\
&= (n-r+1)(n-r+2)\ldots(n-r+r-s)\delta^{i_1\ldots i_s}_{j_1\ldots j_s} \\
&= \tfrac{(n-s)!}{(n-r)!}\delta^{i_1\ldots i_s}_{j_1\ldots j_s}
\end{aligned}$$

If we set $r = s + t$, this can be rewritten as

$$\delta^{i_1\dots i_s i_{s+1}\dots i_{s+t}}_{j_1\dots j_s i_{s+1}\dots i_{s+t}} = \tfrac{(n-s)!}{(n-s-t)!}\delta^{i_1\dots i_s}_{j_1\dots j_s}\,.$$

Setting $s = 1$ gives

$$\delta^{i_1 i_2\dots i_t i_{t+1}}_{j_1 i_2\dots i_t i_{t+1}} = \tfrac{(n-1)!}{(n-1-t)!}\delta^{i_1}_{j_1}$$

and contracting $i_1$ and $j_1$ gives

$$\delta^{i_1 i_2\dots i_t i_{t+1}}_{i_1 i_2\dots i_t i_{t+1}} = \tfrac{(n-1)!}{(n-1-t)!}\delta^{i_1}_{i_1} = \tfrac{(n-1)!}{(n-1-t)!}n = \tfrac{n!}{(n-1-t)!}$$

so that, more simply,

$$\delta^{i_1\dots i_r}_{i_1\dots i_r} = \tfrac{n!}{(n-r)!}$$

and

$$\delta^{i_1\dots i_n}_{i_1\dots i_n} = \tfrac{n!}{(n-n)!} = n!\,.$$

Now let $A_{i_1\dots i_r}$ be any system of scalars indexed by $r$ indices. Then

$$\delta^{i_1\dots i_r}_{j_1\dots j_r}A_{i_1\dots i_r} = \sum_{\pi\in\mathcal{S}_r}\operatorname{sgn}(\pi)A_{j_{\pi(1)}\dots j_{\pi(r)}}$$

since the $\delta$ term is 0 unless $i_1,\dots,i_r$ is a permutation $\pi$ of $j_1,\dots,j_r$, and if this is the case the value is then $\operatorname{sgn}(\pi)$. Applying this formula to

$$A_{i_1\dots i_r} = \delta^{k_1\dots k_r}_{i_1\dots i_r}$$

gives

$$\begin{aligned}
\delta^{i_1\dots i_r}_{j_1\dots j_r}\delta^{k_1\dots k_r}_{i_1\dots i_r} &= \sum_{\pi\in\mathcal{S}_r}\operatorname{sgn}(\pi)\delta^{k_1\ \dots\ k_r}_{j_{\pi(1)}\dots j_{\pi(r)}}\\
&= \sum_{\pi\in\mathcal{S}_r}[\operatorname{sgn}(\pi)]^2\delta^{k_1\dots k_r}_{j_1\dots j_r}\\
&= r!\delta^{k_1\dots k_r}_{j_1\dots j_r}\,.
\end{aligned}$$

It will occasionally prove useful to have a variant of the generalized Kronecker delta for use with increasing permutations. Let $\sigma,\tau\in\mathcal{S}_n$. Then we define

**Def** $$\delta^{\sigma(i_1,\dots,i_r)}_{\sigma(j_1,\dots,j_r)} = \delta^{\sigma(i_1)\sigma(i_2)\dots\sigma(i_r)}_{\tau(j_1)\tau(j_2)\dots\tau(j_r)}\,.$$

If, in addition, $\sigma,\tau\in\mathcal{S}_{n,r}$, we will define

**Def** $$\delta^{\sigma}_{\tau} = \delta^{\sigma(i_1,\dots,i_r)}_{\sigma(j_1,\dots,j_r)} = \delta^{\sigma(i_1)\sigma(i_2)\dots\sigma(i_r)}_{\tau(j_1)\tau(j_2)\dots\tau(j_r)} \qquad \text{for} \quad \sigma,\tau\in\mathcal{S}_{n,r}\,.$$

We then notice the interesting circumstance that

$$\delta^{\sigma}_{\tau} = \begin{cases} 0 & \text{if } \sigma\neq\tau\\ 1 & \text{if } \sigma=\tau\end{cases} \qquad \text{for} \quad \sigma,\tau\in\mathcal{S}_{n,r}$$

because if $\sigma \neq \tau$ then $\{\sigma(1), \dots, \sigma(r)\}$ and $\{\tau(1), \dots, \tau(r)\}$ are distinct as sets and the Kronecker delta must be 0.

Clearly, if $A_\sigma$ is a quantity indexed by $\sigma \in \mathcal{S}_{n,r}$ then

$$\delta^\sigma_\tau A_\sigma = A_\tau \,.$$

We will find the following formula occasionally useful:

$$\sum_{\pi \in \mathcal{S}_{n,r}} \delta^{\pi(1)\dots\pi(r)}_{i_1 \,\dots\, i_r} \delta^{j_1 \,\dots\, j_r}_{\pi(1)\dots\pi(r)} = \delta^{i_1\dots i_r}_{j_1\dots j_r} \,.$$

This formula is almost obvious; we go into detail only to illustrate some technique. Fir, we note that unless $\{i_1, \dots, i_n\}$ and $\{j_1, \dots, j_n\}$ coincide as sets, both sides are 0. Supposing now that $\{i_1, \dots, i_n\}$ and $\{j_1, \dots, j_n\}$ coincide as sets there will be exactly one $\pi_0 \in \mathcal{S}_{n,r}$ having these sets as value, so that for $k = 1, \dots, r$

$$\begin{array}{rcl} i_{\sigma(k)} = \pi_0(k) & \qquad i_k = \pi_0(\sigma^{-1}(k)) & \\ j_{\rho(l)} = \pi_0(l) & \qquad j_l = \pi_0(\rho^{-1}(l)) & \end{array} \qquad \text{for some} \quad \sigma, \rho \in \mathcal{S}_r \,.$$

Then

$$\begin{aligned} \sum_{\pi \in \mathcal{S}_{n,r}} \delta^{\pi(1)\dots\pi(r)}_{i_1 \,\dots\, i_r} \delta^{j_1 \,\dots\, j_r}_{\pi(1)\dots\pi(r)} &= \delta^{\pi_0(1)\dots\pi_0(r)}_{i_1 \,\dots\, i_r} \delta^{j_1 \,\dots\, j_r}_{\pi_0(1)\dots\pi_0(r)} \\ &= \delta^{i_{\sigma(1)}\dots i_{\sigma(r)}}_{i_1 \,\dots\, i_r} \delta^{j_1 \,\dots\, j_r}_{j_{\rho(1)}\dots j_{\rho(r)}} \\ &= \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) \\ &= \operatorname{sgn}(\sigma^{-1})\operatorname{sgn}(\rho^{-1})\delta^{\pi(1)\dots\pi(r)}_{\pi(1)\dots\pi(r)} \\ &= \delta^{\pi_0(\rho^{-1}(1))\dots\pi_0(\rho^{-1}(r))}_{\pi_0(\sigma^{-1}(1))\dots\pi_0(\sigma^{-1}(r))} \\ &= \delta^{i_1\dots i_r}_{j_1\dots j_r} \end{aligned}$$

as desired. We have gone into the matter in such detail to illustrate that bridging the gap between a set of indices $\{i_1, \dots, i_r\}$ and a $\pi \in \mathcal{S}_{n,r}$ may be rigorously accomplished through the action of a $\sigma \in \mathcal{S}_r$. This is seldom necessary but it is comforting to know the technique exists.

Closely related to the generalized Kronecker deltas are the computationally useful $\varepsilon$-systems.

**Def**

$$\varepsilon_{i_1\dots i_n} = \delta^{1\,\dots\,n}_{i_1\dots i_n}$$

$$\varepsilon^{i_1\dots i_n} = \delta^{i_1\dots i_n}_{1\,\dots\,n}$$

Notice that an $\varepsilon$-system has $n = \dim(V)$ indices whereas the generalized Kronecker deltas may have any number of indices. Second, notice that $i_1, \dots i_n$

must be a permutation of $1, \dots, n$ or the $\varepsilon$-symbol has the value 0. Finally, from the properties of the $\delta$-symbol we see

$$\begin{aligned}
\varepsilon_{1\dots n} &= 1 \\
\varepsilon_{i_{\pi(1)}\cdots i_{\pi(n)}} &= \delta^{1\ \ \cdots\ \ n}_{i_{\pi(1)}\cdots i_{\pi(n)}} \\
&= \operatorname{sgn}(\pi)\delta^{1\ \cdots\ n}_{i_1\dots i_n} \\
&= \operatorname{sgn}(\pi)\varepsilon_{i_1\dots i_n}\,.
\end{aligned}$$

Setting $i_j = j$ we have

$$\begin{aligned}
\varepsilon_{\pi(1)\dots\pi(n)} &= \operatorname{sgn}(\pi)\delta^{1\dots n}_{1\dots n} \\
&= \operatorname{sgn}(\pi)\,.
\end{aligned}$$

The calculations are similar for $\varepsilon^{i_1\cdots i_n}$. Thus we have

$$\varepsilon_{i_1\dots i_n} = \varepsilon^{i_1\dots i_n} = \begin{cases} 0 & \text{if } i_1, \dots, i_n \text{ is not a permutation of } 1, \dots, n \\ \operatorname{sgn}(\pi) & \text{if } i_j = \pi(j) \text{ for some } \pi \in \mathcal{S}_n\,. \end{cases}$$

We now establish that

$$\varepsilon^{i_1\dots i_n}\varepsilon_{j_1\dots j_n} = \delta^{i_1\dots i_n}_{j_1\dots j_n}\,.$$

Both sides are 0 if $i_1, \dots, i_n$ are not all distinct. Similarly for $j_1, \dots, j_n$. Thus to have a non-zero result we must have $j_1, \dots, j_n$ a permutation of $1, \dots, n$ and similarly with $i_1, \dots, i_n$. Hence there are permutations $\pi, \sigma \in \mathcal{S}_n$ for which

$$\begin{aligned}
i_k &= \pi(k) \\
j_l &= \sigma(l)
\end{aligned}$$

Then we will have

$$\begin{aligned}
\delta^{i_1\dots i_n}_{j_1\dots j_n} &= \delta^{\pi(1)\dots\pi(n)}_{j_1\ \dots\ j_n} = \operatorname{sgn}(\pi)\delta^{1\ \cdots\ n}_{j_1\dots j_n} \\
&= \operatorname{sgn}(\pi)\delta^{1\ \ \cdots\ \ n}_{\sigma(1)\dots\sigma(n)} = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)\delta^{1\dots n}_{1\dots n} \\
&= \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma) \\
&= \varepsilon^{\pi(1)\dots\pi(n)}\varepsilon_{\sigma(1)\dots\sigma(n)} \\
&= \varepsilon^{i_1\dots i_n}\varepsilon_{j_1\dots j_n}
\end{aligned}$$

We note that

$$\varepsilon^{i_1\dots i_n}_{i_1\dots i_n} = \delta^{i_1\dots i_n}_{i_1\dots i_n} = n!$$

The $\varepsilon$-systems express the same idea as the sign of a permutation with the added advantage that if the indices are repeated then the $\varepsilon$-symbol gives the value 0. The disadvantage is that it is very hard to keep control of any ideas when one is manipulating $\varepsilon$-systems.

An important use of the $\varepsilon$-systems is in the theory of determinants. We recall that

$$\det\begin{pmatrix} \alpha^1_1 & \dots & \alpha^1_n \\ \dots & \dots & \dots \\ \alpha^n_1 & \dots & \alpha^n_n \end{pmatrix} = \sum_{\pi\in\mathcal{S}_n} \operatorname{sgn}(\pi)\alpha^1_{\pi(1)}\cdots\alpha^n_{\pi(n)}\,.$$

We can replace $\text{sgn}(\pi)$ by an $\varepsilon$-symbol, and introduce 0 terms for repeated indices, so that

$$\det(\alpha^i_j) = \varepsilon^{i_1\dots i_n}\alpha^1_{i_1}\alpha^2_{i_2}\cdots\alpha^n_{i_n}\,.$$

We can now rearrange the factors with a permutation $\pi \in \mathcal{S}_n$ and rewrite

$$\begin{aligned}\det(\alpha^i_j) &= \varepsilon^{i_1 i_2\dots i_n}\alpha^{\pi(1)}_{i_{\pi(1)}}\alpha^{\pi(2)}_{i_{\pi(2)}}\cdots\alpha^{\pi(n)}_{i_{\pi(n)}}\\ &= \text{sgn}(\pi)\varepsilon^{i_{\pi(1)} i_{\pi(2)}\dots i_{\pi(n)}}\alpha^{\pi(1)}_{i_{\pi(1)}}\alpha^{\pi(2)}_{i_{\pi(2)}}\cdots\alpha^{\pi(n)}_{i_{\pi(n)}}\,.\end{aligned}$$

We can now relable the $i_{\pi(k)} = j_k$ and get

$$\det(\alpha^i_j) = \text{sgn}(\pi)\varepsilon^{j_1 j_2\dots j_n}\alpha^{\pi(1)}_{j_1}\alpha^{\pi(2)}_{j_2}\cdots\alpha^{\pi(n)}_{j_n}\,.$$

Summing this last equation over $\pi \in \mathcal{S}_n$ we have

$$n!\,\det(\alpha^i_j) = \sum_{\pi\in\mathcal{S}_n}\text{sgn}(\pi)\varepsilon^{j_1 j_2\dots j_n}\alpha^{\pi(1)}_{j_1}\alpha^{\pi(2)}_{j_2}\cdots\alpha^{\pi(n)}_{j_n}$$

and introducing another $\varepsilon$-symbol for $\text{sgn}(\pi)$ we have

$$\begin{aligned}\det(\alpha^i_j) &= \tfrac{1}{n!}\varepsilon_{i_1\dots i_n}\varepsilon^{j_1 j_2\dots j_n}\alpha^{i_1}_{j_1}\cdots\alpha^{i_n}_{j_n}\\ &= \tfrac{1}{n!}\delta^{j_1 j_2\dots j_n}_{i_1\dots i_n}\alpha^{i_1}_{j_1}\cdots\alpha^{i_n}_{j_n}\,.\end{aligned}$$

We want now to use the generalized Kronecker deltas to describe Grassmann algebra coefficients. We can derive all this from previous material, but will first introduce the basic idea from first principles.

Let $w_1 = \alpha^1_1 v_1 + \alpha^2_1 v_2 + \alpha^3_1 v_3$ and $w_2 = \alpha^1_2 v_1 + \alpha^2_2 v_2 + \alpha^3_2 v_3$. Then

$$w_1 \wedge w_2 = \alpha^2_1\alpha^3_2 - \alpha^3_1\alpha^2_2 v_1 \wedge v_2 + \text{two other terms.}$$

The coefficient of $v_1 \wedge v_2$ is

$$\delta^{23}_{ij}\alpha^i_1\alpha^j_2$$

In a similar way, if $w_i = \alpha^j_i v_j$ then the coefficient of $v_{\pi(1)} \wedge \dots \wedge v_{\pi(r)},\ \pi \in \mathcal{S}_{n,r}$ in $w_{\sigma(1)} \wedge \dots \wedge w_{\sigma(r)},\ \sigma \in \mathcal{S}_{n,r}$ is

$$\delta^{\pi(1)\dots\pi(r)}_{i_1\ \dots\ i_r}\alpha^{i_1}_{\sigma(1)}\cdots\alpha^{i_r}_{\sigma(r)}\,.$$

We have treated this problem before, in section 3.3, where we found that the coefficient was

$$\alpha^\pi_\sigma = \det\begin{pmatrix}\alpha^{\pi(1)}_{\sigma(1)} & \dots & \alpha^{\pi(1)}_{\sigma(r)}\\ \dots & \dots & \dots\\ \alpha^{\pi(r)}_{\sigma(1)} & \dots & \alpha^{\pi(r)}_{\sigma(r)}\end{pmatrix}$$

so that the above expression coincides with this determinant:

$$\alpha^\pi_\sigma = \delta^{\pi(1)\dots\pi(r)}_{i_1\ \dots\ i_r}\alpha^{i_1}_{\sigma(1)}\cdots\alpha^{i_r}_{\sigma(r)}\,.$$

We now wish to rewrite the product itself with generalized Kronecker deltas. First, notice that the last equation will remain valid if $\pi(i)$ and $\sigma(j)$ are replaced by sets of distinct (not necessarily increasing) indices:

$$\det\begin{pmatrix} \alpha^{i_1}_{k_1} & \dots & \alpha^{i_1}_{k_r} \\ \dots & \dots & \dots \\ \alpha^{i_r}_{k_1} & \dots & \alpha^{i_r}_{k_r} \end{pmatrix} = \delta^{j_1\dots j_r}_{i_1\dots i_r}\alpha^{i_1}_{k_1}\cdots\alpha^{i_r}_{k_r}$$

because both sides will undergo similar changes in sign if the $j_1\dots j_r$ and $k_1\dots k_r$ are permuted into increasing sequences of indices. Next notice that if $j_1\dots j_r$ are not all distinct then both sides are 0, and similarly if $k_1\dots k_r$ are not all distinct. Thus the relationship is true for any valued of $j_1\dots j_r$ and $k_1\dots k_r$. Now notice that

$$\begin{aligned} v_{\sigma(1)}\wedge\cdots\wedge v_{\sigma(r)} &= \operatorname{sgn}(\pi)\, v_{\sigma(\pi(1))}\wedge\cdots\wedge v_{\sigma(\pi(r))} \\ &= \tfrac{1}{r!}\sum_{\pi\in\mathcal{S}_n}\operatorname{sgn}(\pi) v_{\sigma(\pi(1))}\wedge\cdots\wedge v_{\sigma(\pi(r))} \\ &= \delta^{\,i_1\ \dots\ i_r}_{\sigma(1)\dots\sigma(1)} v_{i_1}\wedge\cdots\wedge v_{i_r} \end{aligned}$$

where all the summands in the sums are equal to $v_{\sigma(1)}\wedge\cdots\wedge v_{\sigma(r)}$, $\sigma\in\mathcal{S}_{n,r}$. We are now in a position to write the Grassmann product with the generalized Kronecker deltas. Indeed

$$\begin{aligned} w_{\sigma(1)}\wedge\cdots\wedge w_{\sigma(r)} &= \sum_{\pi\in\mathcal{S}_{n,r}}\delta^{\pi(1)\dots\pi(1)}_{\ i_1\ \dots\ i_r}\alpha^{i_1}_{\sigma(1)}\cdots\alpha^{i_r}_{\sigma(r)}\, v_{\pi(1)}\wedge\cdots\wedge v_{\pi(r)} \\ &= \tfrac{1}{r!}\sum_{\pi\in\mathcal{S}_{n,r}}\delta^{\pi(1)\dots\pi(1)}_{\ i_1\ \dots\ i_r}\alpha^{i_1}_{\sigma(1)}\cdots\alpha^{i_r}_{\sigma(r)}\,\delta^{\ j_1\ \dots\ j_r}_{\pi(1)\dots\pi(1)}\, v_{j_1}\wedge\cdots\wedge v_{j_r} \\ &= \tfrac{1}{r!}\,\delta^{j_1\dots j_r}_{i_1\dots i_r}\,\alpha^{i_1}_{\sigma(1)}\cdots\alpha^{i_r}_{\sigma(r)}\, v_{j_1}\wedge\cdots\wedge v_{j_r} \end{aligned}$$

where we have used the formula from earlier in the section to eliminate the sum.

# Chapter 6

# Inner Products on $V$ and $V^*$

## 6.1 Introduction

In this chapter we develop Inner Products on the Grassmann algebra. On any vector space, an inner product can be correlated with an isomorphism between the vector space $V$ and its dual space $V^*$. We can extend this isomorphic mapping $\Phi : V \to V^*$ to an isomorphic mapping $\Phi : \Lambda^p(V) \to \Lambda^p(V^*)$ in a natural way, and this can be reinterpreted to give an inner product. In classical Tensor Analysis, this is the content behind raising and lowering indices.

The basic tool to develop the formulas for the inner product will be Grassmann's theorem. It would be possible to simply define the inner product to be the final formula, but the definition then looks rather arbitrary.

We will also develop the $*$-operator in a metric setting. Clasically, $*$ is usually developed by the use of an orthonormal basis, which I feel is improper methodology. The reason for the orthonormal basis is that the metric $*$-operators are really combinations of the $*$-operator of Chapter 5 and the above isomorphisms generated by the inner product. When the basis is orthonormal, the isomorphisms become very well behaved and can be virtually ignored, but this methodology does not work well if the bases are not orthogonal. We show in this chapter that quite natural formulas for $*$ can be developed for any basis and that the difficulties can be completely overcome by separating the roles in the metric versions of $*$ of the above isomorphisms and the $*$ of chapter 5. This leads to the derivation of doublets of formulas for $*$ which then can be used to good advantage, since if one of the doublet is not acting productively the other often will.

We will also see that in the metric setting and *real* scalars the metric form of $*$ is almost uniquely defined; it can do no more than change sign when the basis is changed (and we know when the sign change will occur). In the case of *complex* scalars things are less satisfactory and it will turn out that the metric $*$ operator may accumulate a unimodular ($|\lambda| = 1$) complex factor when the basis is changed. There appears no obvious satisfactory way to insulate $*$ from basis changes in the case of complex scalars.

## 6.2 Exporting the Inner Product on $V$ to $V^*$, $\Lambda^r(V)$ and $\Lambda^r(V^*)$

We first remind the reader of certain of our conventions. If $V$ is a complex vector space with an inner product $(u, v)$, we specify that the inner product is linear in the *second* variable and anti-linear in the *first* variable:

$$(\lambda u, v) = \overline{\lambda}(u, v) \qquad (u, \lambda v) = \lambda(u, v)\,.$$

You may prefer it the opposite way, and it turns out that it *will* be the opposite way for the inner product we will create for $V^*$. I have tried various alternatives here and since there is no way to please everyone I have selected this alternative as being as good as any and better than some.

As with any inner product, we assume it is Hermitian ( $(v, u) = \overline{(u, v)}$ ) if the scalars are the complex numbers and symmetric ( $(v, u) = (u, v)$ ) if the scalars are the real numbers. We can handle both cases at once but considering the conjugate bar to have no effect in the second (real) case. We also assume the inner product is non-degenerate, which means

$$\begin{array}{llll} \text{if} & (u, v) = 0 & \text{for all } v \in V & \text{then} \quad u = 0\,. \\ \text{if} & (u, v) = 0 & \text{for all } u \in V & \text{then} \quad v = 0\,. \end{array}$$

(By the Hermitian or symmetric property, it suffices to assume just one of these.)

An inner product on $V$ creates an anti-isomorphism between $V$ and $V^*$. We recall that a function $\Phi$ is anti-linear if it satisfies

$$\Phi(\lambda u + \mu v) = \overline{\lambda}\Phi(u) + \overline{\mu}\Phi(v)\,.$$

An anti-isomorphism is an anti-linear map which is one-to-one and onto. The mapping $\Phi$ determined by the inner product is defined as follows:

**Def** $\qquad \langle \Phi(u), v\rangle = (u, v) \qquad \text{for all } u, v \in V\,.$

$\Phi(u)$ is clearly in $V^*$ and we also have

$$\begin{aligned} \langle \Phi(\lambda u + \mu v), w\rangle &= (\lambda u + \mu v, w) \\ &= \overline{\lambda}(u, w) + \overline{\mu}(v, w) \\ &= \overline{\lambda}\langle \Phi(u), w\rangle + \overline{\mu}\langle \Phi(v), w\rangle \\ &= \langle \overline{\lambda}\Phi(u) + \overline{\mu}\Phi(v), w\rangle\,. \end{aligned}$$

Since this is true for all $w \in V$, we have

$$\Phi(\lambda u + \mu v) = \overline{\lambda}\Phi(u) + \overline{\mu}\Phi(v) \qquad \text{for all } u, v \in V\,.$$

We now have the basic theorem:

**Theorem** $\Phi : V \to V^*$ defined by

$$\langle \Phi(u), v\rangle = (u, v) \qquad \text{for all } u, v \in V$$

is an anti-isomorphism. **Proof** We have already verified the anti-linearity. We next show that $\Phi$ is one-to-one. Suppose $\Phi(u) = 0$. Then for all $v \in V$

$$(u, v) = \langle \Phi(u), v\rangle = \langle 0, v\rangle = 0\,.$$

Since the inner product is non-degenerate, (and this is the place where we really need it,) we have $u = 0$. Thus $\Phi$ is one-to-one.

To show it is onto, we note the obvious fact that the image $\Phi[V]$ is a subspace of $V^*$. Because $\Phi$ is one-to-one, it is an $n$-dimensional subspace of $V^*$. But $\dim(V^*) = \dim(V) = n$, so that $V^*$ must be the image of $\Phi$ and thus $\Phi$ is onto.

**Exporting the Inner Product from $V$ to $V^*$**

We can use $\Phi^{-1}$ to export the inner product on $V$ to an inner product on $V^*$ in the obvious way; **Def** For $\ell, m \in V^*$ we set

$$(\ell, m) = (\Phi^{-1}(\ell), \Phi^{-1}(m))$$

where the inner product on the right side of the equation is taken in $V$. This inner product is *linear* in the *first* variable and *anti-linear* in the *second* variable; for $\ell, m, n \in V^*$

$$\begin{aligned}
(\lambda\ell + \mu m, n) &= (\Phi^{-1}(\lambda\ell + \mu m), \Phi^{-1}(n)) \\
&= (\overline{\lambda}\Phi^{-1}(\ell) + \overline{\mu}\Phi^{-1}(m), \Phi^{-1}(n)) \\
&= \lambda(\Phi^{-1}(\ell), \Phi^{-1}(n)) + \mu(\Phi^{-1}(m), \Phi^{-1}(n)) \\
&= \lambda(\ell, n) + \mu(m, n)
\end{aligned}$$

and

$$\begin{aligned}
(\ell, \lambda m + \mu n) &= (\Phi^{-1}(\ell), \Phi^{-1}(\lambda m + \mu n)) \\
&= (\Phi^{-1}(\ell), \overline{\lambda}\Phi^{-1}(m) + \overline{\mu}\Phi^{-1}(n)) \\
&= \overline{\lambda}(\Phi^{-1}(\ell), \Phi^{-1}(m)) + \overline{\mu}(\Phi^{-1}(\ell), \Phi^{-1}(n)) \\
&= \overline{\lambda}(\ell, m) + \overline{\mu}(\ell, n)
\end{aligned}$$

The Hermitian or symmetric property is obvious. The non-degeneracy is checked as follows: if $(\ell, m) = 0$ for all $m \in V^*$ then

$$(\Phi^{-1}(\ell), \Phi^{-1}(m)) = (\ell, m) = 0$$

and since $\Phi^{-1}$ is onto we then have

$$(\Phi^{-1}(\ell), v) = 0 \qquad \text{for all } v \in V$$

so $\Phi^{-1}(\ell) = 0$ by the non-degeneracy of the inner product in $V$ and since $\Phi^{-1}$ is one to one, $\ell = 0$. Thus $(\ell, m)$ is an inner product in $V^*$.

To maintain the symmetry between $V$ and $V^*$ we want to derive one other formula:

$$\begin{aligned}(\ell, m) &= (\Phi^{-1}(\ell), \Phi^{-1}(m)) \\ &= \langle \Phi(\Phi^{-1}(\ell)), \Phi^{-1}(m) \rangle \\ &= \langle \ell, \Phi^{-1}(m) \rangle\,.\end{aligned}$$

Thus we have the two symmetric equations

$$\begin{aligned}(u, v) &= \langle \Phi(u), v \rangle && \text{for } u, v \in V \\ (\ell, m) &= \langle \ell, \Phi^{-1}(m) \rangle && \text{for } \ell, m \in V^*\,.\end{aligned}$$

**Exporting the Inner Product from $V$ to $\Lambda^r(V)$** We recall Grassmann's Theorem: $\langle \ell^1 \wedge \ldots \wedge \ell^r, v_1 \wedge \ldots \wedge v_r \rangle = \det(\langle \ell^i, v_j \rangle)$. We also recall that a linear operator $\Phi : V \to V^*$ may be extended to an operator $\Phi : \Lambda^r(V) \to \Lambda^r(V^*)$ by means of the formula $\Phi(v_1 \wedge \ldots \wedge v_r) = \Phi(v_1) \wedge \ldots \wedge \Phi(v_r)$. We use these to export the inner product from $V$ to $\Lambda^r(V)$ by insisting that the formula

$$(u, v) = \langle \Phi(u), v \rangle \qquad u, v \in V$$

remain valid for all $\Lambda^p(V)$:

**Def** $\quad (u_1 \wedge \ldots \wedge u_r, v_1 \wedge \ldots \wedge v_r) = \langle \Phi(u_1 \wedge \ldots \wedge u_r), v_1 \wedge \ldots \wedge v_r \rangle$.

We then have immediately **Grassmann's Theorem; inner product form**

$$(u_1 \wedge \ldots \wedge u_r, v_1 \wedge \ldots \wedge v_r) = \det((u_i, v_j))\,.$$

**Proof**

$$\begin{aligned}(u_1 \wedge \ldots \wedge u_r, v_1 \wedge \ldots \wedge v_r) &= \langle \Phi(u_1 \wedge \ldots \wedge u_r), v_1 \wedge \ldots \wedge v_r \rangle \\ &= \langle \Phi(u_1) \wedge \ldots \wedge \Phi(u_r), v_1 \wedge \ldots \wedge v_r \rangle \\ &= \det(\langle \Phi(u_i), v_j \rangle) \\ &= \det((u_i, v_j))\,.\end{aligned}$$

The inner product is now extended by linearity from products of $r$ vectors to the whole of $\Lambda^r(V)$.

This covers all of the Grassmann algebra of V except $\Lambda^0(V)$ which is defined to be the set of scalars. In this case we define

**Def** $\qquad (\lambda, \mu) = \overline{\lambda}\mu \qquad \text{for } \lambda, \mu \in \Lambda^0(V)\,.$

It is sometimes convenient to extend the inner product from each $\Lambda^r(V)$ to the entire Grassmann Algebra

$$\Lambda(V) = \bigoplus_{r=0}^{n} \Lambda^r(V)$$

This is easily done by setting

**Def** $\quad (A, B) = 0 \quad$ if $A \in \Lambda^r(V)$ and $B \in \Lambda^s(V)$ and $r \neq s$.

It is also sometimes convenient to extend the definition of $\Phi$ to the entire Grassmann algebra. We have already defined $\Lambda^r(V)$ for all positive $r$. It only remains to define $\Phi : \Lambda^0(V) \to \Lambda^0(V^*)$. We first define, for $1 \in \Lambda^0(V)$

$$\Phi(1) = 1 \in \Lambda^0(V^*)$$

and then to preserve the usual antilinearity define

$$\Phi(\lambda) = \Phi(\lambda \cdot 1) = \overline{\lambda}\Phi(1) = \overline{\lambda} \cdot 1 = \overline{\lambda}$$

and similarly we have $\Phi^{-1} : \Lambda^0(V^*) \to \Lambda^0(V)$ defined by

$$\Phi^{-1}(\lambda) = \overline{\lambda}\,.$$

We now have defined the isomorphism $\Phi$ of Grassmann algebras completely:

$$\Phi : \Lambda(V) = \bigoplus_{r=0}^{n} \Lambda^r(V) \to \bigoplus_{r=0}^{n} \Lambda^r(V^*) = \Lambda(V^*)\,.$$

**Exporting the Inner Product from $V^*$ to $\Lambda^r(V^*)$**

There are several equivalent ways to extend the inner product to $\Lambda^r(V^*)$ all leading to the same result. We will do it in analogy to the method we used to go from $V$ to $\Lambda^r(V)$, but here we will use the formula

$$(\ell, m) = \langle \ell, \Phi^{-1}(m) \rangle.$$

We now insist this formula hold in $\Lambda^r(V^*)$, and for $\ell^1, \dots, \ell^r, m^1, \dots, m^r \in V^*$ we define

**Def** $\quad (\ell^1 \wedge \dots \wedge \ell^r, m^1 \wedge \dots \wedge m^r) = \langle \ell^1 \wedge \dots \wedge \ell^r, \Phi^{-1}(m^1 \wedge \dots \wedge m^r) \rangle\,.$

One then has the expected result

$$\begin{aligned}
(\ell^1 \wedge \dots \wedge \ell^r, m^1 \wedge \dots \wedge m^r) &= \langle \ell^1 \wedge \dots \wedge \ell^r, \Phi^{-1}(m^1 \wedge \dots \wedge m^r) \rangle \\
&= \det(\langle \ell^i, \Phi^1(m^j) \rangle \\
&= \det((\ell^i, m^j))\,.
\end{aligned}$$

We then extend by linearity to all of $\Lambda^r(V^*)$, and finally to all of $\Lambda(V^*) = \bigoplus_{r=0}^{n} \Lambda^r(V^*)$ by $(A, B)) = 0$ for $A \in \Lambda^r(V^*)$, $B \in \Lambda^s(V^*)$ and $r \neq s$. We now derive some variants of the basic formulas. By definition, for $u, v \in \Lambda^r(V)$

$$(u, v) = \langle \Phi(u), v \rangle \,.$$

But then

$$\begin{aligned} (u, v) &= \langle \Phi(u), v \rangle = (u, v) = \overline{(v, u)} \\ &= \overline{\langle \Phi(v), u \rangle} \,. \end{aligned}$$

Now, setting $\ell = \Phi(u),\ m = \Phi(v)$, we have $\ell, m \in \Lambda^r(V^*)$ and

$$\begin{aligned} \langle l, v \rangle &= (\Phi^{-1}(\ell), v) \\ &= \overline{\langle \Phi(v), \Phi^{-1}(\ell) \rangle} \end{aligned}$$

giving us the interesting formula

$$\langle \Phi(v), \Phi^{-1}(\ell) \rangle = \overline{\langle l, v \rangle} \,.$$

Before moving on, we wish to note that the the formulas used for defining the inner products on $\Lambda^r(V)$ and on $\Lambda^r(V^*)$ also extend by linearity to all elements of the Grassmann algebra, giving

$$\begin{aligned} (A, B) &= \langle \Phi(A), B \rangle && \text{for all } A, B \in \Lambda^r(V) \\ (A, B) &= \langle A, \Phi^{-1}(B) \rangle && \text{for all } A, B \in \Lambda^r(V^*) \,. \end{aligned}$$

We have defined $\Phi$ on all of the Grassmann algebras $\Lambda(V)$ and $\Lambda(V^*)$ except for the bottom levels $\Lambda^0(V)$ and $\Lambda^0(V^*)$. Recall that the bottom level $\Lambda^0(V)$ is just the scalars, and similarly for $\Lambda^0(V^*)$. A basis in either case is the number 1. The reasonable definition in the circumstances is for $1 \in \Lambda^0(V)$ we define

**Def** $$\Phi(1) = 1 \in \Lambda^0(V^*)$$

Recalling that $\Phi$ has always been an anti-isomorphism, it is reasonable to extend this by

$$\Phi(\lambda) = \Phi(\lambda \cdot 1) = \overline{\lambda}\Phi(1) = \overline{\lambda} \,.$$

We then naturally also have

$$\Phi^{-1}(\lambda) = \overline{\lambda} \,.$$

**Formulas for $\Phi$ and the Inner Products in Coordinates**

Now that we have all the necessary formulas in hand, we wish to find the coordinate forms of the formulas. There are several things to find. We need matrices which express $\Phi$ and $\Phi^{-1}$ and we want matrices for the various inner products.

First we find the matrix of the inner product for $V$. We have $e_1, \ldots, e_n$ a basis for $V$ and $e^1, \ldots, e^n$ the dual basis of $V^*$. (Recall this means that $e^i(e_j) = \langle e^i, e_j \rangle = \delta^i_j$.) The matrix $g_{ij}$ is made from

**Def** $$g_{ij} = (e_i, e_j)$$

so that $g_{ji} = \overline{g_{ij}}$, and $(g_{ij})$ is a Hermitian (or symmetric if $V$ is a real vector space) matrix. Since the inner product is non-degenerate, $\det(g_{ij}) \neq 0$ and we can form $g^{k\ell} = g_{ij}^{-1}$. Now if $u = \rho^i e_i$ and $u = \sigma^j e_j$ we have

$$\begin{aligned}(u, v) &= (\rho^i e_i, \sigma^j e_j) = \overline{\rho^i}\sigma^j (e_i, e_j) \\ &= g_{ij}\overline{\rho^i}\sigma^j\end{aligned}$$

Our next job is to find the formula for $\Phi$ in coordinates, which is easy. The "matrix" of $\Phi$ can now be found, using the bases $e_1, \ldots, e_n$ for $V$ and $e^1, \ldots, e^n$ for $V^*$ in the following way:

$$\begin{aligned}\Phi(e_i) &= \alpha_{ij} e^j && \text{for some } \alpha_{ij} \\ \langle \Phi(e_i), e_k \rangle &= (e_i, e_k) && \text{def of } \Phi \\ \langle \alpha_{ij} e^j, e_k \rangle &= g_{ik} \\ \alpha_{ij} \langle e^j, e_k \rangle &= g_{ik} \\ \alpha_{ij} \delta^j_k &= g_{ik} \\ \alpha_{ik} &= g_{ik}\end{aligned}$$

so that

$$\Phi(e_i) = g_{ik} e^k \,.$$

If now we set $u = \rho^i e_i \in V$ then

$$\Phi(u) = \Phi(\rho^i e_i) = \overline{\rho^i}\Phi(e_i) = g_{ik}\overline{\rho^i} e^k \,.$$

Since $\Phi$ is an anti-isomorphism we would also like to have the formula in coordinates for $\Phi^{-1}$. We have

$$\Phi(e_i) = g_{ik} e^k$$

so that

$$\begin{aligned}e_i &= \Phi^{-1}(g_{ik} e^k) \\ &= \overline{g_{ik}}\Phi^{-1}(e^k) \\ &= g_{ki}\Phi^{-1}(e^k) \\ g^{i\ell} e_i &= g_{ki} g^{i\ell}\Phi^{-1}(e^k) \\ &= \delta^\ell_k \Phi^{-1}(e^k) \\ &= \Phi^{-1}(e^\ell)\end{aligned}$$

and thus for $\ell = \lambda_i e^i \in V^*$ we have

$$\begin{aligned}\Phi^{-1}(\ell) &= \Phi^{-1}(\lambda_i e^i) \\ &= \overline{\lambda_i}\Phi^{-1}(e^i) \\ &= g^{ki}\overline{\lambda_i} e_k \,.\end{aligned}$$

If one wishes to work entirely in coordinates we can set $\ell = \lambda_i e^i$ and $u = \rho^j e_j$ and the formulas that reflect $\ell = \Phi(u)$ and $u = \Phi^{-1}(\ell)$ are

$$\begin{aligned}\lambda_i &= g_{ji}\overline{\rho^j} && \underline{L}\text{ower with } \underline{L}\text{eft index} \\ \rho^j &= g^{ji}\overline{\lambda_i} && \underline{R}\text{aise with } \underline{R}\text{ight index}\end{aligned}$$

These are the formulas for "raising and lowering indices" so familiar from classical tensor analysis. The meaning behind the activity is representing $u \in V$ by the element $\Phi(u) \in V^*$ or vice-versa. The formulas can also be presented in the variant forms

$$\lambda_i = \overline{g_{ij}\rho^j} \qquad \rho^j = \overline{g^{ij}\lambda_i}$$

Since $\Lambda^r(V)$ is an inner product space, it will have metric coefficients corresponding to any basis, and we now wish to determine their form. To this end we recall that

$$e_\pi = e_{\pi(1)} \wedge \ldots \wedge e_{\pi(r)} \quad \text{form a basis of } \Lambda^r(V) \text{ where } \pi \in \mathcal{S}_{n,r}.$$

The metric coefficients are then given by

**Def**
$$\begin{aligned} g_{\pi\sigma} &= (e_\pi, e_\sigma) = (e_{\pi(1)} \wedge \ldots \wedge e_{\pi(r)}, e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(r)}) \\ &= \det((e_{\pi(i)}, e_{\sigma(j)})) = \det(g_{\pi(i)\sigma(j)}), \quad 1 \le i,j \le r\,. \end{aligned}$$

Thus, the metric coefficients of $\Lambda^r(V)$ are the size $r$ subdeterminants of $(g_{ij})$. In terms of basis elements, if $u = \rho^\alpha e_\alpha$ and $v = \sigma^\beta e_\beta$ where $\alpha, \beta \in \mathcal{S}_{n,r}$, then

$$\begin{aligned} (u,v) &= (\rho^\alpha e_\alpha, \sigma^\beta e_\beta) \\ &= \overline{\rho^\alpha}\sigma^\beta (e_\alpha, e_\beta) \\ &= g_{\alpha\beta}\overline{\rho^\alpha}\sigma^\beta \qquad \alpha, \beta \in \mathcal{S}_{n,r} \end{aligned}$$

Now we want the matrix for the inner product in $V^*$. We compute

$$\begin{aligned} (e^i, e^j) &= (\Phi^{-1}(e^i), \Phi^{-1}(e^j)) \\ &= (g^{ki}e_k, g^{\ell j}e_\ell) \\ &= \overline{g^{ki}}g^{\ell j}(e_k, e_\ell) \\ &= \overline{g^{ki}}g^{\ell j}g_{k\ell} \\ &= \overline{g^{ki}}\delta^j_k = \overline{g^{ji}} = g^{ij} \end{aligned}$$

Thus we have the metric coefficients for the inner product on $V^*$ and they turn out to be the inverse of those of the inner product of $V$.

**Remark** Note that we have the following three highly desirable equations:

$$(e_i, e_j) = g_{ij} \qquad (e^i, e^j) = g^{ij} \qquad (g_{ij})(g^{jk}) = (\delta^k_i) = I$$

We are able to get all of these in the case of complex scalars because we have set up the inner products on $V$ and $V^*$ to be antilinear in the opposite slots. No matter how things are set up there will be inconvenience somewhere, and this seems to me to be fairly optimal for computational purposes.

We can now find formulas in coordinates for $(\ell, m)$. Let $\ell = \lambda_i e^i$ and $m = \mu_j e^j$ and we have

$$\begin{aligned} (\ell, m) &= (\lambda_i e^i, \mu_j e^j) \\ &= \lambda_i \overline{\mu_j}(e^i, e^j) \\ &= g^{ij}\lambda_i\overline{\mu_j}\,. \end{aligned}$$

Just as in the case of $\Lambda^r(V)$ we can now derive the coefficients for the inner product in $\Lambda^r(V^*)$. We have

$$e^\pi = e^{\pi(1)} \wedge \ldots \wedge e^{\pi(r)} \quad \text{form a basis of } \Lambda^r(V^*) \text{ where } \pi \in \mathcal{S}_{n,r}.$$

The metric coefficients are then given by

$$\begin{aligned} g^{\pi\sigma} &= (e^\pi, e^\sigma) = (e^{\pi(1)} \wedge \ldots \wedge e^{\pi(r)}, e^{\sigma(1)} \wedge \ldots \wedge e^{\sigma(r)}) \\ &= \det((e^{\pi(i)}, e^{\sigma(j)})) = \det(g^{\pi(i)\sigma(j)}), \quad 1 \le i, j \le r\,. \end{aligned}$$

Thus, the metric coefficients of $\Lambda^r(V)$ are the size $r$ subdeterminants of $(g^{ij})$. In terms of basis elements, if $\ell = \lambda_\alpha e^\alpha$ and $m = \mu_\beta e^\beta$ where $\alpha, \beta \in \mathcal{S}_{n,r}$, then

$$\begin{aligned} (\ell, m) &= (\lambda_\alpha e^\alpha, \mu_\beta e^\beta) \\ &= \lambda_\alpha \overline{\mu_\beta} (e^\alpha, e^\beta) \\ &= g^{\alpha\beta} \lambda_\alpha \overline{\mu_\beta} \qquad \alpha, \beta \in \mathcal{S}_{n,r} \end{aligned}$$

We now wish explicit formulas describing $\Phi : \Lambda^r(V) \to \Lambda^r(V^*)$. We define

**Def**
$$\begin{aligned} E_i &= \Phi(e_i) &&= g_{ik} e^k \in V^* \\ E^i &= \Phi^{-1}(e^i) &&= g^{ji} e_j \in V \end{aligned}$$

and recall that, since $\Phi$ is an anti-isomorphism,

$$\begin{aligned} \Phi(\rho^i e_i) &= \overline{\rho^i} E_i \\ \Phi^{-1}(\lambda_j e^j) &= \overline{\lambda_j} E^j\,. \end{aligned}$$

$\Phi$ extends naturally to $\Lambda^r(V)$ and $\Phi^{-1}$ to $\Lambda^r(V^*)$ so that

$$\begin{aligned} \Phi(\rho^\alpha e_\alpha) &= \overline{\rho^\alpha} E_\alpha && \alpha \in \mathcal{S}_{n,r} \\ \Phi^{-1}(\lambda_\beta e^\beta) &= \overline{\lambda_\beta} E^\beta && \beta \in \mathcal{S}_{n,r}\,. \end{aligned}$$

These $E_\alpha$ and $E^\alpha$ will do no good unless we can decode them, which we do now. Although we have general formulas for this, we will run through it again quickly for those who don't want to read the background material.

$$\begin{aligned} E_\alpha &= E_{\alpha(1)} \wedge \ldots \wedge E_{\alpha(r)} \\ &= g_{\alpha(1)i_1} e^{i_1} \wedge \ldots \wedge g_{\alpha(r)i_r} e^{i_r} \\ &= g_{\alpha(1)i_1} g_{\alpha(2)i_2} \cdots g_{\alpha(r)i_r} e^{i_1} \wedge e^{i_2} \wedge \ldots \wedge e^{i_r} \end{aligned}$$

We now use the method of resolving by permutations, where we group together all terms with the same indices on the $e^i$ which we have arranged so the indices increase. We then have

$$\begin{aligned} E_\alpha &= \sum_{\pi \in \mathcal{S}_{n,r}} \Big( \sum_{\rho \in \mathcal{S}_r} \operatorname{sgn}(\rho) g_{\alpha(1)\pi(\rho(1))} g_{\alpha(2)\pi(\rho(2))} \cdots g_{\alpha(r)\pi(\rho(r))} \Big) e^{\pi(1)} \wedge \ldots \wedge e^{\pi(r)} \\ &= \sum_{\pi \in \mathcal{S}_{n,r}} \det(g_{\alpha(i)\pi(j)}) e^{\pi(1)} \wedge \ldots \wedge e^{\pi(r)} \\ &= g_{\alpha\pi} e^\pi \qquad\qquad \pi \in \mathcal{S}_{n,r}\,. \end{aligned}$$

This gives us formulas for $\Phi$ on $\Lambda^r(V)$ in terms of the subdeterminants of $(g_{ij})$:

$$\Phi(\rho^\alpha e_\alpha) = \overline{\rho^\alpha} E_\alpha = g_{\alpha\pi}\overline{\rho^\alpha} e^\pi \qquad\qquad \alpha, \pi \in \mathcal{S}_{n,r}$$

and similarly

$$\Phi^{-1}(\lambda_\beta e^\beta) = \overline{\lambda_\beta} E^\beta = g^{\pi\beta}\overline{\lambda_\beta} e_\pi \qquad\qquad \beta, \pi \in \mathcal{S}_{n,r}\,.$$

We will use these formulas later to get explicit formulas for the $*$-operators.

## 6.3 The Unit Boxes $\Omega_0$ and $\Omega_0^*$

**The Unit Boxes**

For the definition of of metric forms of the $*$-operator in the next section, it is necessary to fix a *Unit Box* for the vector spaces $V$ and $V^*$. Given any basis $e_1, \ldots, e_n$ for $V$ we may form $\Omega = e_1 \wedge \ldots \wedge e_n$ and then normalize $\Omega$. Since

$$(\Omega, \Omega) = (e_1 \wedge \ldots \wedge e_n, e_1 \wedge \ldots \wedge e_n) = \det(g_{ij}) = g$$

and since $g_{ij}$ is symmetric (for Real vector spaces) or Hermitian (for Complex vector spaces,) we know that $\det(g_{ij})$ is a real number.

Let us recall what happens to $\Omega$ under a change of basis. If $\tilde{e}_i = \alpha_i^j e_j$ then

$$\begin{aligned}
\tilde{\Omega} &= \tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n \\
&= (\alpha_1^{j_1} e_{j_1}) \wedge \ldots \wedge (\alpha_n^{j_n} e_{e_n}) \\
&= \det(\alpha_i^j)\, e_1 \wedge \ldots \wedge e_n \\
&= \det(\alpha_i^j)\, \Omega\,.
\end{aligned}$$

and

$$\begin{aligned}
\tilde{g} &= \det(\tilde{g}_j^i) = \det((\tilde{e}_i, \tilde{e}_j)) = (\tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n, \tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n) \\
&= (\det(\alpha_i^j)\, e_1 \wedge \ldots \wedge e_n, \det(\alpha_k^\ell)\, e_1 \wedge \ldots \wedge e_n) \\
&= \overline{\det(\alpha_i^j)}\, \det(\alpha_k^\ell)(e_1 \wedge \ldots \wedge e_n,\ e_1 \wedge \ldots \wedge e_n) \\
&= |\det(\alpha_i^j)|^2 g\,.
\end{aligned}$$

We can always find an orthonormal basis $f_1, \ldots, f_n$ by the Gram-Schmidt process and arrange things so that

$$\begin{array}{ll}
(f_i, f_i) = +1 & \text{for} \quad 1 \le i \le n - s \\
(f_i, f_i) = -1 & \text{for} \quad n - s + 1 \le i \le n
\end{array}$$

In this case we will have

$$(f_1 \wedge \ldots \wedge f_n, f_1 \wedge \ldots \wedge f_n) = \det((f_i, f_j) = \det \begin{pmatrix}
1 & 0 & \ldots & 0 & & & & \\
 & & \ldots & & & & 0 & \\
0 & 0 & \ldots & 1 & & & & \\
 & & & & -1 & 0 & \ldots & 0 \\
 & & 0 & & & & \ldots & \\
 & & & & 0 & 0 & \ldots & -1
\end{pmatrix} = (-1)^s$$

Now, with our original basis $e_1, \ldots, e_n$, we put

$$e_i = \alpha_i^j f_j, \qquad \Omega = e_1 \wedge \ldots \wedge e_n$$

and using the formula above formula for the change of basis we have

$$\begin{aligned}
g &= (\Omega, \Omega) = (e_1 \wedge \ldots \wedge e_n, e_1 \wedge \ldots \wedge e_n) \\
&= |\det(\alpha_i^j)|^2 (f_1 \wedge \ldots \wedge f_n, f_1 \wedge \ldots \wedge f_n) \\
&= |\det(\alpha_i^j)|^2 (-1)^s\,.
\end{aligned}$$

Hence

$$g = \det(g_{ij}) = |\det(\alpha_i^j)|^2(-1)^s$$

and the *sign* of $\det(g_{ij})$ is fixed by $s$, the number of $f_i$ with $(f_i, f_i) = -1$ in an orthonormal basis. (This $s$ is stable under change of orthonormal basis by Sylvester's Law of Inertia.)

Thus the expression

$$(-1)^s \det(g_{ij}) = (-1)^s g$$

is always a positive number, and we may normalize $\Omega$ to a unit box $\Omega_0$ by dividing $\Omega$ by the square root of this quantity:

$$\Omega_0 = \frac{1}{\sqrt{(-1)^s g}}\Omega\,.$$

Then we have

$$(\Omega_0, \Omega_0) = \frac{1}{(-1)^s g}(\Omega, \Omega) = \frac{1}{(-1)^s g} g = (-1)^s$$

If $V$ is a real vector space, then the one–dimensional real vector space $\Lambda^n(V)$ has room for just two normalized $\Omega_0$, one being the negative of the other. A choice of one of them amounts to choosing an *orientation* for the vector space $V$. Forming an $\tilde{\Omega}_0$ from *any* basis $\tilde{e}_1, \ldots, \tilde{e}_n$ will then result in either $\tilde{\Omega}_0 = \Omega_0$ or $\tilde{\Omega}_0 = -\Omega_0$. A basis $\tilde{e}_1, \ldots, \tilde{e}_n$ is *similarly oriented* to the basis $e_1, \ldots, e_n$ that produced $\Omega_0$ if

$$\tilde{\Omega}_0 = \frac{\tilde{e}_1, \ldots, \tilde{e}_n}{\sqrt{(-1)^s g}} = \Omega_0$$

and *oppositely oriented* if $\tilde{\Omega}_0 = -\Omega_0$. Note that in either case

$$(\tilde{\Omega}_0, \tilde{\Omega}_0) = (\Omega_0, \Omega_0) = (-1)^s\,.$$

Also note that one can shift an oppositely oriented basis to a similarly oriented one by simply changing the sign on any one $\tilde{e}_i$ to $-\tilde{e}_i$.

It is also worth noting that odd–dimensional and even–dimensional vector spaces behave differently if we replace *all* $e_i$ by $-e_i$. If $V$ is even–dimensional, $-e_1, \ldots, -e_n$ is similarly oriented to $e_1, \ldots, e_n$, but if $V$ is odd–dimensional then $-e_1, \ldots, -e_n$ is oppositely oriented to $e_1, \ldots, e_n$. This phenomenon of odd and even–dimensional spaces having differing behavior shows up in a number of different places.

If $V$ is a Complex vector space (the scalars are the Complex numbers) things are not nearly so nice. We have $\Lambda^n(V)$ isomorphic to the Complex numbers. The normalized $\Omega_0$ do not now break down into two easily distinguished classes but instead form a continuum connected to each other by unimodular Complex numbers. (A complex number $\lambda$ is unimodular if and only if $|\lambda| = 1$.) For example, given a basis $e_1, \ldots, e_n$ of $V$ we can form another basis $\tilde{e}_1, \ldots, \tilde{e}_n$ in which

$$\tilde{e}_1 = \lambda e_1, \quad \tilde{e}_2 = e_2, \ldots, \tilde{e}_n = e_n\,.$$

Then

$$\begin{aligned}\tilde{g} &= (\tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n, \tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n) = ((\lambda e_1) \wedge e_2, \wedge \ldots \wedge e_n,\ (\lambda e_1) \wedge e_2, \wedge \ldots \wedge e_n) \\ &= \overline{\lambda}\lambda(e_1 \wedge \ldots \wedge e_n, e_1 \wedge \ldots \wedge e_n) = |\lambda|^2 g\,.\end{aligned}$$

Thus if $|\lambda| = 1$ we have $\tilde{g} = g$ and

$$\tilde{\Omega}_0 = \frac{\tilde{e}_1 \wedge \ldots \wedge \tilde{e}_n}{\sqrt{(-1)^s \tilde{g}}} = \frac{(\lambda e_1) \wedge e_2 \wedge \ldots \wedge e_n}{\sqrt{(-1)^s \tilde{g}}} = \lambda \frac{e_1 \wedge \ldots \wedge e_n}{\sqrt{(-1)^s g}} = \lambda \Omega_0$$

Hence we cannot put an orientation on a Complex vector space; $\Omega_0$ and $-\Omega_0$ are connected by a continuous family $e^{i\theta}\Omega_0$ with $0 \le \theta \le \pi$.

We will use $\Omega_0$ to define the metric $*$-operations in the next section. We have found that they are defined (except for sign) independently of the basis in a *Real* vector space, but in *Complex* vector spaces there is no way to uniquely specify the $\Omega_0$; different bases will result in different $\Omega_0$'s that are connected only through unimodular complex numbers and don't fall into any discrete classes as they do in the real case. We will discuss this further in the next section.

We now consider the analogous unit boxes in $\Lambda^n(V^*)$. We set

$$\Omega^* = e^1 \wedge \ldots \wedge e^n$$

and compute

$$\begin{aligned}(\Omega^*,\ \Omega^*) &= (e^1 \wedge \ldots \wedge e^n,\ e^1 \wedge \ldots \wedge e^n) \\ &= \det((e^i,\ e^j)) \\ &= \det(g^{ij}) = \frac{1}{g} = \frac{(-1)^s}{(-1)^s g}\,.\end{aligned}$$

We normalize by dividing by $\frac{1}{\sqrt{(-1)^s g}}$ to get

$$\Omega_0^* = \sqrt{(-1)^s g}\ \Omega^*$$

so that

$$(\Omega_0^*, \Omega_0^*) = (-1)^s\,.$$

We note that

$$\begin{aligned}\langle \Omega_0^*, \Omega_0 \rangle &= \sqrt{(-1)^s g}\frac{1}{\sqrt{(-1)^s g}}\langle \Omega^*, \Omega \rangle \\ &= \langle e^1 \wedge \ldots \wedge e^n, e_1 \wedge \ldots \wedge e_n \rangle \\ &= \det(\langle e^i, e_j \rangle) = \det(\delta_i^j) \\ &= 1\end{aligned}$$

as we expect for a basis and dual basis element. Next we compute $\Phi$ on these elements $\Omega$ and $\Omega^*$

$$\begin{aligned}\Phi(\Omega) &= \Phi(e_1 \wedge \ldots \wedge e_n) = \Phi(e_1) \wedge \ldots \wedge \Phi(e_n) \\ &= E_1 \wedge \ldots \wedge E_n = g_{1i_1} e^{i_1} \wedge \ldots \wedge g_{ni_n} e^{i_n} \\ &= \det(g_{ij}) e^1 \wedge \ldots \wedge e^n = \det(g_{ij})\Omega^* \\ &= g\Omega^*\end{aligned}$$

and thus

$$\Phi^{-1}(\Omega^*) = \frac{1}{g}\Omega = \det(g^{ij})\Omega$$

and we then have

$$\begin{aligned}\Phi(\Omega_0) &= \tfrac{1}{\sqrt{(-1)^s g}}\Phi(\Omega) = \tfrac{1}{\sqrt{(-1)^s g}} g\Omega^* \\ &= (-1)^s \frac{(-1)^s g}{\sqrt{(-1)^s g}}\Omega^* = (-1)^s\sqrt{(-1)^s g}\,\Omega^* \\ &= (-1)^s\Omega_0^*\end{aligned}$$

and thus

$$\Phi^{-1}(\Omega_0^*) = (-1)^s\Omega_0 = \frac{(-1)^s}{\sqrt{(-1)^s g}}\Omega\,.$$

## 6.4 $*$ Operators Adapted to Inner Products.

When an inner product is available, it is possible to redefine the star operator so as to give an almost involutive isometry between $\Lambda^p(V)$ and $\Lambda^{n-p}(V)$ and similarly for $V^*$. We do this by combining our $* : \Lambda^r(V^*) \to \Lambda^{n-r}(V)$ from the previous chapter with the mapping $\Phi : V \to V^*$, and multiplying by certain constants which force isometry. There is an additional advantage; the reader will recall that $*$ changes by a multiplicative constant when the basis $e_1, \ldots, e_n$ is changed. In the case of a *real* vector space the metric form of $*$ becomes almost independent of the basis; there is a sign change if the bases are oppositely oriented by no other constants appear. In the case of a *complex* vector space $*$ is defined only up to a unimodular ($|\lambda| = 1$) complex number.

We begin our investigation with the following diagram

$$\begin{array}{ccc} \Lambda^p(V^*) & \xrightarrow{*} & \Lambda^{n-p}(V^*) \\ {\scriptstyle \Phi\uparrow\ \downarrow\Phi^{-1}} & \underline{*}\ \overline{*} & {\scriptstyle \Phi^{-1}\downarrow\ \ \uparrow\Phi} \\ \Lambda^p(V) & \xrightarrow[*]{} & \Lambda^{n-p}(V) \end{array}$$

and define

**Def**
$$\begin{array}{rcll} \underline{*}\ell & = & \sqrt{(-1)^s g}(* \circ \Phi^{-1})\ell & \qquad \ell \in \Lambda^p(V^*) \\ \overline{*}v & = & \frac{1}{\sqrt{(-1)^s g}}(* \circ \Phi)v & \qquad v \in \Lambda^p(V) \end{array}$$

The operator $\overline{*}$ is then extended by antilinearity to the whole of $\Lambda^p(V)$, and similarly for $\underline{*}$.

The factors involving $\sqrt{(-1)^s g}$ are inserted to compensate for changes in the $*$-operator when bases are changed. One can determine them by first putting in $k$ and then determining the value of $k$ which will make the next theorem come out as it does, and this is the value of k that leads to invariance. This is not particularly interesting so we have skipped the details.

The reader will note that $\underline{*}$ and $\overline{*}$ could also equally well be defined by $k_1 \Phi \circ *$ and $k_2 \Phi^{-1} \circ *$ with appropriate $k_1$ and $k_2$. We will return to this matter shortly.

To set the stage for the next theorem, which is critical to all that follows let us recall some notation. We have a basis $e_1, \ldots, e_n$ for $V$ with dual basis $e^1, \ldots, e^n$ for $V^*$, $g_{ij} = (e_i, e_j)$, $g = \det(g_{ij})$. If we use the Gram-Schmidt process to form an orthonormal basis $f_1, \ldots, f_n$ from the $e_i$ then $(f_i, f_i)$ is $+1$ for $n - s$ of the $f_i$'s and $-1$ for $s$ of the $f_i$'s. The details are in the previous section. Then we set

$$\begin{array}{rcl} \Omega & = & e_1 \wedge \ldots \wedge e_n \\ \Omega^* & = & e^1 \wedge \ldots \wedge e^n \\ \Omega_0 & = & \left(\dfrac{1}{\sqrt{|(\Omega, \Omega)|}}\right)\Omega = \dfrac{1}{\sqrt{(-1)^s g}}\,\Omega \\ \Omega_0^* & = & \left(\dfrac{1}{\sqrt{|(\Omega^*, \Omega^*)|}}\right)\Omega^* = \sqrt{(-1)^s g}\,\Omega^* \end{array}$$

We now have

**Theorem**

$$\begin{aligned} \ell \wedge \underline{*} m &= (\ell, m)\Omega_0^* & \ell, m \in \Lambda^p(V^*) \\ u \wedge \overline{*} v &= \overline{(u, v)}\Omega_0 & u, v \in \Lambda^p(V) \end{aligned}$$

**Proof**

$$\begin{aligned} \ell \wedge \underline{*} m &= \ell \wedge \sqrt{(-1)^s g}(* \circ \Phi^{-1})m \\ &= \sqrt{(-1)^s g}(\ell \wedge *\Phi^{-1}(m)) \\ &= \sqrt{(-1)^s g}\langle \ell, \Phi^{-1}(m)\rangle \Omega^* \\ &= (\ell, m)\Omega_0^* \\ u \wedge \overline{*} v &= u \wedge \frac{1}{\sqrt{(-1)^s g}}(* \circ \Phi)v \\ &= \frac{1}{\sqrt{(-1)^s g}}(u \wedge *\Phi(v)) \\ &= \frac{1}{\sqrt{(-1)^s g}}\langle \Phi(v), u\rangle \Omega \\ &= (v, u)\Omega_0 \\ &= \overline{(u, v)}\Omega_0 \end{aligned}$$

This theorem conceals a critical fact; since the value of $\overline{*}v$ is completely determined by the values of $u \wedge \overline{*}v$, we see that $\overline{*}v$ is just as well determined as $\Omega_0$. As we saw in the last section, for *real* vector spaces $\Omega_0$ is uniquely determined up to sign, and the same goes for $\Omega_0^*$ and $\underline{*}\ell$. Hence in this case

**Corollary** For a vector space with *real* scalars, $\overline{*}v$ and $\underline{*}\ell$ are uniquely defined up to sign. More precisely, if the bases $e_1, \ldots, e_n$ and $\tilde{e}_1, \ldots, \tilde{e}_n$ are used to compute $\overline{*}_e v$ and $\overline{*}_{\tilde{e}} v$ then

$$\begin{aligned} \overline{*}_{\tilde{e}} v &= +\overline{*}_e v & &\text{if } \tilde{e}_1, \ldots, \tilde{e}_n \text{ and } e_1, \ldots, e_n \text{ are similarly oriented} \\ \overline{*}_{\tilde{e}} v &= -\overline{*}_e v & &\text{if } \tilde{e}_1, \ldots, \tilde{e}_n \text{ and } e_1, \ldots, e_n \text{ are oppositely oriented.} \end{aligned}$$

and the same is true for $\underline{*}\ell$. For the case of a *complex* vector space, as we saw in the previous section, it is not possible to define $\Omega_0$ and $\Omega_0^*$ uniquely; they will change by a unimodular complex number when the basis is changed. Hence in this case $\overline{*}v$ and $\underline{*}\ell$ are defined only up to unimodular complex numbers.

We now return to the alternate possibilities for defining $\underline{*}$ and $\overline{*}$. Once again we skip the dull details of deriving the values of the constants and simply verify the final results.

**Theorem**

$$\begin{aligned} \underline{*}\ell &= \frac{(-1)^s}{\sqrt{(-1)^s g}}(\Phi \circ *)\ell \\ \overline{*}v &= (-1)^s \sqrt{(-1)^s g}\,(\Phi^{-1} \circ *)v \end{aligned}$$

**Proof** We present the proof in detail because care is necessary not to lose minus signs or conjugate bars. For any $\ell \in V^*$

$$\begin{aligned}
\ell \wedge \frac{(-1)^s}{\sqrt{(-1)^s g}}(\Phi \circ *)m &= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\ell \wedge \Phi(*m) \\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\Phi(\Phi^{-1}(\ell) \wedge *m) \\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\Phi(\langle m, \Phi^{-1}(\ell)\rangle \Omega) \\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\overline{\langle m, \Phi^{-1}(\ell)\rangle}\Phi(\Omega) \\
&= \overline{\langle m, \Phi^{-1}(\ell)\rangle}\frac{(-1)^s}{\sqrt{(-1)^s g}}(g\Omega^*) \\
&= \overline{\langle \Phi\Phi^{-1}(m), \Phi^{-1}(\ell)\rangle}\frac{(-1)^s g}{\sqrt{(-1)^s g}}\,\Omega^* \\
&= \langle \ell, \Phi^{-1}(m)\rangle \sqrt{(-1)^s g}\,\Omega^* \\
&= (\ell, m)\Phi_0^* \\
&= \ell \wedge *m
\end{aligned}$$

Since this is true for all $\ell \in \Lambda(V^*)$,

$$*m = \frac{(-1)^s}{\sqrt{(-1)^s g}}(\Phi \circ *)m$$

as required.

In a similar way, for any $u \in \Lambda^p(V)$

$$\begin{aligned}
u \wedge (-1)^s\sqrt{(-1)^s g}(\Phi^{-1} \circ *)(v) &= (-1)^s\sqrt{(-1)^s g}\,(u \wedge \Phi^{-1}(*v)) \\
&= (-1)^s\sqrt{(-1)^s g}\,\Phi^{-1}(\Phi(u) \wedge *v) \\
&= (-1)^s\sqrt{(-1)^s g}\,\Phi^{-1}(\langle \Phi(u), *v\rangle \Omega^*) \\
&= (-1)^s\sqrt{(-1)^s g}\,\overline{\langle \Phi(u), *v\rangle}\Phi^{-1}(\Omega^*) \\
&= (-1)^s\sqrt{(-1)^s g}\,\overline{\langle \Phi(u), *v\rangle}\frac{1}{g}\Omega \\
&= \overline{(u, v)}\frac{\sqrt{(-1)^s g}}{(-1)^s g}\Omega \\
&= \overline{(u, v)}\frac{1}{\sqrt{(-1)^s g}}\Omega \\
&= \overline{(u, v)}\Omega_0 \\
&= u \wedge \overline{*}v
\end{aligned}$$

Since this is true for all $u \in \Lambda^p(V)$

$$\overline{*}v = (-1)^s\sqrt{(-1)^s g}\,(\Phi^{-1} \circ *)(v)$$

as desired.

Using the definition and the two previous theorems, we can easily show that $\underline{*}$ and $\overline{*}$ are almost, but not quite, involutive operators and isometries. Notice in the proof that we use both formulas for $\overline{*}$, one from the definition and one from the theorem; having both formulas is what makes it possible to present an easy basis–free proof of this theorem. Having these two formulas in turn depends on our methodology of factoring $\overline{*} : \Lambda^p(V) \to \Lambda^{n-p}(V)$ into $\Phi : \Lambda^p(v) \to \Lambda^p(V^*)$ and $* : \Lambda^p(V^*) \to \Lambda^{n-p}(V)$. The common methodology for proving this theorem uses an orthonormal basis. The reason this works is that if the basis is orthonormal the $\Phi$ can be almost ignored.

**Theorem**

$$\begin{aligned} \underline{*}\,\underline{*}\ell &= (-1)^{p(n-p)+s}\ell, \quad \ell \in \Lambda^p(V^*) \\ \overline{*}\,\overline{*}v &= (-1)^{p(n-p)+s}v, \quad v \in \Lambda^p(V) \end{aligned}$$

**Proof**

$$\begin{aligned} \underline{*}\,\underline{*}\ell &= \sqrt{(-1)^s g}(*\circ\Phi)(\underline{*}\ell) \\ &= \sqrt{(-1)^s g}(*\circ\Phi)\circ\frac{(-1)^s}{\sqrt{(-1)^s g}}\Phi^{-1}\circ *)\ell \\ &= (-1)^s(*\circ\Phi\circ\Phi^{-1}\circ *)\ell \\ &= (-1)^s(*\circ *)\ell \\ &= (-1)^s(-1)^{p(n-p)}\ell \\ &= (-1)^{p(n-p)+s}\ell \end{aligned}$$

and similarly

$$\begin{aligned} \overline{*}\,\overline{*}v &= \frac{1}{\sqrt{(-1)^s g}}(*\circ\Phi)(\overline{*}v) \\ &= \frac{1}{\sqrt{(-1)^s g}}(*\circ\Phi)((-1)^s\sqrt{(-1)^s g}\Phi^{-1}\circ *)v \\ &= (-1)^s(*\circ\Phi\circ\Phi^{-1}\circ *)v \\ &= (-1)^s(*\circ *)v \\ &= (-1)^s(-1)^{p(n-p)}v \\ &= (-1)^{p(n-p)+s}v \end{aligned}$$

**Theorem** $\underline{*}$ and $\overline{*}$ are anti-isometries:

$$\begin{array}{rcccl} (\underline{*}\ell, \underline{*}m) & = & \overline{(\ell, m)} & = & (m, \ell) \\ (\overline{*}u, \overline{*}v) & = & \overline{(u, v)} & = & (v, u) \end{array}$$

**Proof**

$$\begin{aligned} (\underline{*}\ell, \underline{*}m)\Omega_0^* &= \underline{*}\ell \wedge \underline{*}\,\underline{*}m \\ &= (-1)^{p(n-p)}\underline{*}\ell \wedge m \\ &= m \wedge \underline{*}\ell \\ &= (m, \ell)\Omega_0^* \end{aligned}$$

and similarly

$$\begin{aligned} \overline{(\overline{*}u, \overline{*}v)}\Omega_0 &= \overline{*}u \wedge \overline{*}\,\overline{*}v \\ &= (-1)^{p(n-p)}\overline{*}u \wedge v \\ &= v \wedge \overline{*}u \\ &= \overline{(v, u)}\Omega_0 \end{aligned}$$

For convenience we want to exhibit the formulas for certain special cases. We have, for $\Omega \in \Lambda^n(V)$

$$\begin{aligned} \overline{*}\Omega &= (-1)^s \sqrt{(-1)^s g}\,(\Phi^{-1} \circ *)\Omega \\ &= (-1)^s \sqrt{(-1)^s g}\;\Phi^{-1}(1) \\ &= (-1)^s \sqrt{(-1)^s g}\;1 \\ &= (-1)^s \sqrt{(-1)^s g} \end{aligned}$$

and then

$$\begin{aligned} \overline{*}\Omega_0 &= \overline{*}\Big(\frac{1}{\sqrt{(-1)^s g}}\;\Omega\Big) \\ &= \frac{1}{\sqrt{(-1)^s g}}\;\overline{*}\Omega \\ &= \frac{1}{\sqrt{(-1)^s g}}(-1)^s \sqrt{(-1)^s g} \\ &= (-1)^s\,. \end{aligned}$$

and then for $1 \in \Lambda^0(V)$

$$\begin{aligned} \overline{*}\,1 &= \frac{1}{\sqrt{(-1)^s g}}(* \circ \Phi)\,1 \\ &= \frac{1}{\sqrt{(-1)^s g}}(*\,1) \\ &= \frac{1}{\sqrt{(-1)^s g}}\;\Omega \\ &= \Omega_0 \end{aligned}$$

Similarly, for $\Omega^* \in \Lambda^n(V^*)$

$$\begin{aligned}
\underline{*}\Omega^* &= \frac{(-1)^s}{\sqrt{(-1)^s g}}(\Phi \circ *)\Omega^* \\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\Phi(1) \\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,1 \\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}
\end{aligned}$$

and then

$$\begin{aligned}
\underline{*}\Omega_0^* &= \underline{*}\big(\sqrt{(-1)^s g}\,\Omega^*\big) \\
&= \sqrt{(-1)^s g}\Big(\frac{(-1)^s}{\sqrt{(-1)^s g}}\Big) \\
&= (-1)^s
\end{aligned}$$

and finally for $1 \in \Lambda^0(V^*)$

$$\begin{aligned}
\underline{*}\,1 &= \sqrt{(-1)^s g}\,(* \circ \Phi^{-1})\,1 \\
&= \sqrt{(-1)^s g}\,(*1) \\
&= \sqrt{(-1)^s g}\,\Omega^* \\
&= \Omega_0^*\,.
\end{aligned}$$

## 6.5 Coordinate formulas for $*$-Operators

The $*$-operators are often underutilized in applications because the available technology for them is cumbersome. What we hope to show here that using the metric coefficients $g_{\alpha\beta}$ for $\Lambda^p(V)$ and $g^{\alpha\beta}$ for $\Lambda^p(V^*)$ where $\alpha, \beta \in \mathcal{S}_{n,p}$ we can recapture some of the ease of computation of classical tensor analysis. The reader may recall our earlier contention that using $\mathcal{S}_{n,p}$ as the indexing set is the key to efficient use of Grassmann algebra.

We will now use the formulas from the last section to derive explicit formulas for the $*$-operators, which are important for applications. As usual let $e_1, \dots, e_n$ be a basis of $V$ and $e^1, \dots, e^n$ be the dual basis for $V^*$. We recall the formulas

$$\begin{aligned} \overline{*}v &= (-1)^s\sqrt{(-1)^s g}(\Phi^{-1}\circ *)v \\ \Phi^{-1}(e^i) &= E^i = g^{ji}e_j\,. \end{aligned}$$

Then we have

$$\begin{aligned} \overline{*}e_\sigma &= (-1)^s\sqrt{(-1)^s g}(\Phi^{-1}\circ *)e_\sigma \\ &= (-1)^s\sqrt{(-1)^s g}\,\Phi^{-1}(\operatorname{sgn}(\sigma)e^{\tilde\sigma}) \\ &= (-1)^s\sqrt{(-1)^s g}\,\operatorname{sgn}(\sigma)\Phi^{-1}(e^{\tilde\sigma(p+1)}\wedge\ldots\wedge e^{\tilde\sigma(n)}) \\ &= (-1)^s\sqrt{(-1)^s g}\,\operatorname{sgn}(\sigma)\Phi^{-1}(e^{\tilde\sigma(p+1)})\wedge\ldots\wedge \Phi^{-1}(e^{\tilde\sigma(n)}) \\ &= (-1)^s\sqrt{(-1)^s g}\,\operatorname{sgn}(\sigma)E^{\tilde\sigma(p+1)}\wedge\ldots\wedge E^{\tilde\sigma(n)} \\ &= (-1)^s\sqrt{(-1)^s g}\,\operatorname{sgn}(\sigma)g^{i_1\tilde\sigma(p+1)}\ldots g^{i_{n-p}\tilde\sigma(n)}e_{i_1}\wedge\ldots\wedge e_{i_{n-p}} \\ &= (-1)^s\sqrt{(-1)^s g}\,\operatorname{sgn}(\sigma)g^{\tau\tilde\sigma}e_\tau \qquad \text{where } \tau\in\mathcal{S}_{n,n-p}\,. \end{aligned}$$

In the last line we have used the method of resolving a sum by permutations. See section 3.3 or the end of section 5.1 for more detailed explanations of this method.

Similarly, using the formulas

$$\begin{aligned} \underline{*}\ell &= \frac{(-1)^s}{\sqrt{(-1)^s g}}(\Phi\circ *)\ell \\ \Phi(e_i) &= E_i = g_{ik}e^k \end{aligned}$$

we get, in exactly the same way,

$$\underline{*}e^\sigma = \frac{(-1)^s}{\sqrt{(-1)^s g}}\operatorname{sgn}(\sigma)g_{\tilde\sigma\tau}e^\tau \qquad \tau\in\mathcal{S}_{n,n-p}\,.$$

The natural thing to do next would be to use the alternate formulas for $\overline{*}$ and $\underline{*}$ to get alternate formulas for $\overline{*}e_\sigma$ and $\underline{*}e^\sigma$. However, we think it is worth disturbing the logical flow in order to learn another technique. We will resume the logical flow in a moment.

In many books on differential forms a technique like the following is used for computing $\underline{*}e^\sigma$, (interest in this case being focussed on the dual space $V^*$). We begin with the important equation (sometimes used as the definition of $\underline{*}$).

$$\ell \wedge \underline{*}m = (\ell, m)\Omega_0^*$$

where $\Omega_0^*$ is the unit box in $V^*$:

$$\Omega_0^* = \sqrt{(-1)^s g}\, \Omega^* = \sqrt{(-1)^s g}\, e^1 \wedge \ldots \wedge e^n\,.$$

We use this to compute $\underline{*}e^\sigma$, $\sigma \in \mathcal{S}_{n,p}$ by setting

$$\underline{*}e^\sigma = \alpha_\tau e^\tau \qquad\qquad \tau \in \mathcal{S}_{n,n-p}$$

for some constants $\alpha_\tau$, and then attempt to determine $\alpha_\tau$ by some kind of trickery resembling that used in orthogonal expansions. We substitute $e^\rho$ for $\ell$ and $e^\sigma$ for $m$ in the above formula and then use the expansion of $\underline{*}e^\sigma$ to get

$$\begin{aligned} e^\rho \wedge \underline{*}e^\sigma &= (e^\rho,\, e^\sigma)\Omega_0^* \\ e^\rho \wedge \alpha_\tau e^\tau &= g^{\rho\sigma}\Omega_0^* \\ \alpha_\tau\, e^\rho \wedge e^\tau &= g^{\rho\sigma}\Omega_0^*\,. \end{aligned}$$

Since $\rho \in \mathcal{S}_{n,p}$ and $\tau \in \mathcal{S}_{n,n-p}$, the left side of the last equation is non–zero for exactly one $\tau$, namely $\tau = \tilde{\rho}$. We then have

$$\alpha_{\tilde{\rho}}\, e^\rho \wedge e^{\tilde{\rho}} = g^{\rho\sigma}\Omega_0^* \qquad\qquad (\text{no sum on } \rho)$$

and then

$$\alpha_{\tilde{\rho}}\, \mathrm{sgn}(\rho) e^1 \wedge \ldots \wedge e^n = g^{\rho\sigma} \sqrt{(-1)^s g}\, e^1 \wedge \ldots \wedge e^n$$

and thus

$$\alpha_{\tilde{\rho}} = \mathrm{sgn}(\rho)\, \sqrt{(-1)^s g}\, g^{\rho\sigma} \qquad\qquad \rho \in \mathcal{S}_{n,p}$$

which we can rewrite in a handier form, setting $\tau = \tilde{\rho}$:

$$\begin{aligned} \alpha_\tau &= \mathrm{sgn}(\tilde{\tau})\, \sqrt{(-1)^s g}\, g^{\tilde{\tau}\sigma} \\ &= (-1)^{p(n-p)} \mathrm{sgn}(\tau)\, \sqrt{(-1)^s g}\, g^{\tilde{\tau}\sigma} \qquad\qquad \tau \in \mathcal{S}_{n,n-p} \end{aligned}$$

since $\mathrm{sgn}(\tau)\mathrm{sgn}(\tilde{\tau}) = (-1)^{p(n-p)}$ (see section 3.2). Finally

$$\underline{*}e^\sigma = \alpha_\tau e^\tau = \sqrt{(-1)^s g} \sum_{\tau \in \mathcal{S}_{n,n-p}} \mathrm{sgn}(\tilde{\tau}) g^{\tilde{\tau}\sigma} e^\tau\,.$$

This formula is correct, but the perceptive will note that it bears little resemblance to the previously derived formula. The equivalence of the two formulas is not difficult to prove *provided* one has some specialized tools and knowledge (found in this book) but when working on ones own this sort of thing can be a real source of frustration, and the frustration is augmented when the formulas are expressed with $g^{\tilde{\tau}\sigma}$ written out in full determinental form.

We can derive this formula much more easily by using the other formula for $\underline{*}\ell$:

$$\underline{*}\ell = \sqrt{(-1)^s g}\,(*\circ\Phi^{-1})\ell\,.$$

We then have

$$\begin{aligned}
\underline{*}e^\sigma &= \sqrt{(-1)^s g}\,(*\circ\Phi^{-1})e^\sigma \\
&= \sqrt{(-1)^s g}\ * (\Phi^{-1}(e^{\sigma(1)}) \wedge \ldots \wedge \Phi^{-1}(e^{\sigma(p)}) \\
&= \sqrt{(-1)^s g}\ * \big(g^{i_1\sigma(1)} g^{i_2\sigma(2)} \ldots g^{i_p\sigma(p)} (e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p})\big) \\
&= \sqrt{(-1)^s g}\ * \Big(\sum_{\tau\in\mathcal{S}_{n,p}} \Big[\sum_{\rho\in\mathcal{S}_p} \operatorname{sgn}(\rho) g^{\tau(\rho(1))\sigma(1)} \ldots g^{\tau(\rho(p))\sigma(p)}\Big] e_{\tau(1)} \wedge \ldots \wedge e_{\tau(p)}\Big) \\
&= \sqrt{(-1)^s g} \sum_{\tau\in\mathcal{S}_{n,p}} \operatorname{sgn}(\tau)\, \overline{g^{\tau\sigma}} e^{\tilde{\tau}} \\
&= \sqrt{(-1)^s g} \sum_{\tau\in\mathcal{S}_{n,p}} \operatorname{sgn}(\tau)\, g^{\sigma\tau} e^{\tilde{\tau}}\,.
\end{aligned}$$

Similarly, using

$$\overline{*}v = \frac{1}{\sqrt{(-1)^s g}}(*\circ\Phi)v$$

we derive

$$\overline{*}e_\sigma = \frac{1}{\sqrt{(-1)^s g}} \sum_{\tau\in\mathcal{S}_{n,p}} \operatorname{sgn}(\tau)\, g_{\tau\sigma} e_{\tilde{\tau}}\,.$$

For convenience of reference, we collect here the various formulas involving the star operator.

**Formulas**

$$\begin{aligned}
\underline{*}\ell &= \sqrt{(-1)^s g}(*\circ\Phi^{-1})\ell && \ell\in\Lambda^p(V^*)\\
&= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,(\Phi\circ *)\ell\\
\underline{*}e^\sigma &= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\mathrm{sgn}(\sigma)g_{\tilde\sigma\tau}e^\tau && \tau\in\mathcal{S}_{n,n-p}\\
\underline{*}e^\sigma &= \sqrt{(-1)^s g}\sum_{\tau\in\mathcal{S}_{n,p}}\mathrm{sgn}(\tau)\,g^{\sigma\tau}e^{\tilde\tau}\\
\overline{*}v &= \frac{1}{\sqrt{(-1)^s g}}(*\circ\Phi)v && v\in\Lambda^p(V)\\
&= (-1)^s\sqrt{(-1)^s g}\,(\Phi^{-1}\circ *)v\\
\underline{*}e^\sigma &= \frac{(-1)^s}{\sqrt{(-1)^s g}}\,\mathrm{sgn}(\sigma)g_{\tilde\sigma\tau}e^\tau && \tau\in\mathcal{S}_{n,n-p}\\
\overline{*}e_\sigma &= \frac{1}{\sqrt{(-1)^s g}}\sum_{\tau\in\mathcal{S}_{n,p}}\mathrm{sgn}(\tau)\,g_{\sigma\tau}e_{\tilde\tau}
\end{aligned}$$

$$\begin{aligned}
\ell\wedge\underline{*}m &= (\ell,m)\Omega_0^* && \ell,m\in\Lambda^p(V^*)\\
u\wedge\overline{*}v &= \overline{(u,v)}\Omega_0 && u,v\in\Lambda^p(V)
\end{aligned}$$

$$\begin{aligned}
\Omega^* &= e^1\wedge\ldots\wedge e^n\\
\Omega_0^* &= \sqrt{(-1)^s g}\Omega^*\\
\underline{*}\Omega^* &= \frac{(-1)^s}{\sqrt{(-1)^s g}}\\
\underline{*}\Omega_0^* &= (-1)^s\\
\underline{*}1 &= \Omega_0^*
\end{aligned}$$

$$\begin{aligned}
\Omega &= e_1\wedge\ldots\wedge e_n\\
\Omega_0 &= \frac{1}{\sqrt{(-1)^s g}}\Omega\\
\overline{*}\Omega &= (-1)^s\sqrt{(-1)^s g}\\
\overline{*}\Omega_0 &= (-1)^s\\
\overline{*}1 &= \Omega_0
\end{aligned}$$

## 6.6 Formulas for Orthogonal bases

Although one of the basic principles of this book is never to derive formulas using orthogonal or orthonormal bases, we recognize the importance of these objects in applications. Hence this section will be devoted to the calculation of the $*$-operator in $V^*$ for orthogonal bases.

We are going to do this in two different ways. First, we will show how to specialize our general formulas to this case. Second, we will show how to calculate these formulas directly from the definition. The latter method is included for the convenience of those persons who wish to utilize the $*$-operator in classes but do not have the time (or perhaps the inclination) to go through the general procedure.

We will derive all our formulas for $V^*$. The case of $V$ is handled analogously, and the formulas are predictable from those of $V^*$.

To make the formulas as simple as possible we will simplify the notation, using $h_i$ for $\sqrt{g^{ii}}$. This destroys the systematic applicability of the summation convention, which is suspended for this section. We will continue to use increasing permutations as one of our basic tools, since without them the $*$-operator is very hard to manage.

What characterizes an orthogonal basis is that $g^{ij} = 0$ for $i \neq j$. We will set things up like this:

$$(g^{ij}) = \begin{pmatrix} h_1^2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & 0 & \dots & 0 \\ 0 & \dots & h_r^2 & 0 & \dots & 0 \\ 0 & \dots & 0 & -h_{r+1}^2 & \dots & 0 \\ 0 & \dots & 0 & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & -h_n^2 \end{pmatrix}$$

Thus the last $s = n - r$ diagonal coefficients have negative signs. A typical example would be the inner product used in special relativity with coordinates $t, x, y, z$ (in that order) and matrix

$$(g^{ij}) = \begin{pmatrix} (\frac{1}{c})^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Here $n = 4$ and $s = 3$.

The general formula for the $\underline{*}$-operator is given by

$$\underline{*}e^\sigma = \sqrt{(-1)^s g} \sum_{\tau \in \mathcal{S}_{n,p}} \text{sgn}(\sigma) g^{\sigma\tau} e^{\tilde{\tau}}$$

where the $g^{\sigma\tau}$ are the $p \times p$ subdeterminants of $(g^{ij})$. (We have chosen this formula for $\underline{*}e^\sigma$ because it uses the entries from the $(g^{ij})$ and is thus most convenient.)

We first discuss the situation for $g^{\sigma\tau}$ where $\sigma \neq \tau$. We will have $g^{\sigma\tau} = 0$ as we see from the following example which demonstrates how a row of zeros will appear in this case. We take $\sigma, \tau \in \mathcal{S}_{n,2}$

$$\sigma = \left(\begin{array}{cc|cccc} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 3 & 1 & 4 & \dots & n \end{array}\right) \qquad \tau = \left(\begin{array}{cc|cccc} 1 & 2 & 3 & 4 & \dots & n \\ 2 & 4 & 1 & 3 & \dots & n \end{array}\right).$$

We then have

$$g^{\sigma\tau} = \det \begin{pmatrix} h_2^2 & 0 \\ 0 & 0 \end{pmatrix}$$

because the entries for the second row must come from the third $(= \sigma(2))$ row of the $(g^{ij})$ and the only non zero entry in that row is in the third column. However, $\tau$ selects columns 2 and 4, thus missing the only non-zero entry.

Hence $g^{\sigma\tau}$ is diagonal; for $\sigma = \tau \in \mathcal{S}_{n,p}$ it selects a set of rows and equally numbered columns to form a diagonal submatrix of $(g^{ij})$. For example, with the above $\sigma$ and $\tau$ we have

$$g^{\sigma\sigma} = \det \begin{pmatrix} h_2^2 & 0 \\ 0 & h_3^2 \end{pmatrix} \qquad g^{\tau\tau} = \det \begin{pmatrix} h_2^2 & 0 \\ 0 & h_4^2 \end{pmatrix}$$

(We have assumed here that $4 \leq r$). Recall now that $(g^{ij}) = (g_{ij})^{-1}$ and $g = \det(g_{ij})$ from which we obtain

$$g = \frac{1}{h_1^2 \dots h_{n-r}^2(-h_{n-r+1}^2) \dots (-h_n^2)} = \frac{1}{(-1)^s (h_1 \dots h_n)^2}$$

so that

$$\sqrt{(-1)^s g} = \frac{1}{h_1 \dots h_n}.$$

Now suppose $\sigma \in \mathcal{S}_{n,p}$ and

$$\sigma(1), \dots, \sigma(p-b) \leq r \text{ and } \sigma(p-b+1), \dots, \sigma(p) > r.$$

Then exactly $p - (p-b+1) + 1 = b$ of the entries of

$$\begin{pmatrix} h_{\sigma(1)}^2 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & 0 & \dots & 0 \\ 0 & \dots & h_{\sigma(p-b)}^2 & 0 & \dots & 0 \\ 0 & \dots & 0 & -h_{\sigma(p-b+1)}^2 & \dots & 0 \\ 0 & \dots & 0 & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & -h_{\sigma(p)}^2 \end{pmatrix}$$

will be negative, so that the determinant of this matrix will be

$$g^{\sigma\sigma} = (-1)^b h_{\sigma(1)}^2 \dots h_{\sigma(p)}^2$$

and

$$\begin{aligned} \sqrt{(-1)^s g}\, g^{\sigma\sigma} &= \frac{1}{h_1 \dots h_n} (-1)^b h_{\sigma(1)}^2 \dots h_{\sigma(p)}^2 \\ &= (-1)^b \frac{h_{\sigma(1)} \dots h_{\sigma(p)}}{h_{\sigma(p+1)} \dots h_{\sigma(n)}}. \end{aligned}$$

Thus finally we have

$$\begin{aligned}\underline{*}e^{\sigma} &= \sqrt{(-1)^s g}\sum_{\tau\in\mathcal{S}_{n,p}} \mathrm{sgn}(\sigma) g^{\sigma\tau} e^{\tilde{\tau}}\\ &= \sqrt{(-1)^s g}\,\mathrm{sgn}(\sigma) g^{\sigma\sigma} e^{\tilde{\sigma}}\\ &= (-1)^b \frac{h_{\sigma(1)}\dots h_{\sigma(p)}}{h_{\sigma(p+1)}\dots h_{\sigma(n)}}\mathrm{sgn}(\sigma)e^{\tilde{\sigma}}\,.\end{aligned}$$

Just as a quick check, let us calculate $\underline{*}\,\underline{*}e^{\sigma}$. We have

$$\begin{aligned}\underline{*}\,\underline{*}e^{\sigma} &= (-1)^b \frac{h_{\sigma(1)}\dots h_{\sigma(p)}}{h_{\sigma(p+1)}\dots h_{\sigma(n)}}\mathrm{sgn}(\sigma)\underline{*}e^{\tilde{\sigma}}\\ &= (-1)^b \frac{h_{\sigma(1)}\dots h_{\sigma(p)}}{h_{\sigma(p+1)}\dots h_{\sigma(n)}}\mathrm{sgn}(\sigma)(-1)^{s-b}\frac{h_{\tilde{\sigma}(1)}\dots h_{\tilde{\sigma}(n-p)}}{h_{\tilde{\sigma}(n-p+1)}\dots h_{\tilde{\sigma}(n)}}\mathrm{sgn}(\tilde{\sigma})e^{\sigma}\\ &= (-1)^s\mathrm{sgn}(\sigma)\mathrm{sgn}(\tilde{\sigma})\frac{h_{\sigma(1)}\dots h_{\sigma(p)}}{h_{\sigma(p+1)}\dots h_{\sigma(n)}}\frac{h_{\sigma(p+1)}\dots h_{\sigma(n)}}{h_{\sigma(1)}\dots h_{\sigma(p)}}e^{\sigma}\\ &= (-1)^s(-1)^{p(n-p)}e^{\sigma} = (-1)^{s+p(n-p)}e^{\sigma}\end{aligned}$$

which is correct, as we have seen in section 5.3.

Now we want to apply this formula, as an example, to the special relativity metric

$$\begin{pmatrix} (\frac{1}{c})^2 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}.$$

However, in order to be consistent with established custom and our future needs for this material we are going to modify some of our usual notations. In this context it is usual to index with the numbers $0, 1, 2, 3$ rather than $1, 2, 3, 4$ and to replace the basis vectors $e_1, e_2, e_3, e_4$ with $dt, dx^1, dx^2, dx^3$. The reasons for this will become clear in Chapter 9. We know that for normally indexed increasing permutations we can get the sign of the permutation by

$$\mathrm{sgn}(\sigma) = (-1)^{\sum_{j=1}^{p}\sigma(j)-T_p} = (-1)^{\sum_{j=1}^{p}(\sigma(j)-j)}\,.$$

If we reset the origin from 1 to 0, the exponent becomes

$$\sum_{i=0}^{p-1}\big((\sigma(i)+1)-(i+1)\big) = \sum_{i=0}^{p-1}(\sigma(i)-i) = \sum_{i=0}^{p-1}\sigma(i) - T_{p-1}$$

and

$$\mathrm{sgn}(\sigma) = (-1)^{\sum_{i=0}^{p-1}\sigma(i)-T_{p-1}}$$

In this formula we put $T_0 = 0$ when $p = 1$.
We give some examples:

$$\sigma = \left(\begin{array}{c|ccc} 0 & 1 & 2 & 3\\ 1 & 0 & 2 & 3\end{array}\right) \qquad \mathrm{sgn}(\sigma) = (-1)^{1-T_0} = (-1)^{1-0} = -1$$

$$\sigma = \begin{pmatrix} 0 & 1 & \vline & 2 & 3 \\ 0 & 3 & \vline & 1 & 2 \end{pmatrix} \qquad \mathrm{sgn}(\sigma) = (-1)^{0+3-T_1} = (-1)^{3-1} = +1$$

$$\sigma = \begin{pmatrix} 0 & 1 & \vline & 2 & 3 \\ 1 & 3 & \vline & 0 & 2 \end{pmatrix} \qquad \mathrm{sgn}(\sigma) = (-1)^{1+3-T_1} = (-1)^{4-1} = -1$$

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & \vline & 3 \\ 0 & 1 & 3 & \vline & 2 \end{pmatrix} \qquad \mathrm{sgn}(\sigma) = (-1)^{0+1+3-T_2} = (-1)^{4-3} = -1$$

With these and the formula for $\underline{*}$ we can compute

$$\underline{*}\, dx^1 = (-1)^1 \frac{1}{\frac{1}{c}\cdot 1 \cdot 1}(-1)\, dt \wedge dx^2 \wedge dx^3 = +c\; dt \wedge dx^2 \wedge dx^3$$

$$\underline{*}\, dt \wedge dx^3 = (-1)^1 \frac{\frac{1}{c}\cdot 1}{1\cdot 1}(+1)\, dx^1 \wedge dx^2 = -\frac{1}{c}\; dx^1 \wedge dx^2$$

$$\underline{*}\, dx^1 \wedge dx^3 = (-1)^2 \frac{1\cdot 1}{\frac{1}{c}\cdot 1}(-1)\, dt \wedge dx^2 = -c\; dt \wedge dx^2$$

$$\underline{*}\, dt \wedge dx^1 \wedge dx^3 = (-1)^2 \frac{\frac{1}{c}\cdot 1 \cdot 1}{1}(-1)\, dt \wedge dx^2 = -\frac{1}{c}\; dx^1$$

Now, using the formula $\underline{*}\,\underline{*} e^\sigma = (-1)^{s+p(n-p)} e^\sigma$ each of the above four gives us a second formula. For example from the second of the above four we have

$$\begin{aligned} \underline{*}\,\underline{*}\, dt \wedge dx^3 &= -\frac{1}{c}\; \underline{*}\, dx^1 \wedge dx^2 \\ (-1)^{3+2(4-2)}\, dt \wedge dx^3 &= -\frac{1}{c}\; \underline{*}\, dx^1 \wedge dx^2 \\ \underline{*}\, dx^1 \wedge dx^2 &= (-1)\cdot(-c)\, dt \wedge dx^3 \\ &= c\, dt \wedge dx^3\,. \end{aligned}$$

We can now digest all this and more in the following table

$$\underline{*}\, 1 = c\, dt \wedge dx^1 \wedge dx^2 \wedge dx^3$$

$$\begin{array}{rclrcl} \underline{*}\, dt &=& +(1/c) dx^1 \wedge dx^2 \wedge dx^3 & \underline{*}\, dx^1 \wedge dx^2 \wedge dx^3 &=& +c\, dt \\ \underline{*}\, dx^1 &=& +c\, dt \wedge dx^2 \wedge dx^3 & \underline{*}\, dt \wedge dx^2 \wedge dx^3 &=& +(1/c)\, dx^1 \\ \underline{*}\, dx^2 &=& -c\, dt \wedge dx^1 \wedge dx^3 & \underline{*}\, dt \wedge dx^1 \wedge dx^3 &=& -(1/c)\, dx^2 \\ \underline{*}\, dx^3 &=& +c\, dt \wedge dx^1 \wedge dx^2 & \underline{*}\, dt \wedge dx^1 \wedge dx^2 &=& +(1/c)\, dx^3 \end{array}$$

$$\begin{array}{rclrcl} \underline{*}\, dt \wedge dx^1 &=& -(1/c) dx^2 \wedge dx^3 & \underline{*}\, dx^2 \wedge dx^3 &=& +c\, dt \wedge dx^1 \\ \underline{*}\, dt \wedge dx^2 &=& +(1/c) dx^1 \wedge dx^3 & \underline{*}\, dx^1 \wedge dx^3 &=& -c\, dt \wedge dx^2 \\ \underline{*}\, dt \wedge dx^3 &=& -(1/c) dx^1 \wedge dx^2 & \underline{*}\, dx^1 \wedge dx^2 &=& +c\, dt \wedge dx^3 \end{array}$$

$$\underline{*}\, dt \wedge dx^1 \wedge dx^2 \wedge dx^3 = -(1/c)$$

We would now like to derive the formula

$$\underline{*}e^\sigma = (-1)^b \frac{h_{\sigma(1)}\dots h_{\sigma(p)}}{h_{\sigma(p+1)}\dots h_{\sigma(n)}} \operatorname{sgn}(\sigma) e^{\tilde\sigma}$$

by a less sophisticated method. To do this we recall the formulas

$$\ell \wedge \underline{*}m = (\ell, m)\Omega_0^*$$

and

$$\Omega_0^* = \sqrt{(-1)^s g}\,\Omega^* = \frac{1}{h_1\dots h_n} e^1 \wedge \dots \wedge e^n .$$

If we now set,

$$\underline{*}\,e^\sigma = \sum_{\tau\in\mathcal{S}_{n,n-p}} \alpha_{\sigma\tau} e^\tau$$

where the $\alpha_{\sigma\tau}$ are to be determined, then we have

$$e^\rho \wedge \underline{*}\,e^\sigma = \sum_{\tau\in\mathcal{S}_{n,n-p}} \alpha_{\sigma\tau} e^\rho \wedge e^\tau \qquad\qquad \rho\in\mathcal{S}_{n,p} .$$

The term $e^\rho \wedge e^\tau$ is 0 unless $\tau = \tilde\rho$. Thus

$$e^\rho \wedge \underline{*}\,e^\sigma = \alpha_{\sigma\tilde\rho} .$$

Next we note by the above equation and Grassmann's theorem

$$\begin{aligned} e^\rho \wedge \underline{*}\,e^\sigma &= (e^\rho, e^\sigma)\Omega_0^* \\ &= \det\begin{pmatrix} (e^{\rho(1)}, e^{\sigma(1)}) & \dots & (e^{\rho(1)}, e^{\sigma(p)}) \\ \dots & \dots & \dots \\ (e^{\rho(p)}, e^{\sigma(1)}) & \dots & (e^{\rho(p)}, e^{\sigma(p)}) \end{pmatrix} \Omega_0^* \end{aligned}$$

and by the othogonality we have the right side equal to 0 unless $\{\rho(1)\dots\rho(p)\}$ and $\{\sigma(1)\dots\sigma(p)\}$ coincide as sets (so that no column will be all zeros). But then, since they are both in $\mathcal{S}_{n,p}$, we must have $\sigma = \rho$. Thus $\alpha_{\sigma\tau}$ is 0 except for the terms $\alpha_{\sigma\tilde\sigma}$. We now know

$$\underline{*}\,e^\sigma = \alpha\, e^{\tilde\sigma}$$

for some $\alpha$ (depending on $\sigma$) and we must determine $\alpha$. Once again

$$\begin{aligned} e^\sigma \wedge \underline{*}\,e^\sigma &= \alpha\, e^\sigma \wedge e^{\tilde\sigma} \\ (e^\sigma, e^\sigma)\Omega_0^* &= \alpha \operatorname{sgn}(\sigma) e^1 \wedge \dots \wedge e^n \end{aligned}$$

$$\det\begin{pmatrix} ((e^{\sigma(1)}, e^{\sigma(1)}) & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & ((e^{\sigma(p)}, e^{\sigma(p)}) \end{pmatrix} = \alpha \operatorname{sgn}(\sigma)\Omega^*$$

Now $(e^{\sigma(i)}, e^{\sigma(i)}) = +h_i$ for $\sigma(i) \leq r = n - s$ and $(e^{\sigma(i)}, e^{\sigma(i)}) = -h_i$ for $\sigma(i) > r = n - s$. Let $b$ be the number of $\sigma(i)$ satisfying the latter condition. We then have

$$\begin{aligned} (-1)^b h^2_{\sigma(1)} \ldots h^2_{\sigma(p)}\, \Omega^*_0 &= \alpha \operatorname{sgn}(\sigma)\, h_1 \ldots h_n\, \Omega^*_0 \\ &= \alpha \operatorname{sgn}(\sigma)\, h_{\sigma(1)} \ldots h_{\sigma(n)}\, \Omega^*_0 \end{aligned}$$

giving

$$\alpha = (-1)^b \frac{h_{\sigma(1)} \ldots h_{\sigma(p)}}{h_{\sigma(p+1)} \ldots h_{\sigma(n)}}$$

which is our previous result.

# Chapter 7

# Regressive Products

## 7.1 Introduction and Example

The Regressive Product is Grassmann's extension of the wedge product. Unlike the wedge product, the extended product has not been much appreciated but it has many uses in vector algebra and projective geometry. In this introductory section we will show how natural a construction it is by looking at the problem of finding the intersection of two planes in $\mathbb{R}^3$. We examine the standard solution of this problem by vector algebra, and then by Grassmann's regressive product, and show the methodological superiority of the latter.

I must mention at this point that the regressive product does have something of a defect; it is somewhat basis dependent. If the basis is changed the regressive product will pick up a multiplier. This is not a problem in projective geometry and there are fixes in metric geometry. We will discuss this matter at greater length in the next section.

Let $\dim(V) = n$. For the wedge product, if $A \in \Lambda^r(V)$ and $B \in \Lambda^s(V)$ and $r+s > n$ then $A \wedge B = 0$. This is convenient for many purposes but also a little dull. Grassmann found and "extension" of the wedge product which gives more interesting information, which we will now present.

To motivate our construction we will look at a familiar problem of vector algebra; finding a vector along the line of intersection of two planes. We will work here in $V = \mathbb{R}^3$ with the standard inner product.

Let the two planes be

$$\begin{array}{lrcl} P_1 : & -2x^1 - 3x^2 + x^3 & = & 0 \\ P_2 : & 3x^1 - x^2 - 2x^3 & = & 0 \end{array}$$

The normal vectors to these planes are

$$n_1 = (-2, -3, 1)^\mathsf{T} \qquad (7.1)$$

$$n_2 = (3, -1, -2)^\mathsf{T} \qquad (7.2)$$

The vector $v$ along the line of intersection must be perpendicular to the two normals $n_1$ and $n_2$ and thus we can take for $V$

$$v = n_1 \times n_2 = (7, -1, 11)^\mathsf{T}$$

From this and the point $(0, 0, 0)$ on both planes we could write down the equation for the line of intersection, if we were interested in this.

Now we want to look at the calculation in a very different way. Recall the basic equations for the $*$-operator:

$$\begin{array}{rcl} \Xi \wedge *A & = & \langle \Xi, A \rangle \, \Omega^* \\ A \wedge *\Xi & = & \langle \Xi, A \rangle \, \Omega \end{array}$$

We will now redo the vector algebra calculation in $\mathbb{R}^3$ but we will NOT use the inner product in $\mathbb{R}^3$ which means we will work in $\mathbb{R}^3$ and $\mathbb{R}^{3*}$. Elements of $\mathbb{R}^3$

will be written as column vectors and elements of $\mathbb{R}^{3*}$ will be written as row vectors. We will use the standard basis for $\mathbb{R}^{3*}$ and its dual basis

$$e^1 = (1,0,0) \qquad e^2 = (0,1,0) \qquad e^3 = (0,0,1)$$

for $\mathbb{R}^{3*}$. So we have

$$\langle e^i, e_j \rangle = \delta^i_j$$

as usual.

Now we want to represent the planes $P_1$ and $P_2$ as wedge products of two vectors in each plane. For $P_1$ we take the two vectors

$$\begin{array}{rcl} v_1 &=& (1,0,2)^\intercal \\ &=& 1e_1 + 0e_2 + 2e_3 \end{array} \qquad \begin{array}{rcl} v_2 &=& (0,1,3)^\intercal \\ &=& 0e_1 + 1e_2 + 3e_3 \end{array}$$

and form

$$v_1 \wedge v_2 = -2e_2 \wedge e_3 - 3e_3 \wedge e_1 + 1e_1 \wedge e_2$$

and then

$$\lambda_1 = *(v_1 \wedge v_2) = -2e^1 - 3e^2 + 1e^3$$

Notice that we have counterfeited normal vector $n_1$ in the dual space. Notice also that a vector $w = (x^1, x^2, x^3)^\intercal$ in in $p_1 \Leftrightarrow$

$$\begin{array}{rcl} \lambda_1(w) &=& \langle \lambda_1, w \rangle \\ &=& \langle -2e^1 - 3e^2 + e^3,\ x^1e_1 + x^2e_2 + x^3e_3 \rangle \\ &=& -2x^1 - 3x^2 + x^3 \\ &=& 0 \end{array}$$

In a similar way, using the vectors $w_1$ and $w^2$ in $P_2$

$$\begin{array}{rcl} w_1 &=& (2,0,3)^\intercal \\ w_2 &=& (0,-1,\frac{1}{2})^\intercal \end{array}$$

we get

$$\begin{array}{rcl} w_1 \wedge w_2 &=& 3e^2 \wedge e^3 - 1e^3 \wedge e^1 - 2e^1 \wedge e^2 \\ \lambda_2 \ = \ *(w_1 \wedge w_2) &=& 3e_1 - e_2 - 2e_3 \end{array}$$

and vector $w = (x^1, x^2, x^3)^\intercal$ is in $P_2 \Leftrightarrow \lambda_2(w) = 0$.

Now let us form

$$\begin{array}{rcl} \lambda_1 \wedge \lambda_2 &=& (-2e^1 - 3e^2 + 1e^3) \wedge (3e^1 - 1e^2 - 2e^3) \\ &=& 7e^2 \wedge e^3 - 1e^3 \wedge e^1 + 11e^1 \wedge e^2 \\ v \ = \ *(\lambda_1 \wedge \lambda_2) &=& 7e^1 - 1e^2 + 11e^3 \\ &=& (7,-1,11)^\intercal \end{array}$$

Now let us verify in this context that $v$ satisfies our desire, namely that it lies in both planes which is equivalent to $\lambda_1(v) = \lambda_2(v) = 0$. We calculate

$$\begin{aligned}\langle\lambda_1, v\rangle\, \Omega^* &= \langle\lambda_1, *(\lambda_1 \wedge \lambda_2\rangle\, \Omega^* \\ &= \lambda_1 \wedge * * (\lambda_1 \wedge \lambda_2) \\ &= \lambda_1 \wedge \lambda_1 \wedge \lambda_2 \\ &= 0\end{aligned}$$

and similarly for $\langle\lambda_2, v\rangle = 0$. (Here we have used that for $A \in \Lambda^r(\mathbb{R}^3)$ we have $* * A = (-1)^{r(3-r)}A) = A$.)

The perceptive reader will notice that we have gotten the same vector $v$ which we originally obtained from the cross product $n_1 \times n_2$. Why is this impressive? Notice that the problem of finding a vector along the line of intersection of two planes has *nothing to do* with an *inner product.* The vector method uses the inner product in an essential way. Using Grassmann techniques we have eliminated the extraneous inner product from the problem, which is methodologically desirable. Careful examination of the calculation will clarify how we were able to counterfeit the activities of the inner product by using the dual space. This is valuable methodology, since if we want to apply Grassmann algebra to, say, projective geometry we definitely do NOT want an inner product hanging around marring the beauty of the landscape.

Grassmann was able to find a modification of the wedge product which gets to the $v$ we found from the products $v_1 \wedge v_2$ and $w_1 \wedge w_2$ representing the planes. He writes

$$v = [v_1 \wedge v_2(w_1 \wedge w_2)]$$

or more elegantly

$$v = [v_1 \wedge v_2.w_1 \wedge w_2]$$

The definition and properties of this new product $[AB]$ are the subject of the following sections.

Since Grassmann was unaware of the dual space as a separate entity, his construction of the above product was based on an inner product methodology. This is contrary to the spirit of modern mathematics (virtually the only part of Grassmann's work where this is the case) and so our development in the following sections will modify Grassmann's construction. It is rather remarkable how little this changes things. The modifications are not very significant.

## 7.2 Definition of the Regressive Product

Based on our experience in the last section we will now define the regressive and combinatorial products. We will work with a vector space $V$ of dimension $n$, its dual space $V^*$ and the duality operators

$$\begin{aligned} *: \Lambda^r(V) &\to \Lambda^{n-r}(V^*) \\ *: \Lambda^r(V^*) &\to \Lambda^{n-r}(V) \end{aligned}$$

Let $\{e_1, \dots, e_n\}$ be a basis of $V$ and $\{e^1, \dots, e^n\}$ the dual basis of $V^*$. Recall that if $\pi \in \mathcal{S}_{n,r}$ then

$$*\, e_\pi = *(e_{\pi(1)} \wedge \dots \wedge e_{\pi(r)}) = \mathrm{sgn}(\pi) e^{\pi(r+1)} \wedge \dots \wedge e^{\pi(n)}$$

and similarly for $*e^\pi$. Recall that these formula are valid for any permutation, not just the ones from $\mathcal{S}_{n,r}$. Finally recall that for $A \in \Lambda^r(V)$

$$*\,*\,A = (-1)^{r(n-r)}\,A$$

Now we define the regressive product:

**Def** Let $A \in \Lambda^r(V)$ and $B \in \Lambda^s(V^*)$ where $0 \le r, s < n$
if $r + s < n$ then $[AB] = A \wedge B$
if $r + s \ge n$ then $*[AB] = *A \wedge *B$
In the latter case we can compute $[AB]$ by using the formula above to compute $*\,*\,[AB]$. Also, since $r, s < n$ we have $n - r,\ n - s > 0$ and $*A \in \Lambda^{n-r}(V^*)$ and $*B \in \Lambda^{n-s}(V^*)$ so

$$\begin{aligned} *[AB] &= *\,A \wedge *B \in \Lambda^{n-r+n-s}(V^*) = \Lambda^{2n-(r+s)}(V^*) \\ [AB] &\in \Lambda^{n-(2n-(r+s))}(V^*) = \Lambda^{(r+s)-n}(V^*) \end{aligned}$$

Since $n \le r + s < 2n$, we have $0 \le (r + s) - n < n$ so $[AB]$ is in the range we like. Summing up
if $r + s < n$ then $[AB] \in \Lambda^{r+s}(V^*)$
if $r + s \ge n$ then $[AB] \in \Lambda^{r+s-n}(V^*)$
The vocabulary is
if $r + s < n$ then the product $[AB]$ is called *progressive*
if $r + s \ge n$ then the product $[AB]$ is called *regressive*

For $A \in \Lambda^r(V^*)$ and $B \in \Lambda^s(V^*)$ exactly the same formulas are used, although the $*$ in these formulas goes the opposite way from the $*$ used above.

Grassmann identified $\Lambda^n(V)$ with $\Lambda^0(V)$ by identifying $a\Omega \in \Lambda^n(V)$ with $a \in \Lambda^0(V)$. This is natural in our circumstances, which we now demonstrate. Let $A = e_1 \wedge \dots \wedge e_r$ and $B = e_{r+1} \wedge \dots \wedge e_n$. First note that $*B$ uses the permutation

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n-r & n-r+1 & n-r+2 & \dots & n \\ r+1 & r+2 & \dots & n & 1 & 2 & \dots & r \end{pmatrix}$$

We see that $\text{sgn}(\pi) = (-1)^{r(n-r)}$ so that $*B = (-1)^{r(n-r)} e_1 \wedge \ldots \wedge e_r$. According to the above rules we must calculate $[AB]$ regressively so that

$$\begin{aligned}
*[AB] &= *A \wedge *B \\
&= (e^{r+1} \wedge \ldots \wedge e^n) \wedge (-1)^{r(n-r)} (e^1 \wedge \ldots \wedge e^r) \\
&= (e^1 \wedge \ldots \wedge e^r) \wedge (e^{r+1} \wedge \ldots \wedge e^n) \\
&= \Omega^* \\
&= *1 \\
[AB] &= 1
\end{aligned}$$

Since $A \wedge B = \Omega$ this is pretty good evidence that $\Omega$ is going to act like 1 in the regressive world. We can also see this if we relax the restrictions on the $A$ and $B$ so that we allow $A = \Omega$ and compute regressively

$$\begin{aligned}
*[\Omega B] &= *\Omega \wedge *B \\
&= 1 \wedge *B \\
&= *B \\
[\Omega B] &= B
\end{aligned}$$

and thus again $\Omega$ acts like 1. However, it seems simpler to me to just restrict $A$ and $B$ to the range $\Lambda^r(V)$ with $0 \leq r \leq n-1$ and leave $\Lambda^n(V)$ completely out of the system, its job being done by $\Lambda^0(V)$.

Later we will have to work with products of more than two elements.

Notice that the formula when $r + s \geq n$, $\quad *[AB] = *A \wedge *B$ has an automorphism like quality to it. It was clever of Grassmann to build this into the theory long before automophisms had been defined, and is another example of Grassmann's profound instincts.

It is always nice to know how a new concept applies to the basis vectors so let us now deal with that. Let $\pi \in \mathcal{S}_{n,r}$ and $\sigma \in \mathcal{S}_{n,s}$:

$$\begin{aligned}
e_\pi &= e_{\pi(1)} \wedge \ldots \wedge e_{\pi(r)} \\
e_\sigma &= e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(s)}
\end{aligned}$$

If $r + s < n$ then $[e_\pi e_\sigma] = e_\pi \wedge e_\sigma$ and there is nothing new.
If $r + s = n$ there are two possibilities.
First

$$e_{\pi(1)}, \ldots, e_{\pi(r)}, e_{\sigma(1)}, \ldots, e_{\sigma(s)}$$

are all distinct. Since both $\pi$ and $\sigma$ are increasing permutations, this can happen only if $\sigma = \tilde{\pi}$. By the rules, we must compute this product regressively:

$$\begin{aligned}
* \, e_\pi &= \text{sgn}(\pi) \, e^{\tilde{\pi}} \\
* \, e_\sigma &= * \, e_{\tilde{\pi}} = \text{sgn}(\tilde{\pi}) \, e^{\pi} \\
* \, [e_\pi e_\sigma] &= * \, e_\pi \wedge * \, e_\sigma = * \, e_\pi \wedge * \, e_{\tilde{\pi}} \\
&= \text{sgn}(\pi) \text{sgn}(\tilde{\pi}) \, e^{\tilde{\pi}} \wedge e^{\pi}
\end{aligned}$$

$$
\begin{aligned}
&= (-1)^{r(n-r)} e^{\tilde{\pi}} \wedge e^{\pi} \\
&= e^{\pi} \wedge e^{\tilde{\pi}} \\
&= \operatorname{sgn}(\pi)\, e^1 \wedge \ldots \wedge e^n \\
&= \operatorname{sgn}(\pi)\, \Omega^* \\
&= \operatorname{sgn}(\pi)\, *1 \\
[e_\pi e_\sigma] &= \operatorname{sgn}(\pi)
\end{aligned}
$$

Second, if $r+s=n$ and

$$e_{\pi(1)}, \ldots, e_{\pi(r)}, e_{\sigma(1)}, \ldots, e_{\sigma(s)}$$

are *not* all distinct, then

$$e^{\tilde{\pi}(1)}, \ldots, e^{\tilde{\pi}(s)}, e^{\tilde{\sigma}(1)}, \ldots, e^{\tilde{\sigma}(r)}$$

also cannot be all distinct, and so

$$
\begin{aligned}
*\,[e_\pi e_\sigma] &= *\, e_\pi \wedge *\, e_\sigma \\
&= *\, e_\pi \wedge *\, e_\sigma \\
&= \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)\, e^{\tilde{\pi}} \wedge e_{\tilde{\sigma}} \\
&= 0 \\
[e_\pi e_\sigma] &= 0
\end{aligned}
$$

We can sum this up conveniently by

$$\text{If } r+s=n \text{ then } [e_\pi e_\sigma] = \operatorname{sgn}(\pi)\, \delta_{\pi\sigma}$$

Notice that

$$e_\pi \wedge e_\sigma = \operatorname{sgn}(\pi) \delta_{\pi\sigma} \Omega$$

So everything is nicely consistent, and we can even write

$$e_\pi \wedge e_\sigma = [e_\pi e_\sigma] \Omega$$

The case $r+s>n$ is best approached in the following way. There must be some repetitions among

$$e_{\pi(1)}, \ldots, e_{\pi(r)}, e_{\sigma(1)}, \ldots, e_{\sigma(s)}$$

Collect the repetitions together and arrange in increasing order to get a $\rho \in \mathcal{S}_{n,r}$, By rearrangement we can write

$$
\begin{aligned}
e_\pi &= \pm\, e_\rho\, e_{\pi_1} \\
e_\sigma &= \pm\, e_\rho\, e_{\sigma_1}
\end{aligned}
$$

where $e_{\pi_1}$ and $e_{\sigma_1}$ have no elements in common. Then compute

$$[(e_\rho\, e_{\pi_1})(e_\rho\, e_{\sigma_1})]$$

In certain cases this can be simplified, but we will handle this when we get to products of three or more elements.

Our next project is to discuss the basis dependence of the regressive product.

## 7.3 Change of Basis

We now wish to discuss what happens to the Regressive Product when the basis is changed. It will turn out that the regressive product accumulates a constant multiplier, which in simple cases is predictable.

Let $\{e_1, \ldots, e_n\}$ be the original basis of $V$ with dual basis $\{e^1, \ldots, e^n\}$ of $V^*$ and let $\{f_1, \ldots, f_n\}$ be the new basis of $V$ and $\{f^1, \ldots, f^n\}$ be the new basis of $V^*$. We want to know the effect of the change of basis on $[AB]$.

We have, with $\Omega = e_1 \wedge \ldots \wedge e_n$,

$$\begin{aligned} f_i &= \alpha_i^j e_j \\ \tilde{\Omega} &= f_1 \wedge \ldots \wedge f_n \\ &= \sum_\pi \operatorname{sgn} \alpha_1^{\pi(1)} \alpha_2^{\pi(2)} \ldots \alpha_n^{\pi(n)} \, e_1 \wedge \ldots \wedge e_n \\ &= \det(\alpha_i^j) \, e_1 \wedge \ldots \wedge e_n \\ &= \det(\alpha_i^j) \, \Omega \end{aligned}$$

The crudest way to find the equation connecting $\tilde{\Omega}^*$ and $\Omega^*$ is as follows. We know

$$f^k = \beta_j^k e^j$$

for some matrix $(\beta_j^k)$ so

$$\begin{aligned} \delta_\ell^k &= \langle f^k, f_\ell \rangle \;=\; \langle \beta_m^k e^m, \alpha_\ell^j e_j \rangle \\ &= \beta_m^k \alpha_\ell^j \langle e^m, e_j \rangle \;=\; \beta_m^k \alpha_\ell^j \delta_j^m \\ &= \beta_j^k \alpha_\ell^j \end{aligned}$$

so

$$\begin{aligned} I &= (\beta_j^k)(\alpha_\ell^j) \\ (\beta_j^k) &= (\alpha_\ell^j)^{-1} \end{aligned}$$

Thus

$$\begin{aligned} \tilde{\Omega}^* &= f^1 \wedge \ldots \wedge f^n \;=\; \det(\beta_j^k) e^1 \wedge \ldots \wedge e^n \\ &= [\det(\alpha_j^i)]^{-1} \Omega^* \end{aligned}$$

Of course, it would have been easier to use $\langle \Omega^*, \Omega \rangle = 1$ and $\langle \tilde{\Omega}^*, \tilde{\Omega} \rangle = 1$ and hence with $\tilde{\Omega}^* = \beta \Omega$ for some $\beta$

$$\begin{aligned} 1 \;=\; \langle \tilde{\Omega}^* \tilde{\Omega} \rangle &= \langle \beta \Omega^*, \det(\alpha_j^i) \Omega \rangle \\ &= \beta \det(\alpha_j^i) \langle \Omega^*, \Omega \rangle \;=\; \beta \det(\alpha_j^i) \\ \beta &= [\det(\alpha_j^i)]^{-1} \\ \tilde{\Omega}^* &= [\det(\alpha_j^i)]^{-1} \Omega^* \end{aligned}$$

Notice the second compuation short circuits a lot of trivia. These computations can also be done in matrix form but we will leave this for the reader to do herself.

Recall now the basic equations for the $*$-operator, which are, with $A \in \Lambda^r(V)$ and $\Xi \in \Lambda^r(V^*)$,

$$\begin{aligned} \Xi \wedge *A &= \langle \Xi, A\rangle \Omega^* \\ A \wedge *\Xi &= \langle \Xi, A\rangle \Omega \end{aligned}$$

The basis enters here in the definitions of $\Omega$ and $\Omega^*$. Let us set $\alpha = \det(\alpha^i_j)$. Then $\tilde{\Omega} = \alpha\,\Omega$ and $\tilde{\Omega}^* = \alpha^{-1}\Omega^*$. We will designate with $\tilde{*}A$ the result of calculationg the $*$-operator using the basis $\{f_1, \ldots, f^n\}$. Then we have

$$\begin{aligned} \Xi \wedge *A &= \langle \Xi, A\rangle\, \Omega^* \\ \Xi \wedge \tilde{*}A &= \langle \Xi, A\rangle\, \tilde{\Omega}^* \\ &= \langle \Xi, A\rangle \alpha^{-1}\Omega^* \end{aligned}$$

Then

$$\Xi \wedge \tilde{*}A = \alpha^{-1}\,\langle \Xi, A\rangle\,\Omega^* = \alpha^{-1}\Xi \wedge A = \Xi \wedge \alpha^{-1} * A$$

Since this is true for all $\Xi \in \Lambda^r(V^*)$ we have

$$\tilde{*}\, A = \alpha^{-1} * A$$

Using exactly similar methods the reader will have no difficulty showing

$$\tilde{*}\, \Xi = \alpha * \Xi$$

for $\Xi \in \Lambda^r(V^*)$.

It is now time to look at a simple example to get a feeling for how this all works and fix the methodology in our minds. Let $V = \mathbb{R}^3$, $\{e_1, \ldots, e_n\}$ the standard basis (written as columns). The new basis will be $\{f_1, \ldots, f_n\}$ where $f_i = 2e_i$. Then $(\alpha^i_j) = 2I$ and $\alpha = \det(\alpha^i_j) = 8$. Note also for the dual basis that $f^i = \frac{1}{2}e^i$. We now have

$$\begin{aligned} *e_1 &= e^2 \wedge e^3 \\ *\tfrac{1}{2}f_1 &= 2f^2 \wedge 2f^3 \\ *f_1 &= 8f^2 \wedge f^3 \end{aligned}$$

whereas

$$\tilde{*}\, f_1 = f^2 \wedge f^3 = \tfrac{1}{8} * f_1$$

just as we predicted from theory. As a second example

$$\begin{aligned} *(e_2 \wedge e_3) &= e^1 \\ *(\tfrac{1}{2}f_2 \wedge \tfrac{1}{2}e_3) &= 2f^1 \\ *(f_2 \wedge e_3) &= 8f^1 \\ \tilde{*}\,(f_2 \wedge f_3) &= f^1 \;=\; \tfrac{1}{8}8f^1 \;=\; \tfrac{1}{8} * (f_2 \wedge f_3) \end{aligned}$$

again as we predicted.

It is now trivial to see what happens to the regressive product under basis change of $A \in \Lambda^r$ and $B \in \Lambda^s$ when $r + s \geq n$ We will label the regressive products $[AB]_*$ and $[AB]_{\tilde{*}}$. Then

$$\begin{aligned}
\tilde{*}[AB]_{\tilde{*}} &= [\tilde{*}A\tilde{*}B]_{\tilde{*}} \\
&= \tilde{*}A \wedge \tilde{*}B \\
&= \frac{1}{\alpha^2} * A \wedge *B \\
\frac{1}{\alpha} * [AB]_{\tilde{*}} &= \frac{1}{\alpha^2} * [AB]_* \\
[AB]_{\tilde{*}} &= \frac{1}{\alpha}[AB]_*
\end{aligned}$$

Naturally this equation is true only when two elements are multiplied and only when the multiplication is regressive.

Now let us again practise to see this at work. Suppose, as before, $V = \mathbb{R}^3$ and $A = e_1 \wedge e_2$, $B = e_1 \wedge e_3$. Then

$$\begin{aligned}
*[(e_1 \wedge e_2)(e_1 \wedge e_3)] &= [*(e_1 \wedge e_2) * (e_1 \wedge e_3)] \\
&= [e^3(-e^2)] \\
&= [e^2e^3] = e^2 \wedge e^3 \\
&= *e_1
\end{aligned}$$

Thus

$$[(e_1 \wedge e_2)(e_1 \wedge e_3)]_* = e_1$$

Similarly

$$[(f_1 \wedge f_2)(f_1 \wedge f_3)]_{\tilde{*}} = f_1$$

Now using $f_i = 2e_i$ and recalling that $\alpha = 8$, we have from the last equation

$$\begin{aligned}
[(2e_1 \wedge 2e_2)(2e_1 \wedge 2e_3)]_{\tilde{*}} &= 2e_1 \\
16[(e_1 \wedge e_2)(e_1 \wedge e_3)]_{\tilde{*}} &= 2e_1 \\
[(e_1 \wedge e_2)(e_1 \wedge e_3)]_{\tilde{*}} &= \tfrac{1}{8}e_1 \\
&= \tfrac{1}{8}[(e_1 \wedge e_2)(e_1 \wedge e_3)]_*
\end{aligned}$$

as predicted.

Thus the regressive product of *two* elements of $\Lambda(V)$ picks up a factor of $\frac{1}{\det(\alpha^i_j)}$ under basis change $f_j = \alpha^i_j e_i$, and similarly the regressive product of two elements of $\Lambda(V*)$ will pick up a factor $\det(\alpha^i_j)$. If more than two elements are involved and both progressive and regressive products show up in the computation it can be difficult to keep track of the factors. We will look at this briefly in a later section.

However, in applications in projective geometry the factors usually do not matter since things are usually defined only up to a constant multiplier. Also, looking back at section one where we found the line of intersection of two planes, the actual length of the vector along the line was of no relevance, so again the constant multiplier is of little importance.

In applications where the length of the vector is important one approach would be to restrict ourselves to base change where $\det(\alpha^i_j) = 1$, that is, to require $(\alpha^i_j) \in \mathrm{SO}(n, \mathbb{R})$. The appearance of $\mathrm{SO}(n, \mathbb{R})$ here suggests that in this approach a metric lurks in the background, which could be defined by declaring the given basis $\{e_1, \ldots, e_n\}$ to be orthonormal.

## 7.4 Dot Notation

This section has almost no content; it is about notation. Because the combinatorial product (which requires both progressive and regressive products in its computation) is highly non associative, it is important to invent a notation that clarifies how the various elements are to be combined in the product. We could, of course, use parentheses, but this turns out to be cumbersome in practise. Thus I have adapted a notation using dots which I learned from W. V. O. Quine long ago. Quine adapted it from Russel and Whitehead, and that is as much history as I know. Similar notations have no doubt been invented many times.

In this section $A, B, \ldots, M$ are elements of $\Lambda(V)$ which are not necessarily monomials. The basic principle which we use is left associativity, which means that

$$[AB \cdots M] = [\cdots [[[AB]C]D] \cdots M]$$

That is, unless otherwise indicated, products are computed by first computing the first 2, then computing with that product and the third element, etc.

If we wish the multipliaction to associate elements differently, we set off groupb of left associated multiplications with dots. Left associativity begins anew with each dot. Thus for example

$$\begin{aligned} [A.BC] &= [A[BC]] \\ [AB.CD] &= [[AB][CD]] \\ [AB.CD.EF] &= [[[AB][CD]][EF]] \end{aligned}$$

Already the reader may note some advantage in the notation. However, we cannot do everything with a single dot. Often multiple dots are necessary; the more dots the stronger the association. There is even a rule for decoding, which we will see later. As examples of where this is necessary, consider the two

$$\begin{array}{ccc} [[[AB][CD]][EF]] & \text{and} & [[AB][[CD][EF]]] \\ & \text{with dots} & \\ [AB.CD.EF] & \text{and} & [AB:CD.EF] \end{array}$$

In the first product, all the dots are of equal weight and so left associativity takes over; first compute $[[AB][CD]] = X$ and then compute $[X[EF]]$. In the second product the double dot indicates that left associativity starts over at the double dot, so first one computes $[[CD][EF]] = Y$ and then computes $[[AB]Y]$.

The principle in general is left associativity starts at an n-dot symbol and continues until the next n-dot symbol, a higher dot symbol, or a final ]. After the new n-dot or higher dot symbol left associativity begins anew. It is most important to realize that the number of dots is equal to the number of left bracket symbols at that point. Using this, one can mechanically fill in the brackets when one sees the dots. Observe this carefully in the following examples.

$$[ABC.EF] = [[[AB]C][EF]]$$

$$\begin{aligned}
[A.BC.EF] &= [[A[BC]][EF]] \\
[A:BC.EF] &= [A[[BC]][EF]]] \\
[A:B.C.EF] &= [A[B[C[EF]]]] \qquad \textbf{WRONG}
\end{aligned}$$

This expression on the left might be psychologically helpful but it violates the rules; the double dot is not necessary as one sees by counting only one left bracket [ at that point. This is a source of confusion and should be avoided. The proper expression is

$$[A.B.C.EF] = [A[B[C[EF]]]] \qquad \textbf{CORRECT}$$

Here are more examples, but in order to increase comprehensibility I will write $[ABC]$ for the correct $[[AB]C]$. Notice my abbreviation conforms to the left associative rule.

$$\begin{aligned}
[ABC.DEF] &= [[ABC][DEF]] \\
[ABC.DEF.GHI] &= [[[ABC][DEF]][GHI]] \\
[ABC:DEF.GHI] &= [[ABC][[DEF]][GHI]] \\
[AB.CD.EF.GH] &= [[[[AB][CD]][EF]][GH]] \\
[AB.CD:EF.GH] &= [[[AB][CD]][[EF]][GH]]]] \quad \text{common} \\
[AB:CD.EF.GH] &= [[AB][[CD][EF][GH]]]
\end{aligned}$$

Note that I am here using the left associativity convention in the expression $[[CD][EF][GH]]$. This is necessary to make the count come out right.

A couple more examples:

$$\begin{aligned}
[AB:CD:EF.GH] &= [[AB][[CD][[EF][GH]]]] \\
[AB:.CD:EF.GH:IJ.KL] &= [[AB][[[CD][[EF][GH]][[IJ][KL]]]]
\end{aligned}$$

## Chapter 8

# Applications to Vector Algebra

## 8.1 Introduction

What is vector algebra? As it is ordinarily taught in the United States, it consists of the elementary theory of the dot product in three dimensional space over $\mathbb{R}$ (largely trivial), a component involving the cross product which is based on the property that $v, w \perp v \times w$ and the equality $||v \times w|| = ||v|| ||w|| \sin\theta$, and a much less trivial part which centers around the vector triple product law $u \times (v \times w) = (u \cdot w)v - (v \times w)u$. In the first section of this chapter we will essentially duplicate this construction, but to maintain interest we will do it in $n$-dimensions. In particular, we will construct an analogy of the cross product that functions in $n$-dimensions. This construction is well known in Russia as the "skew product" but is less well known in the west. This section assumes the standard inner or dot product on $\mathbb{R}^n$.

In the second section we will geometrically interpret (in $\mathbb{R}^n$) the elements of $\Lambda^r(V)$ and in the third we will look at $\Lambda^r(V^*)$. Then we will examine the meaning of the duality operator $* : \Lambda^r(V) \to \Lambda^{n-r}(V^*)$ in the fourth section. Finally in the fifth section we will look at the interpretation of $\underline{*} : \Lambda^r(V) \to \Lambda^{n-r}(V)$ in the presence of a metric.

## 8.2 Elementary $n$-dimensional Vector Algebra

The inner (or dot) product part of $n$-dimensional vector has been sketched out in Chapter 1. It differs not at all from the theory of the 3-dimensional case which we assume known. It allows us to find lengths of vectors and angles between vectors.

We turn now to the analog of the "cross product". We recall that in the 3-dimensional case the cross product is characterized by

1. $v, w \perp v \times w$
2. $||v \times w|| = ||v|| ||w|| \sin\theta$
3. $v, w,$ and $v \times w$ form a right handed system.

To produce an analog in $n$-dimensions, we must discuss the third item first. Handedness is determined by choosing an element $0 \neq \Omega \in \Lambda^n(V)$, and we consider $\Omega_1$ and $\Omega_2$ to determine the same handedness if $\Omega_2 = \alpha\Omega_1$ where $\alpha > 0$. (This makes sense because $\Lambda^n(V)$ is 1-dimensional.) There are two handedness classes represented by $\Omega$ and $-\Omega$. We will refer to the chosen element $\Omega$ (really its handedness class) as the *positive* orientation or *right handed* orientation; the other will be the *negative* or *left handed* orientation. A basis $e_1, e_2, \ldots, e_n$ is then positively oriented or negatively oriented according to whether $e_1 \wedge e_2 \wedge \ldots \wedge e_n = \alpha\Omega$ where $\alpha > 0$ or $\alpha < 0$. If a reader feels that he has a right handed basis $e_1, e_2, \ldots, e_n$ he need only take $\Omega = e_1 \wedge e_2 \wedge \ldots \wedge e_n$ as his chosen element of $\Lambda^n(V)$. There is of course no *mathematical* way to distinguish right from left; only the distinction between the two has meaning.

Let us suppose a choice of basis $\Omega$ has been made. We select an *orthonormal* basis $e_1, e_2, \ldots, e_n$ and form $e_1 \wedge e_2 \wedge \ldots \wedge e_n = \alpha\Omega$ If $\alpha < 0$, we may reverse the sign on any $e_i$ and we will now have a right handed orthonormal basis. We will always consider that this adjustment has been made. We then set $e_1 \wedge e_2 \wedge \ldots \wedge e_n = \Omega_0$ and have $\Omega_0 = \alpha\Omega$ with $\alpha > 0$, so $e_1, e_2, \ldots, e_n$ is a right handed system and $\Omega_0$ is in the right handedness class. We then have

$$(\Omega_0, \Omega_0) = \det(\,(e_i, e_j)\,) = \det \begin{pmatrix} 1 & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & 1 \end{pmatrix} = 1\,.$$

We then have a duality operator $\underline{*} : \Lambda^r(V) \to \Lambda^{n-r}(v)$ which, as we know from section 5.3, satisfies

$$A \wedge \underline{*}B = (A, B)\,\Omega_0 \qquad \text{for } A, B \in \Lambda^r(V)\,.$$

Since in this section we will not be discussing $V^*$ at all, we can simplify the notation by dropping the underline on $\underline{*}$ and writing simply $*$. We are now in a position to define the vector product. **Def** Let $v_1, v_2, \ldots v_{n-1} \in V$. The *cross product* of these vectors, which we shall denote by $\{v_1, v_2, \ldots v_{n-1}\}$, is defined by

$$\{v_1, v_2, \ldots v_{n-1}\} = *(v_1 \wedge v_2 \wedge \ldots \wedge v_{n-1})\,.$$

The perceptive reader will note that the cross product involves $n-1$ vectors. For $n=3$, this is two vectors and gives the familiar $v_1 \times v_2$. However, the $\times$ notation is not well adapted to any dimension except $n=3$.

We remark at this point that it would be possible to define the cross product as a determinant in a way analogous to the case of three dimensions. The theory can then be developed from this, but there is a difficulty in getting the length of the cross product by this route. We will not adopt this method; we will develop the cross product out of the theory of the $*$ operator and get the determinantal formula later.

Notice that the cross product is indeed a vector; $v_1 \wedge v_2 \wedge \ldots \wedge v_{n-1} \in \Lambda^{n-1}$ and $* : \Lambda^{n-1}(V) \to \Lambda^1(V)$. Let us next note that if

$$\sigma = \left( \begin{array}{cccc|c} 1 & 2 & \ldots & n-1 & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n-1) & \sigma(n) \end{array} \right) \in \mathcal{S}_{n,n-1}$$

then

$$\begin{aligned} \{e_{\sigma(1)} \ldots e_{\sigma(n-1)}\} &= *(e_{\sigma(1)} \wedge \ldots \wedge e_{\sigma(n-1)}) \\ &= \operatorname{sgn}(\sigma) e_{\sigma(n)} \end{aligned}$$

Recall that the reverse $\tilde{\sigma}$ of $\sigma$ is given by

$$\tilde{\sigma} = \left( \begin{array}{c|cccc} 1 & 2 & 3 & \ldots & n \\ \sigma(n) & \sigma(1) & \sigma(2) & \ldots & \sigma(n-1) \end{array} \right) \in \mathcal{S}_{n,1}$$

Then $\operatorname{sgn}(\tilde{\sigma}) = (-1)^{\sigma(n)-T_1} = (-1)^{\sigma(n)-1}$ and since $\operatorname{sgn}(\sigma)\operatorname{sgn}(\tilde{\sigma}) = (-1)^{1\cdot(n-1)}$ we have

$$\operatorname{sgn}(\sigma) = (-1)^{n-1-[\sigma(n)-1]} = (-1)^{n-\sigma(n)}$$

so that

$$\{e_{\sigma(1)} \ldots e_{\sigma(n-1)}\} = (-1)^{n-\sigma(n)} e_{\sigma(n)} \,.$$

Now we can do some examples.

$n=2:$ Then the cross product has only one input element $\{e_{\sigma(1)}\}$ and we have

$$\begin{aligned} \{e_1\} &= (-1)^{2-2} e_2 = e_2 \\ \{e_2\} &= (-1)^{2-1} e_1 = -e_1 \end{aligned}$$

Thus for $n=2$ the cross product rotates the vector $\frac{\pi}{2}$ positively.

$n=3:$ The cross product now has two elements and we have

$$\begin{aligned} \{e_1 e_2\} &= (-1)^{3-3} e_3 = e_3 \\ \{e_1 e_3\} &= (-1)^{3-2} e_2 = -e_2 \\ \{e_2 e_3\} &= (-1)^{3-1} e_1 = e_3 \end{aligned}$$

Since $\{e_1 e_3\} = *(e_1 \wedge e_3) = - * (e_3 \wedge e_1) = -\{e_3 e_1\}$, the second equation is often written

$$\{e_3 e_1\} = e_2 \,.$$

These last equations show that $\{e_i e_j\} = e_i \times e_j$ and thus $\{vw\} = v \times w$ as the latter is ordinarily defined, since the value of $\times$ on the basis vectors determines it completely. Hence our cross product is indeed a generalization of the ordinary three dimensional cross product.

To return to the general theory, we first have

$$\begin{aligned} **(v_1 \wedge \ldots \wedge v_{n-1}) &= (-1)^{(n-1)(n-[n-1])} v_1 \wedge \ldots \wedge v_{n-1} \\ &= (-1)^{(n-1)} v_1 \wedge \ldots \wedge v_{n-1} \end{aligned}$$

and then, for $1 \le i \le n-1$,

$$\begin{aligned} (v_i, \{v_1 \ldots v_{n-1}\})\Omega_0 &= (v_i, *(v_1 \wedge \ldots \wedge v_{n-1}))\Omega_0 \\ &= v_i \wedge **(v_1 \wedge \ldots \wedge v_{n-1}) \\ &= (-1)^{n-1} v_i \wedge v_1 \wedge \ldots \wedge v_{n-1} \\ &= 0 \end{aligned}$$

and thus $(v_i, \{v_1 \ldots v_{n-1}\}) = 0$ and

$$v_i \perp \{v_1 \ldots v_{n-1}\} \qquad \text{for } 1 \le i \le n-1\,.$$

This is perhaps the most important property of the cross product.

Next we have

$$\begin{aligned} v_1 \wedge \ldots \wedge v_{n-1} \wedge \{v_1 \ldots v_{n-1}\} &= (v_1 \wedge \ldots \wedge v_{n-1}) \wedge *(v_1 \wedge \ldots \wedge v_{n-1}) \\ &= (v_1 \wedge \ldots \wedge v_{n-1}, v_1 \wedge \ldots \wedge v_{n-1})\,\Omega_0\,. \end{aligned}$$

Since $(v_1 \wedge \ldots \wedge v_{n-1}, v_1 \wedge \ldots \wedge v_{n-1}) \ge 0$, we have either $v_1, \ldots, v_{n-1}, \{v_1 \ldots v_{n-1}\}$ are linearly dependent or $v_1, \ldots, v_{n-1}, \{v_1 \ldots v_{n-1}\}$ is a right handed system.

The first possibility, that $v_1, \ldots, v_{n-1}, \{v_1 \ldots v_{n-1}\}$ is linearly dependent, is equivalent to $v_1, \ldots, v_{n-1}$ being linearly dependent, as we now show. Indeed, since $*$ is an isometry in $\mathbb{R}$ we have

$$\begin{aligned} (\{v_1 \ldots v_{n-1}\}, \{v_1 \ldots v_{n-1}\}) &= (*(v_1 \wedge \ldots \wedge v_{n-1}), *(v_1 \wedge \ldots \wedge v_{n-1})) \\ &= (v_1 \wedge \ldots \wedge v_{n-1}, v_1 \wedge \ldots \wedge v_{n-1}) \\ &= ||v_1 \wedge \ldots \wedge v_{n-1}||^2 \end{aligned}$$

Thus

$$v_1 \wedge \ldots \wedge v_{n-1} \wedge \{v_1 \ldots v_{n-1}\} = ||v_1 \wedge \ldots \wedge v_{n-1}||^2\,\Omega_0$$

and we have

$$\begin{aligned} v_1, \ldots, v_{n-1}, \{v_1 \ldots v_{n-1}\} \text{ lin. ind.} \quad &\text{iff} \quad v_1 \wedge \ldots \wedge v_{n-1} \wedge \{v_1 \ldots v_{n-1}\} = 0 \\ &\text{iff} \quad v_1 \wedge \ldots \wedge v_{n-1} = 0 \\ &\text{iff} \quad v_1, \ldots, v_{n-1} \text{ linearly independent}\,. \end{aligned}$$

Thus the third characteristic property of the cross product remains valid in the generalization.

Another useful property, obtained from Grassmann's theorem, is

$$\begin{aligned}
(\{v_1 \dots v_{n-1}\}, \{w_1 \dots w_{n-1}\}) &= (*(v_1 \wedge \dots \wedge v_{n-1}), *(w_1 \wedge \dots \wedge w_{n-1})\,) \\
&= (v_1 \wedge \dots \wedge v_{n-1}, w_1 \wedge \dots \wedge w_{n-1}) \\
&= \det(\,(v_i, w_j)\,) \qquad (1)
\end{aligned}$$

so that

$$\begin{aligned}
||v_1 \wedge \dots \wedge v_{n-1}||^2 &= (\{v_1 \dots v_{n-1}\}, \{v_1 \dots v_{n-1}\}) \\
&= \det(\,(v_i, v_j)\,)\,. \qquad (2)
\end{aligned}$$

The generalization of the second property is less straightforward. The formula

$$||v \times w|| = ||v||\,||w||\sin\theta$$

does not generalize in a simple way and we cannot discuss it further here. However, the geometric idea behind the formula does generalize, although we will need to be a bit informal about it at this time. The expression above for $||v \times w||$ gives the absolute value of the area of the parallelogram spanned in 3-space by the $v$ and $w$. This we can generalize. Let $v_1, \dots, v_n$ be $n$ linearly independent vectors in $n$-space. They span an $n$-parallelopiped $P$ whose $n$-volume is (by definition) $||v_1 \wedge \dots \wedge v_n||$. (We will discuss this more completely in section 7.5). Let $P_1$ be the "face" of $P$ spanned by $v_1, \dots, v_{n-1}$. Then it is plausible that

$$(n\text{-volume of } P) = (\,(n-1)\text{-volume of } P_1)\,||v_n||\,|\cos\theta|$$

where $\theta$ is the angle between $\{v_1 \dots v_{n-1}\}$ (which is perpendicular to the face) and $v_n$. But now we note that

$$\begin{aligned}
||\{v_1 \dots v_{n-1}\}||\;||v_n||\,\cos\theta\;\Omega_0 &= (v_n, \{v_1 \dots v_{n-1}\})\,\Omega_0 \\
&= (v_n, *(v_1 \wedge \dots \wedge v_{n-1})\,)\,\Omega_0 \\
&= v_n \wedge **\,(v_1 \wedge \dots \wedge v_{n-1})\,\Omega_0 \\
&= v_n \wedge **\,(v_1 \wedge \dots \wedge v_{n-1}) \\
&= (-1)^{(n-1)(n-(n-1))} v_n \wedge (v_1 \wedge \dots \wedge v_{n-1}) \\
&= (-1)^{(n-1)} v_n \wedge v_1 \wedge \dots \wedge v_{n-1} \\
&= v_1 \wedge \dots \wedge v_n
\end{aligned}$$

So, taking norms and recalling that $||\Omega_0|| = 1$, we have

$$\begin{aligned}
||\{v_1 \dots v_{n-1}\}||\;||v_n||\,|\cos\theta| &= ||v_1 \wedge \dots \wedge v_n|| \\
&= (n\text{-volume of } P)\,.
\end{aligned}$$

Comparing this with the equation above, we see that

$$||\{v_1 \dots v_{n-1}\}|| = (\,(n-1)\text{-volume of } P)$$

We can extract from the above proof that

$$(\{v_1 \dots v_{n-1}\}, v_n)\, \Omega_0 = v_1 \wedge \dots \wedge v_n \tag{3}$$

which is a useful formula.

Next we wish to generalize the formula

$$v \times w = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ v^1 & v^2 & v^3 \\ w^1 & w^2 & w^3 \end{vmatrix}.$$

This form is bad for two reasons. First, we would like to write our vector components in column form, and second, the sign comes out wrong in even dimensional spaces. The proper way to write this for purposes of generalization is

$$v \times w = \begin{vmatrix} v^1 & w^1 & \hat{i} \\ v^2 & w^2 & \hat{j} \\ v^3 & w^3 & \hat{k} \end{vmatrix}.$$

We can now generalize, when $v_1, \dots, v_{n-1}$ are written in terms of an orthogonal coordinate system $e_1, \dots, e_n$ as $v_i = v_i^j e_j$, by

$$\{v_1 \dots v_{n-1}\} = \begin{vmatrix} v_1^1 & v_2^1 & \cdots & e_1 \\ \vdots & \vdots & \vdots & \vdots \\ v_1^n & v_2^n & \cdots & e_n \end{vmatrix}. \tag{4}$$

The proof is very easy; let $w$ be any vector in $V$. Then

$$\begin{aligned}
(\{v_1 \dots v_{n-1}\}, w)\, \Omega_0 &= (w, \{v_1 \dots v_{n-1}\})\, \Omega_0 \\
&= (w, *(v_1 \wedge \dots \wedge v_{n-1})\,)\, \Omega_0 \\
&= w \wedge **\,(v_1 \wedge \dots \wedge v_{n-1}) \\
&= (-1)^{(n-1)(n-(n-1))} w \wedge (v_1 \wedge \dots \wedge v_{n-1}) \\
&= (-1)^{(n-1)} w \wedge v_1 \wedge \dots \wedge v_{n-1} \\
&= v_1 \wedge \dots \wedge v_{n-1} \wedge w \\
&= \begin{vmatrix} v_1^1 & v_2^1 & \cdots & v_{n-1}^1 & w^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1^n & v_2^n & \cdots & v_{n-1}^n & w^n \end{vmatrix} \Omega_0 \\
&= \left( (-1)^{n+1} \begin{vmatrix} v_1^2 & \cdots & v_{n-1}^2 \\ \vdots & \vdots & \vdots \\ v_1^n & \cdots & v_{n-1}^n \end{vmatrix} w^1 + \cdots\cdots \right) \Omega_0 \\
&= \left( (-1)^{n+1} \begin{vmatrix} v_1^2 & \cdots & v_{n-1}^2 \\ \vdots & \vdots & \vdots \\ v_1^n & \cdots & v_{n-1}^n \end{vmatrix} e_1 + \cdots\cdots,\ w^1 e_1 + \cdots + w^n e_n \right) \Omega_0 \\
&= \left( \begin{vmatrix} v_1^1 & v_2^1 & \cdots & v_{n-1}^1 & e_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1^n & v_2^n & \cdots & v_{n-1}^n & e_n \end{vmatrix}, w \right) \Omega_0
\end{aligned}$$

from which, since $w$ is arbitrary, the desired equation follows.

Methodologically speaking, it is useful to note that one could take (4) as the *Definition* of $\{v_1 \dots v_{n-1}\}$ of one were only interested in having the cross product and not the entire Grassmann process available. If one goes this route it is not so easy to derive (1) or (2) unless one assumes that $\det(AB) = \det(A)\det(B)$ where $A$ and $B$ are matrices with both scalar and vector entries, like (4), in which case different multiplications are being used for different entries.

To see what interesting and important results may be derived with just the equipment in this section, the reader might profitably consult the splendid book Rosenblum[1].

Equation (1) can be looked at as the source of various special laws resembling the vector triple product law. For example, if $n = 3$ we have by (3)

$$(\{uv\}, z)\,\Omega_0 = u \wedge v \wedge z = v \wedge z \wedge u = (\{vz\}, u)\,\Omega_0$$

so that

$$(\{uv\}, z) = (\{vz\}, u) = (u, \{vz\})$$

or, in more familiar notation

$$u \times v \cdot z = v \times z \cdot u = u \cdot v \times z$$

with the familiar interchange of $\cdot$ and $\times$. Then (1) gives, upon substituting $\{uw\}$ for $u$ and using Grassmann's theorem,

$$\begin{aligned} (\{\{uw\}v\}, z) &= (\{vz\}, \{uw\}) \\ &= \det\begin{pmatrix} (v,u) & (v,w) \\ (z,u) & (z,w) \end{pmatrix} \\ &= (u,v)(w,z) - (w,v)(u,z) \\ &= (\,(u,v)w - (w,v)u,\ z\,) \end{aligned}$$

Since $z$ is arbitrary,

$$\{\{uw\}v\} = (u,v)w - (w,v)u$$

or in more familiar notation

$$(u \times w) \times v = (u,v)w - (w,v)u$$

which is the familiar vector triple product law.

It is clear that the vector triple product law can now be generalized to $n$ dimensions by using similar techniques. First we have

$$\begin{aligned} (\{uv_2 \dots v_{n-1}\}, z)\,\Omega_0 &= u \wedge v_2 \wedge \dots \wedge v_{n-1} \wedge z \\ &= (-1)^{n-1} v_2 \wedge \dots \wedge v_{n-1} \wedge z \wedge u \\ &= (-1)^{n-1}(\{v_2 \dots v_{n-1} z\}, u)\,\Omega_0 \end{aligned}$$

and then

$$
\begin{aligned}
(\{\{u_1\dots u_{n-1}\},v_2\dots v_{n-1}\},z) &= (-1)^{n-1}(\{v_2\dots v_{n-1}z\},\{u_1\dots u_{n-1}\}) \\
&= (-1)^{n-1}\det\begin{pmatrix} (v_2,u_1) & \cdots & (v_2,u_{n-1}) \\ \vdots & \vdots & \vdots \\ (v_{n-1},u_1) & \cdots & (v_{n-1},u_{n-1}) \\ (z,u_1) & \cdots & (z,u_{n-1}) \end{pmatrix} \\
= \det\begin{pmatrix} (v_2,u_2) & \cdots & (v_2,u_{n-1}) \\ \vdots & \vdots & \vdots \\ (v_{n-1},u_2) & \cdots & (v_{n-1},u_{n-1}) \end{pmatrix}(u_1,z) \\
- \det\begin{pmatrix} (v_2,u_1) & (v_2,u_3) & \cdots & (v_2,u_{n-1}) \\ \vdots & \vdots & \vdots & \\ (v_{n-1},u_1) & (v_{n-1},u_3) & \cdots & (v_{n-1},u_{n-1}) \end{pmatrix}(u_2,z) + \cdots
\end{aligned}
$$

where we have expanded the determinant by using the bottom row. This shows, since $z$ was arbitrary, that

$$
\begin{aligned}
\{\{u_1\dots u_{n-1}\},v_2\dots v_{n-1}\} &= \det\begin{pmatrix} (v_2,u_2) & \cdots & (v_2,u_{n-1}) \\ \vdots & \vdots & \vdots \\ (v_{n-1},u_2) & \cdots & (v_{n-1},u_{n-1}) \end{pmatrix}u_1 \\
&- \det\begin{pmatrix} (v_2,u_1) & (v_2,u_3) & \cdots & (v_2,u_{n-1}) \\ \vdots & \vdots & \vdots & \\ (v_{n-1},u_1) & (v_{n-1},u_3) & \cdots & (v_{n-1},u_{n-1}) \end{pmatrix}u_2 + \cdots
\end{aligned}
$$

This is the $n$-dimensional analog of the vector triple product law.

## 8.3 Standard Interpretation of the Grassmann Algebra $\Lambda^r(V)$ of a Real Vector Space $V$

We assume for this section that the field is the real numbers $\mathbb{R}$. Things will be similar over other fields but more difficult to visualize.

We must review some matters before beginning to avoid possible confusion. First, we recall that if $V$ is an $n$-dimensional vector space then $\Lambda^n(V)$ is one dimensional. Hence any two non-zero elements will be real multiples of each other, and are thus to that extent comparable; for example, one may be twice the other.

Now suppose that we have a 2-dimensional subspace $W$ of a 3-dimensional space $V$. Then any two elements of $\Lambda^2(V)$ which are constructed from elements of $W$ are also in $\Lambda^2(W)$ and hence are *comparable*. We are quite used to this in the case of $\Lambda^1(V)$. We interpret elements of $\Lambda^1(V)$ as directed line segments ("vectors") and if we have two collinear vectors (that is, two vectors in the same 1-dimensional subspace $W$,) then one will be a multiple of the other;

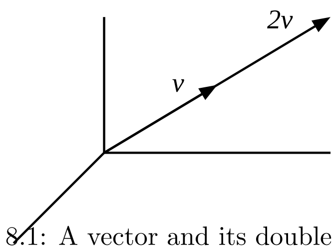



Figure 8.1: A vector and its double

for example in the picture we see $v$ and $2v$. The standard way to describe this situation is to say the $2v$ is *twice as long as* $v$. However, there is *no metric* involved here; the reason we can speak this way is because both $v$ and $2v$ are in the same 1-dimensional space $\Lambda^1(W)$ where $W$ is the 1-dimensional subspace spanned by $v$.

We could push this even further if we wished; we could select a vector $e$ to be the basis of $W$. Then $v = \alpha e$ for any $v \in W$ and we could assign a "length" $|\alpha|$ to the vector $v$, thus counterfeiting a metric on $W$. The flaw in this plan is that there is no sensible relationship between vectors in $V$ which are not collinear, so we do not have an actual metric.

I mention in passing that this method of counterfeiting a metric is often used in projective geometry to define the cross ratio in a situation in which there normally would be no mention of a metric.

Although it is possible to object to the use of language like "2v is twice as long as v," it is very handy to use this language for our purposes, and indeed to avoid it would require creating some clumsy circumlocutions. Therefore we will use the phrase "2v is twice as long as v" in this and the following sections but the reader must remember that no metric has been introduced. More to the point, we will use analogous expressions for higher dimensions, for example "$2A$

has twice the area of $A$," and the reader will understand that what is meant here is that we are working in some $\Lambda^2(W)$ where $W$ is a two dimensional subspace of $V$, or working in some $\Lambda^r(W)$ where $W$ is an $r$-dimensional subspace of $V$,

It is worth noting that should a metric become available then "$w$ is twice as long as $V$" in the sense of this section really does imply that $||w|| = ||2v||$, but the converse is of course false. The situation is similar for higher dimensions.

The situation $w = -2v$ is very common and the description "$w$ is twice as long as $v$ and oppositely oriented" is clumsy, but nothing better is obviously available and we must muddle through with it as best we can.

These preliminaries out of the way, we can now get to the main objective. Vectors in $V$ are represented pictorially as directed line segments in the usual way. Now let $V$ be a three dimensional vector space and $e_1, e_2, e_3$ a fixed basis, which we here draw as orthonormal for artistic convenience.

Let

$$\begin{aligned} v_1 &= e_1 + e_2 \\ v_2 &= e_1 + e_2 + e_3 \end{aligned}$$

so that the picture looks like:

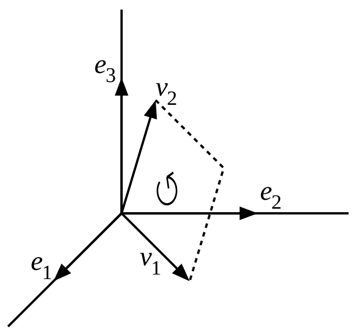



Figure 8.2: Rectangle represented as product

We think of $v_1 \wedge v_2$ as represented by the parallelogram two of whose sides are $v_1$ and $v_2$ as pictured above. The parallelogram has an orientation (first $v_1$ then $v_2$) which is not easy to represent pictorially. Then $v_2 \wedge v_1$ is represented by the same parallelogram which we think of as oriented oppositely, so that $v_1 \wedge v_2 + v_2 \wedge v_1 = 0$. We have drawn the customary circular arrows on the parallelogram to indicate orientation. These are occasionally helpful for visualization. We will often omit the orientation arrows from the figures if they are not relevant to the discussion. This orientation is not easily controlled geometrically, so it is fortunate that the corresponding algebra controls it adequately.

In elementary vector analysis orientation is controlled by using a normal vector $n$ so that $v_1, v_2, n$ form a right handed system. Certain generalizations of this method will function in restricted circumstances but in general normal vectors are not available for Grassmann objects so we will not pursue this method.

Besides the orientation problem, there is a second difficulty in the representation of elements of $\Lambda^2(V)$. For vectors there is essentially just one geometrical object representing the element of $\Lambda^1(V)$, but for $\Lambda^r(V)$ with $r > 1$ this is no longer true. (It is not even true for $\Lambda^1(V)$ unless we insist that all the vector tails are at the origin, which we have tacitly done. Without this requirement a vector is an "equivalence class" of directed line segments where equivalence means same "length" and same direction.) The situation is similar for $\Lambda^r(V)$; even a element $v_1 \wedge v_2$ which is a pure product will have many equivalent representations as a parallelogram, and the same will be true for higher dimensional objects. This makes the pictorial representation less useful than it is for vectors, but still much can be learned from it. We illustrate this now.

Recall that to express $v_1 \wedge v_2$ in terms of $e_1, e_2, e_3$ we use the subdeterminants of the matrix of coefficients

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

of $v_1$ and $v_2$. We have

$$\begin{aligned} v_1 \wedge v_2 &= 0e_1 \wedge e_2 + 1e_1 \wedge e_3 + 1e_2 \wedge e_3 \\ &= (e_1 + e_2) \wedge e_3 \end{aligned}$$

This gives us the picture:

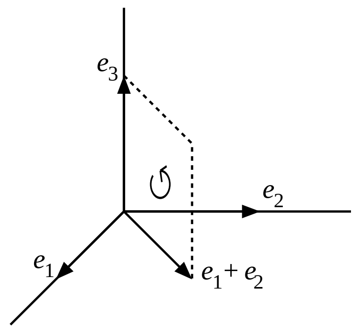



Figure 8.3: Rectangle represented as product

which the reader will note is not the same as the previously pictured parallelogram. Thus the elements $v_1 \wedge v_2$ and $(e_1 + e_2) \wedge e_3$ are equal as elements of the Grassmann algebra but have differing pictorial representations, illustrating the non-uniqueness of the representation. We would like to have some sense of when two parallelograms do indeed represent the same element of the Grassmann algebra, and this is true if we have the following

1. The Parallelograms lie in the same plane
2. The parallelograms have the same area
3. The parallelograms have the same orientation

For number 2. we can deal with the question of area in some naive way, for example dissection. (Recall the discussion of this earlier in this section.) Remember that we are dealing with a pictorial representation which gets progressively less adequate as the dimensions grow, and we must not expect too much of it. For 3., if the dimension of $V$ is 3, one can get a sense of the orientation by curling the fingers of ones right hand in the direction from the first vector to the second, and noting the direction of ones thumb; the orientation is the same if the thumb points out of the plane the same way each time. For higher dimensions orientation is best controlled algebraically.

The reader should now have the sense that any two two pairs of vectors $v_1, v_2$ and $w_1, w_2$ which lie in the same plane, create parallelograms with the same area and have the same orientation will satisfy $v_1 \wedge v_2 = w_1 \wedge w_2$. Here is another example: Let

$$\begin{aligned} v_1 &= (e_1 + e_2) & \qquad w_1 &= e_1 + e_2 - e_3 \\ v_2 &= 2e_3 & \qquad w_2 &= e_1 + e_2 + e_3 \end{aligned}$$

We then have the pictures

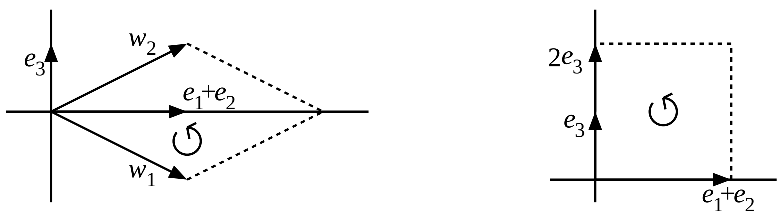



Figure 8.4: Different representations of the same product

The equality of the areas show can be seen by an elementary dissection. Computing algebraically the matrix of coefficients for $w_1$ and $w_2$ is

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}$$

and this gives

$$w_1 \wedge w_2 = 0e_1 \wedge e_2 + 2e_1 \wedge e_3 + 2e_2 \wedge e_3 = (e_1 + e_2) \wedge 2e_3$$

We now note that the above analysis, though taking place in a space $V$ of dimension 3, would also be valid if it all took place in a subspace of dimension 3 in a space $V$ of arbitrary dimension. Products of two vectors would still be represented by parallelograms and everything in the above analysis would remain correct with the exception of the remark concerning the use of a normal vector, which would not be available in the more general case.

Let us illustrate these ideas by considering a sum of the form $v_1 \wedge v_2 + v_3 \wedge v_4$. If $v_1 \wedge v_2$ and $v_3 \wedge v_4$ determine the same plane, the problem is not interesting, so we suppose this is not occurring. In an arbitrary vector space $V$ the most common case occurs when the two planes only intersect at the origin, and this case is also uninteresting because we can do nothing further. However, if the two planes intersect in a line, and this is the case which must occur for distinct planes if $\dim(V) = 3$, there is then an interesting special case. We may then find a vector $w$ contained in the intersection of the planes. We illustrate in the picture with $v_1 = e_1 + e_2 - \frac{1}{2}e_3, \quad v_2 = e_3, \quad v_3 = e_1,$ and $v_4 = e_2$

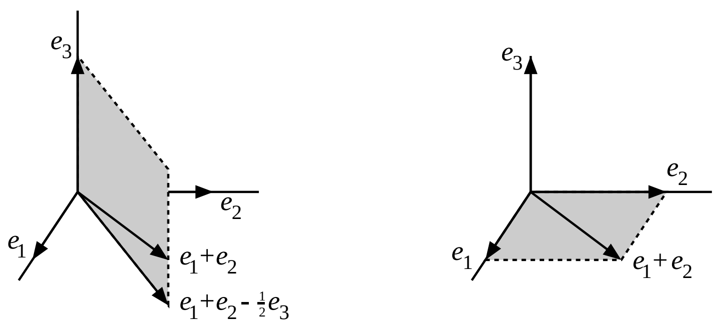



Figure 8.5: Adding two rectangles

In these pictures we see the parallelograms for $v_1 \wedge v_2$ and $v_3 \wedge v_4$ and we choose the vector $w = e_1 + e_2$ in the intersection of the two planes. It is then possible to rewrite the products using this $w$, so that $v_1 \wedge v_2 = w \wedge w_1$ and $v_3 \wedge v_4 = w \wedge w_2$. Of course, $w_1$ and $w_2$ are not uniquely determined. They need only be chosen so as to conserve the area. In our case this can be accomplished by

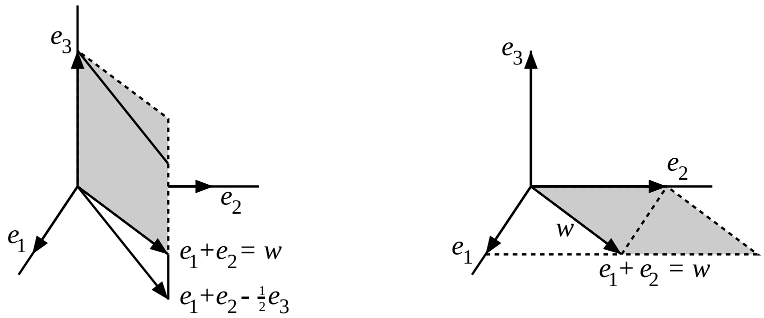



Figure 8.6: Adding two rectangles

taking

$$\begin{aligned} v_1 \wedge v_2 &= (e_1 + e_2 - \tfrac{1}{2}e_3) \wedge e_3 = (e_1 + e_2) \wedge e_3 = w \wedge e_3 \\ v_3 \wedge v_4 &= e_1 \wedge e_2 = (e_1 + e_2) \wedge e_2 = w \wedge e_2\,. \end{aligned}$$

We then have

$$\begin{aligned} v_1 \wedge v_2 + v_3 \wedge v_4 &= w \wedge e_3 + \wedge e_2 \\ &= w \wedge (e_2 + e_3) \end{aligned}$$

which we can picture as

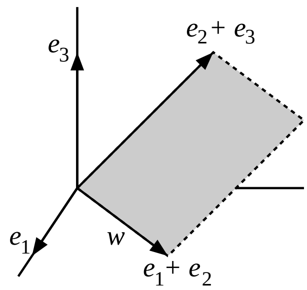



Figure 8.7: Sum of two rectangles

Thus, in the sense of Grassmann algebra, we have added the two parallelograms.

In a similar way, products of three vectors can be represented by parallelopipeds. Higher dimensional objects can be imagined in analogy to those described.

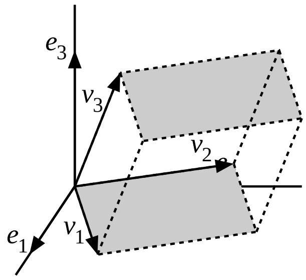



Figure 8.8: Representation of product of 3 vectors

An arbitrary element $A \in \Lambda^r(V)$ will be a sum of products of $r$ vectors which in general will not collapse into a single prodict. In this case our geometric intuition for $A$ remains weak at best. (Those who are familiar with homology theory will sense the similarity to chains in that context.) We best we can do is visualize the individual $r$-parallelograms corresponding to the summands of $A$.

We will now *attempt* to clarify pictorially the orientation on $v_1 \wedge v_2 \wedge v_3$. Below we have shown the parallelopiped representing $v_1 \wedge v_2 \wedge v_3$ and we have drawn in the orientations

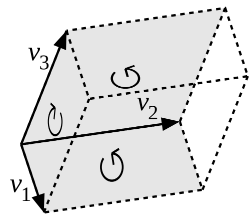



Figure 8.9: Representation showing orientations

and now we show and exploded view:

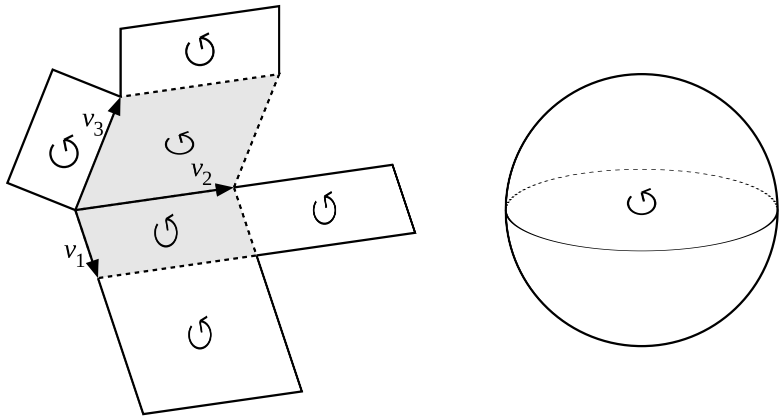



Figure 8.10: Exploded view showing orientations.

Arrows for the faces determined by $v_1 \wedge v_2$, $v_2 \wedge v_3$ and $v_3 \wedge v_1$. These particular orientations are determined by the fact that

$$\begin{aligned}(v_1 \wedge v_2) \wedge v_3 &= (v_2 \wedge v_3) \wedge v_1 \\ &= (v_3 \wedge v_1) \wedge v_2\end{aligned}$$

so that in each of the orders we the vectors form a same handed system as $v_1, v_2, v_3$. We can now order all the sides by following the rule that orientations must cancel when they meet at an edge. Thus in exploded view we have

## 8.4 Geometrical Interpretation of $V^*$

We wish in this section to provide a geometrical interpretation of $V^*$. While certain aspects of this interpretation are incomplete, it nevertheless gives a way of visualizing the elements of $\Lambda^p(V^*)$ and, surprisingly, was discovered over a hundred years ago.

We will begin our discussion with more generality than is really necessary for this section but which will be helpful later in the chapter on manifolds. In this Chapter we will assume the field is always the real numbers **R**. We first consider a *one dimensional vector space* $V$ and then we will interpret an element $f \in \Lambda^1(V^*) = V^*$ to be pictured by a series of vertical lines.

Figure 8.11: An element $f$ of $V^*$

This representation of $f$ we now superimpose on the a horizontal line representing $V$ on which a basis element $e_1$ has been chosen.

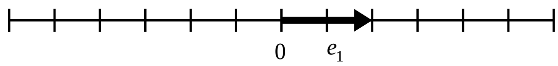



Figure 8.12: $f$ and an element of $e_1$ of $V$

This particular example is set up to represent the element $f = 2e^1 \in V^*$, and the picture is interpreted by counting the number of intervals crossed by $e_1$ to give the value $f(e_1) =< 2e^1, e_1 >= 2$ We further interpret $f$ to be a the description of a (constant) linear density 2. If we wish to illustrate with units, the $e^1$ represents one centimeter and $f$ represents a density of 2 grams/cm. Then $< f, e_1 >$ represents the mass of two grams. Similarly, we represent the situation $v = 2e_1$ as

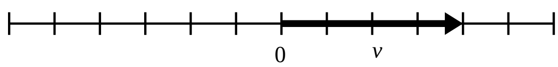



Figure 8.13: An element $f$ of $V^*$

For a physical example, suppose we have a wire of density 4 gm/cm and we wish to find the mass of a wire of length 2 cms. We have $g = 4e^1$ for the density and $v = 2e_1$ for the length, which we diagram in Figure 4. which illustrates that $< g, v >=< 4e^1, 2e_1 >= 8$; that is, the wire has mass 8 gm. If

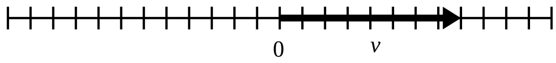



Figure 8.14: $\langle g, v \rangle = 8$

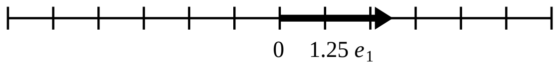



Figure 8.15: $< 2e^1, 1.25e_1 >= 2.5$

we stretch our imaginations a bit we can make this work in general; for example $< 2e^1, 1.25e_1 >= 2.5$ is illustrated in Figure 5.
We could refine the system by putting subdivisions in to make it easier to count fractions of a unit:

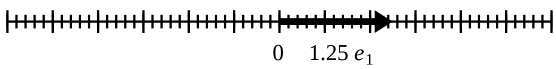



Figure 8.16: $< 2e^1, 1.25e_1 >= 2.5$ in high res

but we are not really interested in refining the system to this degree. We could also change the base unit from cm to inches and correspondingly change the spacing between the lines to represent ounces per inch.

In terms of our previous theory, choosing a "distance" between the vertical lines is choosing a basis element $\Omega^*$ for $\Lambda^1(V)$:

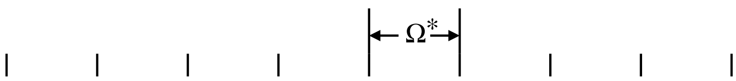



Figure 8.17: Diagram of basis element

and then $\frac{1}{2}\Omega^*$ is represented by

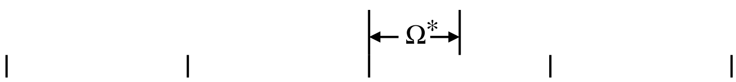



Figure 8.18: Diagram of $\frac{1}{2}$ basis element

and $2\Omega^*$ is represented by

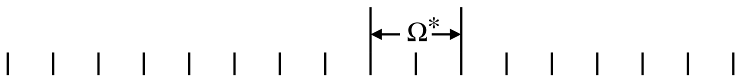



Figure 8.19: Diagram of twice basis element

Associated with the choice of $\Omega^* \in V^*$ is a choice of $\Omega \in V$ which fits precisely between the two vertical lines of $\Omega^*$:

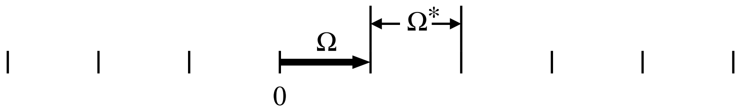



Figure 8.20: Diagram of both basis elements

In our example, since $\Omega^*$ represents 2 gm/cm, $\Omega$ will be a vector half a centimeter long.

We now turn to the case $\dim(V) = 2$. We first interpret elements of $\Lambda^1(V^*)$. We choose a basis in $V$ and a corresponding dual basis $e^1, e^2$ in $V^*$ and ask how

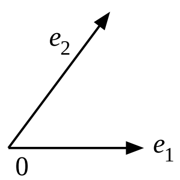



Figure 8.21: Diagram of $e^1$ and $e^2$ from $V$

to interpret $f = \lambda_1 e^1 + \lambda_2 e^2$. We interpret $f$ as a series of equidistant parallel lines generated as follows. The first line is the straight line connecting $\frac{1}{\lambda_1}e_1$ and $\frac{1}{\lambda_2}e_2$. For clarity we will use a specific example; if we take $f = 2e^1 + e^2$ this is the line shown on the left side of Figure 12. The second line is the line parallel to this and through the origin as shown on the right side of Figure 12, and then a whole series of lines is determined by these two as in Figure 13.

The series of parallel lines represents $f = 2e^1 + e^2$. The extra mark has a significance which we will explain shortly.

We now want to relate the this picture of $f \in \Lambda^1(V)$ to the value of f on a vector v, that is we want to use the picture with the vector included to determine the value of $\langle f, v\rangle$. To do this we define a *stripe* as the region between two of the parallel lines. We then draw the vector into the picture and count the number of stripes through which the vector passes. This is then the value of $\langle f, v\rangle$.

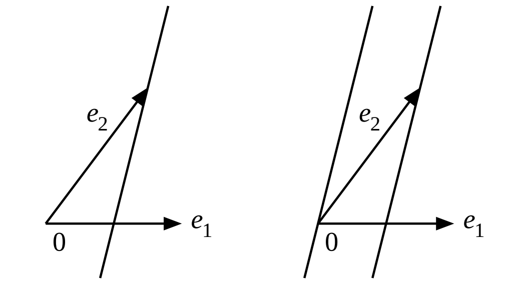



Figure 8.22: Partial Diagrams of $2e^1 + e^2$

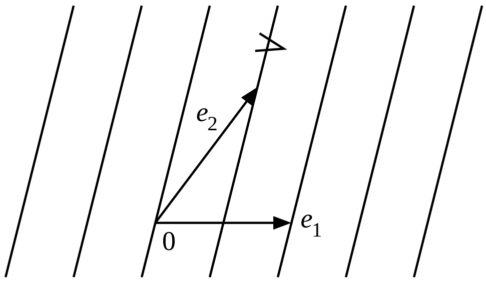



Figure 8.23: Complete Diagram of $2e^1 + e^2$

We take as examples the vectors $v = e_1 + e_2$, $u = -e_1 + e_2$, and $w = 1.5e_1 + .5e_2$. First we concentrate on $v$, where we see in the picture that $v$ crosses three stripes:

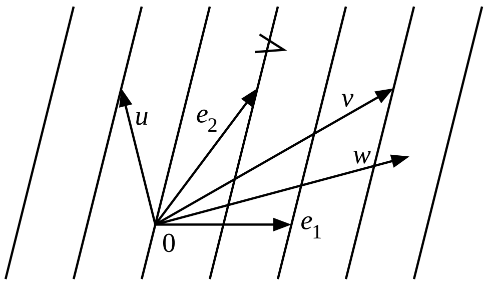



Figure 8.24: Value of $2e^1 + e^2$ on vectors $u, v, w \in V$

and this tells us that $\langle f, v\rangle = 3$. If we calculate this we find

$$\langle f, v\rangle = \langle 2e^1 + e^2, e_1 + e_2\rangle = 2 + 0 + 0 + 1 = 3\,.$$

in agreement with the result from the picture.

To justify the pictorial representation, we note that it is clear, first that $\langle f, e_1\rangle = 2$ and $\langle f, e_2\rangle = 1$ and, second, that the method of computing $\langle f, v\rangle$ is linear in $v$. Thus our method correctly represents $f = 2e^1 + e^2$.

For the other vectors $u$ and $w$, pictorially we have

$$\langle f, w\rangle = 3\tfrac{1}{2} \qquad \langle f, u\rangle = -1$$

and computationally

$$\begin{aligned}\langle f, w\rangle &= \langle f = 2e^1 + e^2, 1.5e_1 + .5e_2\rangle = 2\cdot 1.5 + 1\cdot .5 = 3.5\\ \langle f, u\rangle &= \langle f = 2e^1 + e^2, -1e_1 + 1e_2\rangle = 2\cdot(-1) + 1\cdot(1) = -1\end{aligned}$$

Notice that in the visual computation of $\langle f, u\rangle = -1$ we obtained the result $-1$. This is where the small sign ($>$) plays its role. This sign indicates the *positive* direction of the series of parallel lines. Vectors like $v$ and $w$ which cut the lines in this direction (rightwards in this case) count the strips positively. If the vector cuts the strips in the opposite direction, like $u$, then the result is negative. The element $-f = -2e^1 - e^2 \in V$ would have the same series of parallel lines, but the positive direction of the lines would be reversed.
The diagram for $-f$ then looks like

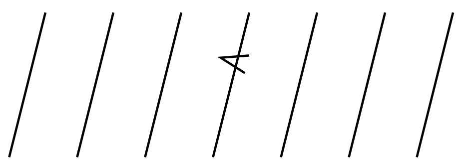


Figure 8.25: Diagram for $-f = -2e^1 - e^2$

Finally, we note the patterns of parallel lines that correspond to the basis vectors $e^1$ and $e^2$ of $V^*$. They are Before going on, it might be helpful to provide a

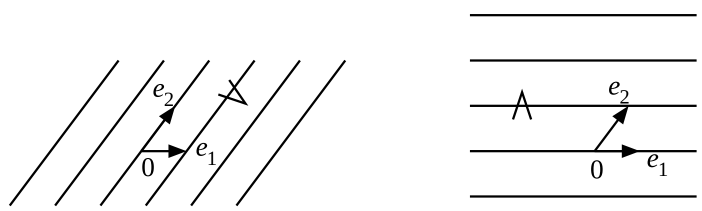



Figure 8.26: Diagrams for $e^1$ and $e^2$

physical example to of how the parallel lines of a linear functional could be interpreted. There are many ways to do this but perhaps temperature is the most familiar. We could interpret $\langle f, v\rangle$ as the *temperature* at the head end of the vector $v$, or, more generally and since everything is linear, we could also say that $\langle f, v\rangle$ gives the difference in temperature between the head and tail ends of the vector. The parallel lines of $f$ then can be interpreted as the isotherms of the temperature (lines of constant temperature) and the symbol ($>$) indicates the direction of increasing temperature.

Another possible interpretation would be electrostatic potential where the parallel lines are equipotential lines.

Next we consider $\Lambda^2(V)$. Let us consider a product $f \wedge g$ where $f$ is the same as in the previous part of the section and $g = e^1 + 3e^2$; we diagram both below. Hence we might reasonably represent $f \wedge g$ by putting both sets of lines

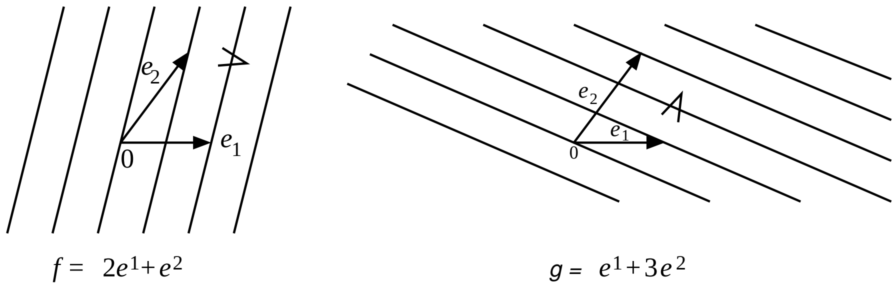



Figure 8.27: Diagrams for $f$ and $g$

on the same graph: The crossed lines representing $f \wedge g$ can now be used in the

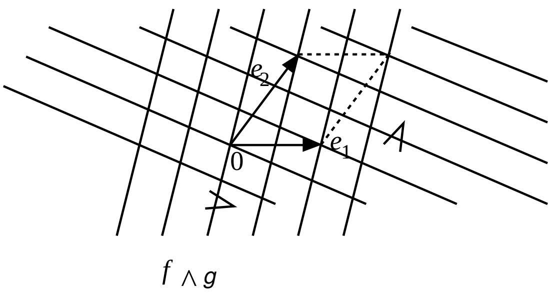



Figure 8.28: Diagram for $f \wedge g$

following manner. If we have a parallelogram representing $v \wedge w$ in the plane, the value of $\langle f \wedge g, v \wedge w\rangle$ can be found by counting the number of areas formed by the crossed lines which are *inside* the the parallelogram of $v \wedge w$. For example, we have shown the parallelogram for $e_1 \wedge e_2$. If we count the number of areas formed by the crossed lines and inside the parallelogram of $e_1 \wedge e_2$ we see there are approximately 5 such units of area. Hence $\langle f \wedge g, e_1 \wedge e_2\rangle$ is approximately

5. If we explicitly compute this we have

$$\begin{aligned}\langle f \wedge g, e_1 \wedge e_2 \rangle &= \langle (e^1 + 2e^2) \wedge (3e^1 + e^2), e_1 \wedge e_2 \rangle \\ &= \langle 5e^1 \wedge e^2, e_1 \wedge e_2 \rangle \\ &= 5\end{aligned}$$

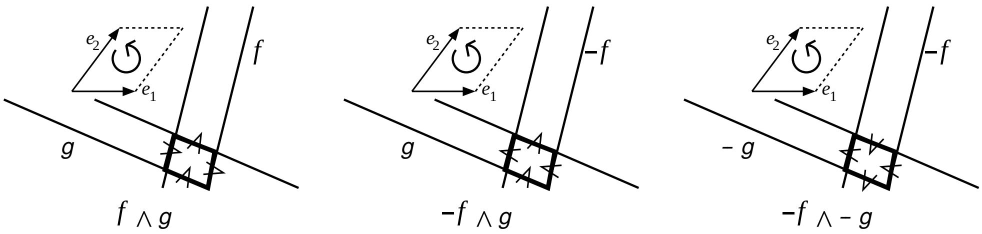



Figure 8.29: Diagram for $f \wedge g$

There is still the question of orientation to deal with. We will just deal with this cursorially, leaving the interested reader to fill in the details. If we want to discover whether a given system of crossed lines corresponding to $f \wedge g$ is positively or negatively oriented, we start with one of the lines of $f$ and go around a parallelopided in the direction considered positive in comparison to the basic choice of $\Omega = e_1 \wedge e_2$ (which is counterclockwise in the above picture). As we go round the parallelopided in the proper direction (starting with an $f$ line) the arrows will point IN or OUT. Form the sequence of INs and OUTs from the arrows:

| | | | | |
|---|---|---|---|---|
| left picture: | $I$ | $O$ | $O$ | $I$ |
| middle picture: | $I$ | $I$ | $O$ | $O$ |
| right picture: | $O$ | $I$ | $I$ | $O$ |

The rule is this. If the first two in the sequence are the same, then $f \wedge g$ is negatively oriented and $f \wedge g = \alpha\Omega^*$ with $\alpha < 0$. If the first two in the sequence are different, then $f \wedge g$ is positively oriented and $f \wedge g = \alpha\Omega^*$ with $\alpha > 0$.

We now turn the case of $\dim(V) = 3$. In this case a linear functional $f \in \Lambda^1(V^*) = V^*$ is represented by a system of equally spaced planes determined analogously to the previous case; if

$$f = \lambda_1 e^1 + \lambda_3 e^3 + \lambda_3 e^3$$

then one of the planes generating the system goes through the three points $(1/\lambda_1, 0, 0)$, $(0, 1/\lambda_2, 0)$, $(0, 0, 1/\lambda_3)$ and the second is parallel to it and through the origin. The others are parallel and spaced at identical intervals. Figure 20 shows the planes for $2e^1 + 2e^2 + e^3$. It would be possible to use some sort of little cone to indicate the increasing direction for the planes in analogy to the two dimensional examples, but we will not pursue this option.

Just as in two space, it is possible to determine $\langle f, v \rangle$ by counting the number of layers between the planes a vector passes through, as illustrated in

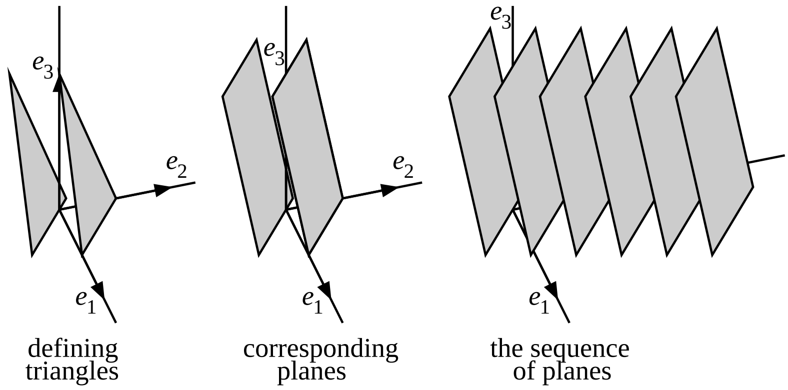



Figure 8.30: Diagram for $f = 2e^1 + 2e^3 + e^3$

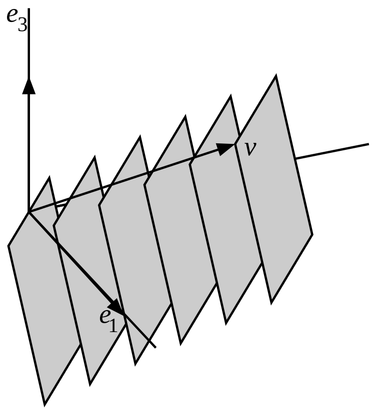



Figure 8.31: Calculating $\langle f, v\rangle$

Figure 21 for $v = e_1 + e_2 + e_3$ where we have shifted the planes downward to allow viewing of the vector. As is visible, the vector $v$ cuts through 5 layers giving the value $\langle f, v\rangle = 5$. This is also the result of calculation:

$$\begin{aligned} \langle f, v\rangle &= \langle 2e^1 + 2e^2 + e^3, e_1 + e_2 + e_3 \\ &= 2 + 2 + 1 = 5 \end{aligned}$$

We now turn to representing elements of $\Lambda^2(V^*)$. For example if we take the elements $f = 2e^1 + 2e^2 + e^3$ and $g = 5e^3$ of $\Lambda^1(V^*)$ illustrated in Figure 22 and form their product $f \wedge g$ we get which shows how an element of $\Lambda^2(V^*)$ determines a system of rectangular tubes. There is an interesting historical circumstance here; In the 1860's James Clerk Maxwell was working on the mathematical description of electric force fields which he referred to as "tubes

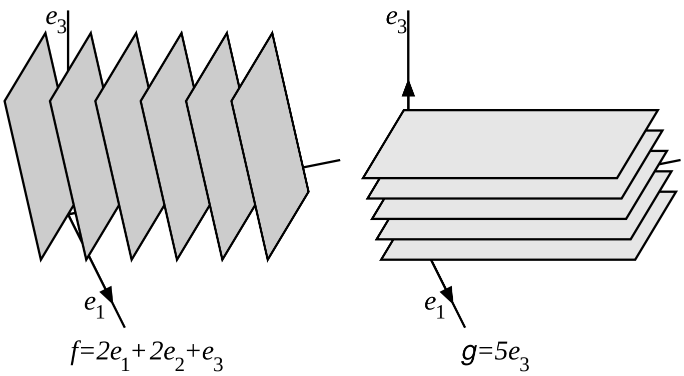



Figure 8.32: Diagrams for $f$ and $g$

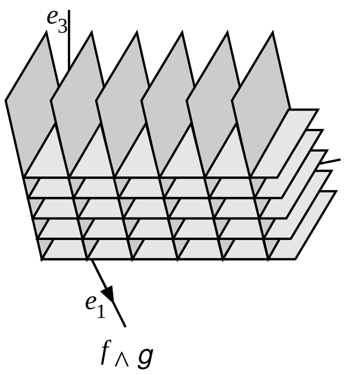



Figure 8.33: Diagram for $f \wedge g$

of force". He then described surfaces as cutting these tubes, and the "flux" was a measure of how many tubes the surface cut. The tubes to which he was referring are exactly the tubes visible in the above figure. Thus in some sense none of this material is really new. We will return to this in Chapter 9.

It is important to realize that the representation of $f$ as a picture is highly non-unique. Also, there are issues of orientation to worry about, which are a little difficult to represent pictorially. For example, in the above example

$$\begin{aligned} f \wedge g &= (2e^1 + 2e^2 + e^3) \wedge 5e^3 \\ &= 10e^1 \wedge e^3 + 10e^2 \wedge e^3 + 5e^3 \wedge e^3 \\ &= 10e^1 \wedge e^3 + 10e^2 \wedge e^3 \\ &= (e^1 + e^2) \wedge 10e^3 \end{aligned}$$

which in some ways would make for a simpler picture.

We now wish to examine the interaction $\langle f, v\rangle$ pictorially. To do this we will make the situation a bit simpler. We will take $v = e_2 \wedge (-e_1 + e_3) \in \Lambda^2(V)$ and $f = 2e^2 \wedge 3e^3 \in \Lambda^2(V^*)$. These are illustrated in Figure 24. To illustrate

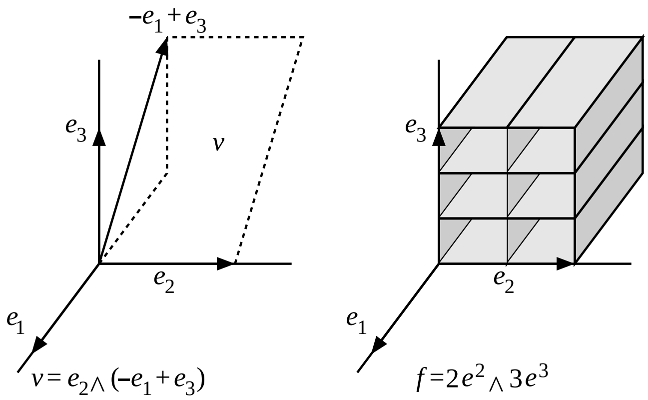



Figure 8.34: Diagram for $\langle f, v\rangle$

the interaction of the two objects, we move $v$ forward and extend the tubes so that the just reach $v$. The value of $\langle f, v\rangle$ is then equal to the number of tubes

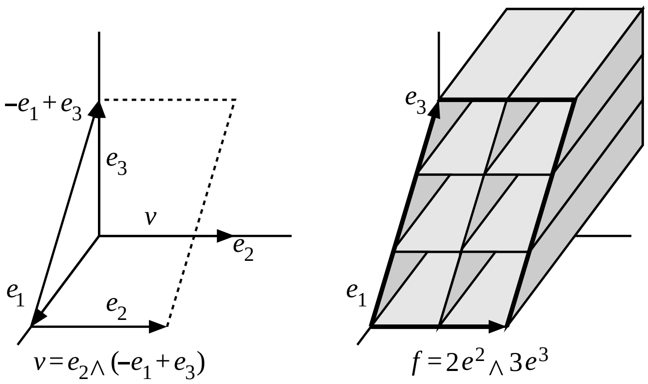



Figure 8.35: Better Diagram for $\langle f, v\rangle$

cut, which we see is 6. Hence $\langle f, v\rangle = 6$ which we now verify by computation:

$$\begin{aligned}\langle f, v\rangle &= \langle 2e^2 \wedge 3e^3, e_2 \wedge (-e_1 + e_3)\rangle \\ &= -6\langle e^2 \wedge e^3, e_2 \wedge e_1\rangle + 6\langle e^2 \wedge e^3, e_2 \wedge e_3\rangle \\ &= -6\delta^{23}_{21} + 6\delta^{23}_{23} = -6 \cdot 0 + 6 \cdot 1 = 6\,.\end{aligned}$$

Our last case is concerns $\Lambda^3(V^*)$ which we illustrate with $f = 2e^1 \wedge 2e^2 \wedge 3e^3$. The picture is from which we see that the value of $f$ on the unit cube $e_1 \wedge e_2 \wedge e_3$

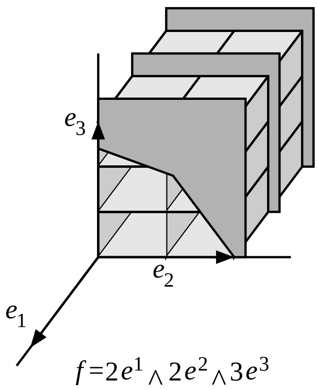



Figure 8.36: Diagram for a 3-form

is 12 since the unit cube would enclose 12 of the cells of $f$, which is then easily verified by computation

$$\begin{aligned} \langle f, e_1 \wedge e_2 \wedge e_3 \rangle &= \langle 2e^1 \wedge 2e^2 \wedge 3e^3, e_1 \wedge e_2 \wedge e_3 \rangle \\ &= 12 \langle e^1 \wedge e^2 \wedge e^3, e_1 \wedge e_2 \wedge e_3 \rangle \\ &= 12 \langle \Omega^*, \Omega \rangle \\ &= 12\,. \end{aligned}$$

This concludes our attempt to render the meaning of $\Lambda^r(V^*)$ visually. While a lot of the subtleties are lost in the higher dimensional examples, I nevertheless feel that the geometric entities we have illustrated are are quite helpful in understanding the action of $\Lambda^r(V^*)$ on $\Lambda^r(V)$.

## 8.5 Geometrical Interpretation of $* : \Lambda^r(V^*) \to \Lambda^{n-r}(V)$ and $* : \Lambda^r(V) \to \Lambda^{n-r}(V^*)$

*Figures need relabeling and maybe some repositioning*

Recall from section 4.5 that the $*$-operators were defined by

$$
\begin{aligned}
* : \Lambda^r(V) \to \Lambda^{n-r}(V^*) \qquad & \ell \wedge *v = \langle \ell, v \rangle \Omega^* \\
* : \Lambda^r(V^*) \to \Lambda^{n-r}(V) \qquad & u \wedge *\ell = \langle \ell, u \rangle \Omega
\end{aligned}
$$

for $\ell \in \Lambda^r(V^*)$, $\quad u, v \in \Lambda^r(V)$ and $\Omega,\ \Omega^*$ satisfying the condition $\langle \Omega^*, \Omega \rangle = 1$. These basic equations make it possible to easily interpret the geometric meaning. We begin with $\dim(V) = 2$ and $\ell \in \Lambda^1(V^*) = V^*$ so that $*\ell \in \Lambda^1(V)$. The basic equation $u \wedge *\ell = \langle \ell, u \rangle \Omega$ then becomes a condition on the area of a parallelogram formed from $u$ and $*\ell$. We represent $\ell$ by the usual sequence of parallel lines

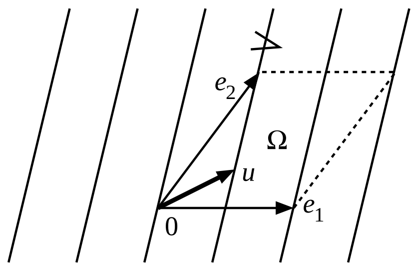



Figure 8.37: $u \wedge *\ell = \langle \ell, u \rangle \Omega$

and also illustrate $\Omega = e_1 \wedge e_2$ and a vector $u$. See figure 1. Since $u$ crosses exactly one stripe, we have $\langle \ell, u \rangle = 1$, and this is true for any vector whose arrow end lies on the line through (the arrow end of) $e_2$. We think of $u$ as representing any such vector. Finding a representation for $*\ell$ then reduces to finding a vector $*\ell$ so that the parallelogram formed from $u$ and $*\ell$ (including orientation) will have the same area as the parallelogram formed from $\Omega = e_1 \wedge e_2$. This is simple; it is only required to point $*\ell$ along the line of $\ell$ through the origin, and adjust its length so that it will have the required area. This is illustrated below in Figure 2. The exact position of $u$ is of no consequence, since sliding its arrow end up and down the line does not change the area of the parallelogram.

It is interesting to use the picture as a guide to deriving the formula for $*\ell$. If $\ell = \lambda_1 e^1 + \lambda_2 e^2$ then the series of parallel lines is generated by the line through (the arrow ends of) $\frac{1}{\lambda_1} e^1$ and $\frac{1}{\lambda_2} e^2$. Thus $*\ell = \alpha(\frac{1}{\lambda_2} e_2 - \frac{1}{\lambda_1} e_1)$. We may determine $\alpha$ if $\lambda_2 \neq 0$ by multiplication by $e_1$:

$$\langle \ell, e_1 \rangle \Omega \quad = \quad e_1 \wedge *\ell$$

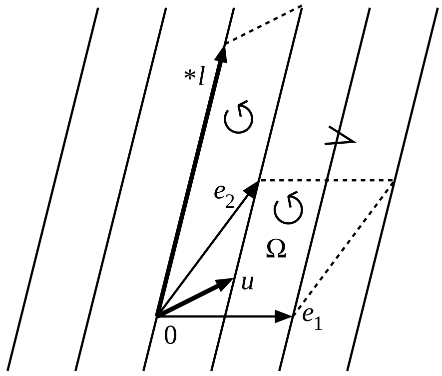



Figure 8.38: $u \wedge *\ell = \langle \ell, u \rangle \Omega$

$$\begin{aligned}\langle \lambda_1 e^1 + \lambda_2 e^2, e_1 \rangle \Omega &= e_1 \wedge \alpha(\frac{1}{\lambda_2} e_2 - \frac{1}{\lambda_1} e_1) \\ \lambda_1 \Omega &= \frac{\alpha}{\lambda_2} e_1 \wedge e_2 = \frac{\alpha}{\lambda_2} \Omega\end{aligned}$$

so that $\alpha = \lambda_1 \lambda_2$ and thus

$$\begin{aligned}*\ell = *(\lambda_1 e^1 + \lambda_2 e^2) &= \alpha(\frac{1}{\lambda_2} e_2 - \frac{1}{\lambda_1} e_1) \\ &= \lambda_1 \lambda_2 (\frac{1}{\lambda_2} e_2 - \frac{1}{\lambda_1} e_1) \\ &= \lambda_1 e_2 - \lambda_2 e_1\end{aligned}$$

which is the formula given by the algebraic methods of section 4.5.

We now turn our attention to the opposite problem of geometrically representing $*v$ as a series of parallel lines. This is a little harder because our intuition for families of parallel lines is not so well developed, so we use a little algebraic help. Let

$$v = \alpha^1 e_1 + \alpha^2 e_2$$

so

$$\begin{aligned}*v &= \alpha^1 * e_1 + \alpha^2 * e_2 \\ &= \alpha^1 e^2 - \alpha^2 e^2 .\end{aligned}$$

Our standard method for drawing the parallel lines of a linear functional $\ell = \beta_1 e^1 + \beta_2 e^2$ is to generate them by the line through the arrow ends of $\frac{1}{\beta_1} e_1$ and $\frac{1}{\beta_2} e_2$ and a parallel line through the origin. In our case these are $-\frac{1}{\alpha^2} e_1$ and $\frac{1}{\alpha^1} e_2$. A vector going from the first to the second is then

$$\frac{1}{\alpha^1} e_2 - (-\frac{1}{\alpha^2} e_1) = \frac{1}{\alpha^1 \alpha^2} (\alpha^2 e_2 + \alpha^1 e_1)$$

$$= \frac{1}{\alpha^1\alpha^2}v\,.$$

Thus the lines of $*v$ are parallel to $v$, and the above recipe determines the spacing. All that remains is the orientation, which we determine as follows. Select $w$ so that $v\wedge w = \alpha\Omega = \alpha\, e_1\wedge e_2$ with $\alpha>0$, (that is, $v$ and $w$ are same handed as $e_1$ and $e_2$). Then we have

$$\begin{aligned} w\wedge *\ell &= \langle \ell, w\rangle\Omega \qquad \text{for } \ell\in\Lambda^1(V^*)\\ w\wedge **v &= \langle *v, w\rangle\Omega\\ w\wedge(-1)^{1\cdot(n-1)}v &= \\ -w\wedge v &= \\ v\wedge w &= \\ \alpha\Omega &= \end{aligned}$$

This shows that $\langle *v, w\rangle = \alpha > 0$ so that $w$ is in the increasing direction for the parallel lines, and this determines the orientation, which we see in the following picture:

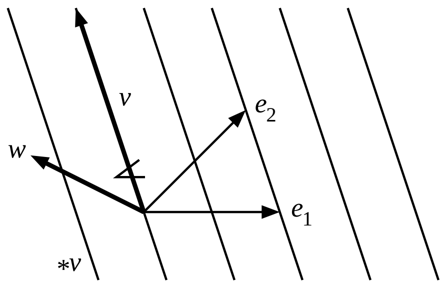



Figure 8.39: $u\wedge *\ell = \langle \ell, u\rangle\Omega$

The (<) always points out the side of $v$ that $e_2$ lies on with respect to $e_1$. (The above picture above uses $v=-2e_1+2e_2,\quad *v=-2e^1-2e^2$ for an example.)

Now we present a more complete picture to illustrate the equation

$$\ell\wedge *v = \langle \ell, v\rangle\Omega^*\,.$$

The following picture includes the linear functional $\ell = e^1$ and shows the parallelogram representing $\Omega^* = e^1\wedge e^2$ and also the parallelogram representing $\ell\wedge *v$. The width of the
strips between the lines representing $*v$ has been set so that the parallelogram representing $\ell\wedge *v$ has exactly half the area of that representing $\Omega^*$. This means that that $\ell\wedge *v = \pm 2\Omega^*$. To see why this is so, recall how an element in $\Lambda^2(V^*)$ is evaluated on an element of $\Lambda^2(V)$; one counts the number of cells representing the the first which lie in the area representing the second. Hence, if one cell is

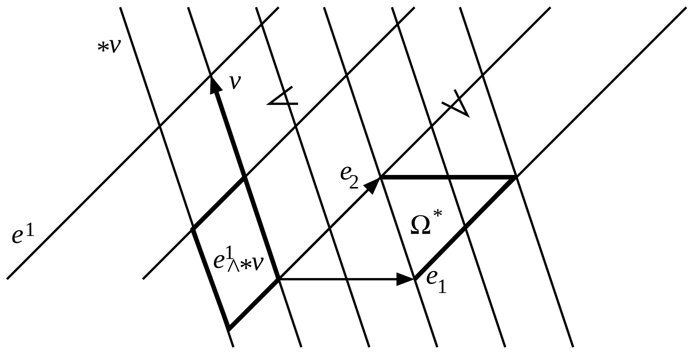



Figure 8.40: $u \wedge *\ell = \langle \ell, u \rangle \Omega$

twice as big as another, it's value on a given area is half as big. In fact, as we now calculate, $\ell \wedge *v = -2\Omega^*$ but we will leave it to the reader to sort out the orientation for himself. The calculation runs

$$\begin{aligned}\ell \wedge *v = e^1 \wedge (-2e^1 - 2e^2) &= -2(e^1 \wedge e^1 + e^1 \wedge e^2)\\ &= -2(0 + e^1 \wedge e^2)\\ &= -2\,\Omega^* = \langle \ell, v\rangle \Omega^*\,,\end{aligned}$$

since

$$\begin{aligned}\langle \ell, v\rangle &= \langle e^1, -2e_1 + 2e_2\rangle &(8.1)\\ &= -2 &(8.2)\end{aligned}$$

so that indeed $\ell \wedge *v = \langle \ell, v\rangle \Omega^*$. The last calculation is represented geometrically by the parallelogram representing $e^1 \wedge *v$ having a side half as long as $v$.

Next we turn out attention to the case of $\dim(V) = 3$ and begin with an $\ell \in \Lambda^1(V^*)$. We will draw the axes orthonormal and use the example $\ell = e^1 + e^2 + e^3$ so that the situation will be easy to visualize. We represent $\ell$ as usual as a sequence of planes,

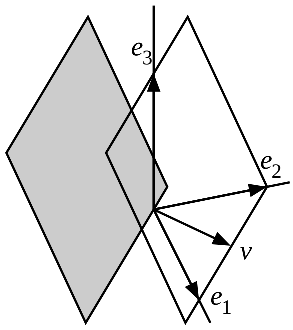



Figure 8.41: $u \wedge *\ell = \langle \ell, u\rangle \Omega$

two of which are shown above. From our experience in the $\dim(V) = 2$ case we suspect that $*\ell$ will be an area which we can represent (in many ways) in the shaded plane and which, when multiplied with a vector $v$ going from the origin to the transparent plane (for which $\langle \ell, v\rangle = 1$) will form the element $\Omega = e_1 \wedge e_2 \wedge e_3$. In contrast to the two dimensional case, it is not totally obvious how to do this geometrically, although it could be so done. We will proceed algebraically and illustrate the results geometrically.

We find $*\ell = *(e^1 + e^2 + e^3)$ in the usual way:

$$\begin{aligned}*\ell &= *(e^1 + e^2 + e^3)\\ &= e_2 \wedge e_3 - e_1 \wedge e_3 + e_1 \wedge e_2\,.\end{aligned}$$

It is not obvious that this element actually lies in the shaded plane, and to see this we must make some modifications.

$$
\begin{aligned}
*\ell &= e_2 \wedge e_3 - e_1 \wedge e_3 + e_1 \wedge e_2 \\
&= (e_2 - e_1) \wedge e_3 + e_1 \wedge e_2 \\
&= (e_2 - e_1) \wedge e_3 - (e_2 - e_1) \wedge e_2 \\
&= (e_2 - e_1) \wedge (e_3 - e_2)
\end{aligned}
$$

In this form it is obvious that the representing parallelogram does indeed lie in the shaded plane.

## Chapter 9

# Applications to Projective Geometry

## 9.1 Introduction

A very fruitful place for Grassmann's ideas is projective geometry. Grassmann's original presentation of his theory was projective, but we have taken a vector space approach as this is a more familiar environment for most mathematicians now, and improves the flexibility of application. The two approaches are entirely equivalent being merely different interpretations of the same symbols. However, analytic projective geometry is extremely important, lying as it does in the foothills of Algebraic Geometry, and every mathematician should be acquainted with its basic ideas.

Our approach to projective geometry is slightly non-standard because our concern is the application of Grassmann Algebra. However, the differences from a standard treatment might be considered enrichment by a fair minded person. Also, while the treatment might look original it is only a trivial elaboration of the ideas of Grassmann and those of standard projection geometry; no serious innovation is claimed.

There is no ideal way to approach this subject. Every presentation has serious drawbacks. The approach I take keeps related ideas together and is not overwhelming in bringing on too many ideas at once, but is seriously inefficient. My attempt at an efficient treatment mixed up the concepts so thoroughly I feared nothing would be clear, but I apologize to experienced persons who will be annoyed at having to go through essentially the same material twice.

## 9.2 Standard Projective Geometry

In this section I will introuduce homogeneous coordinates and simple analytica projective geomery in two dimensions. In a later section I will relate this to the set of lines through the origina in 3 space, but we will not use that interpretation in this section. Higher dimestions are treated analogously. We need a nodding familiarity with the standard treatment in order to see how it relates to Grassmann's treatment, and it is enormously interesting in its own right.

Two dimensional Euclidean Geometry, essentially $\mathbb{R} \times \mathbb{R}$ with the standard inner product, is seriously asymmetric with respect to points and lines.

two points determine a line

two *non parallel* lines determine a point

Desargue (1591-1661) sugggested the addition of "points at infinity" so that parallel lines would have a point at which to meet. Each family of parallel lines determines a unique point at $\infty$ through which all lines of the family go. The technical terminology is "All lines of the family are *incident* with the point at $\infty$." This adds a "circle" of points at $\infty$ except that antipodal points on the circle are identified. This is a very crude picture of the *Project Plane* $\mathbb{P}^2(\mathbb{R})$ but it can be helpful in getting an initial feeling for the subject. Since the "circle" is at infinite "distance", its radius is infinite and so its curvature is 0. Thus it can also be thought of as the "line at $\infty$," and this way of thinking is also useful. Note that since two points determine a line, two points a $\infty$ determine the line at $\infty$. All of this can be analytically confirmed once we bring on the equipment. After adding in the points at $\infty$ and the one line at $\infty$ we have much more symmetry between lines and points;

two points determine a line

two lines determine a point

The first, but not the only, casualty of adding in the new points and line are the loss of the inner product; the new structure has no simple geometrically understandable metric or inner product; we must forego the idea of any two points having a finite distance between them.

Another gain is that projections from the new structure $\mathbb{P}^2(\mathbb{R})$ to itself are extraordinarly well behaved and are handled well by the familar GL(3, $\mathbb{R}$) (the 3 is not a misprint) as we will later show.

However, there is another serious loss. Two dimensional Euclidean space $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is the direct product of $\mathbb{R}$ with itself, but this is not true for $\mathbb{P}^2(\mathbb{R})$. $\mathbb{P}^1(\mathbb{R})$ is formed by adding a single point at $\infty$ to the real line, and this point, residing as it does at either "end" of $\mathbb{R}$, connects the two "ends" so that $\mathbb{P}^1(\mathbb{R})$ is topologically a circle. Thus, topologically, $\mathbb{P}^1(\mathbb{R}) \times \mathbb{P}^1(\mathbb{R})$ is the direct product of two circles and thus a torus (hollow donut, like an inner tube). It is intuitively clear that this is not the same as $\mathbb{P}^2(\mathbb{R})$:

$$\mathbb{P}^2(\mathbb{R}) \neq \mathbb{P}^1(\mathbb{R}) \times \mathbb{P}^1(\mathbb{R})$$

So there is both gain and loss in moving from $\mathbb{R}^2$ to $\mathbb{P}^2(\mathbb{R})$; it would be unwise to consider either of them "better" or "more natural;" mathematics needs both. In the context of algebraic geometry $\mathbb{P}^2(\mathbb{R})$ is the more natural basic object, but even there tangent spaces to varieties are usually thought of as $\mathbb{R}^2$.

Now a critical person might suggest that the whole idea of points at $\infty$ is nonsense; in fact many students feel this way when introduced to this subject. However, most of the students can be beaten into sullen acquiescence when the points at $\infty$ are given coordinates. If objects have coordinates which are numbers, it is harder to claim they make no sense. Plücker and Grassmann both gave coordinates to the points at $\infty$ around 1842, in slightly different ways. The standard treatment which I will present here is closer to Plücker's approach; we will look at Grassmann's approach later, and give a generalized treatment of which both are special cases.

In the standard treatment a point in $\mathbb{P}^2(\mathbb{R})$ is given by a triple $[x^0, x^1, x^2]$ of real numbers. (There is an analogous theory using complex numbers which is used in algebraic geometry; there is no difficulty in switching from $\mathbb{P}^2(\mathbb{R})$ to $\mathbb{P}^2(\mathbb{C})$.) The tricky point is that any *non-zero multiple* of the original triple $[\alpha x^0, \alpha x^1, \alpha x^2]$ refers to the *same point.* Thus the triples

$$[2,4,6], \quad [1,2,3], \quad [\tfrac{1}{2},1,\tfrac{3}{2}], \quad [100,200,300]$$

all refer to the same point. And what point is that? This question asks for the *affine* coordinates $(X^1, X^2)$ of the point. We find these in the $x^1$ and $x^2$ positions when the triple is adjusted to that $x^0 = 1$, or put another way, if we divide $[x^0, x^1, x^2]$ by $x^0$ we get $[1, \frac{x^1}{x^0}, \frac{x^2}{x^0}]$ so that we have

$$X^1 = \frac{x^1}{x^0} \qquad X^2 = \frac{x^2}{x^0}$$

In our example the triple that begins with 1 is $[1,2,3]$ and thus the point referred to is $(2,3)$, which can also be obtained from the triple $[100,200,300]$ by dividing by 100. Thus we can go back and forth between the triples and the affine coordinates with ease, the point $(-7,5)$ having triple $[1,-7,5]$.

It will be useful to have a notation for triples that refer to the same point. We wil use $\sim$ for this so we write

$$\begin{aligned}[\alpha x^0, \alpha x^1, \alpha x^2] &\sim [x^0, x^1, x^2] \\ [2,4,6] &\sim [100,200,300]\end{aligned}$$

But, you say, what if $x^0 = 0$? To investigate this question and get some practise with the notation, let us look at the typical line in affine coordinates

$$4X^1 + 7X^2 - 3 = 0$$

We parametrize the line by setting $X^1 = \frac{7t}{4}$ and thus $X^2 = -t + \frac{3}{7}$ and thus the affine coordinates are $(\frac{7t}{4}, -t + \frac{3}{7}$. Hence the projective coordinates are

$$[1, \frac{7t}{4}, -t + \frac{3}{7}] \sim [\frac{1}{t}, \frac{7}{4}, -1 + \frac{3}{7t}] \sim [\frac{4}{t}, 7, -4 + \frac{12}{7t}]$$

Now it is clear from the forms of the affine and projective coordinates that as $t \to \infty$

$$\begin{array}{rcl} (\frac{7t}{4}, -t + \frac{3}{7}) & \longrightarrow & \infty \\ [\frac{4}{t}, 7, -4 + \frac{12}{7t}] & \longrightarrow & [0, 7, -4] \end{array}$$

and thus a reasonable interpretation is that $[0, 7, -4]$ are the projective coordinates of the point at $\infty$ on the line $4X^1 + 7X^2 - 3 = 0$. In a similar way

$$[0, \mu_2, -\mu_1] \text{ is the point at } \infty \text{ on } \mu_0 + \mu_1 X^1 + \mu_2 X^2 = 0$$

The family of lines parallel to $\mu_0 + \mu_1 X^1 + \mu_2 X^2 = 0$ all have the form $\tilde{\mu}_0 + \mu_1 X^1 + \mu_2 X^2 = 0$ and we see that they all have the same point $[0, \mu_2, -\mu_1]$ at $\infty$.

We also see that $x^0 = 0$ is characteristic of points at $\infty$. Let us look at this a little more closely. Using the equations $X^1 = \frac{x^1}{x^0}, X^2 = \frac{x^2}{x^0}$ we can get the *projective* equation of a line easily by multiplying through by $x^0$;

$$\begin{array}{rcl} \mu_0 + \mu_1 X^1 + \mu_2 X^2 & = & 0 \\ \mu_0 + \mu_1 \frac{x^1}{x^0} + \mu_2 \frac{x^2}{x^0} & = & 0 \\ \mu_0 x^0 + \mu_1 x^1 + \mu_2 x^2 & = & 0 \end{array}$$

The last is the projective equation of the line, and it has the advantage that we can now include the line at $\infty$ by taking $\mu_0 = 1$ and $\mu_1 = \mu_2 = 0$. Thus $x_0 = 0$ is the projective equation of the line at $\infty$.

Looked at another way, the intersection of the line $\mu_0 x^0 + \mu_1 x^1 + \mu_2 x^2 = 0$ and the line at $\infty$ is formed by setting $x^0 = 0$ leaving us with $\mu_1 x^1 + \mu_2 x^2 = 0$ with obvious solutions $[0, \mu_2, -\mu_1]$ and $[0, -\mu_2, \mu_1]$ which of course give the same point, being non-zero multiples of one another.

The last equation suggests an extemely important idea. We note that $[\mu_0, \mu_1, \mu_2]$ and $[x^0, x^1, x^2]$ enter into the equation $\mu_0 x^0 + \mu_1 x^1 + \mu_2 x^2 = 0$ in a symmetrical manner. This suggests that the points (represented by $[x^0, x^1, x^2]$) and the lines (represented by $[\mu_0, \mu_1, \mu_2]$, with multiplication by a non-zero constant giving the same line) enter into projective geometry ina symmetrical manner, which we have already indicated by the symmetry of

two points determine a line

two lines determine a point

This is the famous *Principle of Duality* which is fundamental in projective geometry and can be systematized by

**A theorem remains valid if the words "point" and "line" are replaced by the words "line" and "point".**

as in the above example.

Above we found the intersection of the line $\mu_0 x^0 + \mu_1 x^1 + \mu_2 x^2 = 0$ with the line at $\infty$. It is natural, then, to want a formula for the intersection of any two distinct lines

$$\begin{aligned} \mu_0 x^0 + \mu_1 x^1 + \mu_2 x^2 &= 0 \\ \nu_0 x^0 + \nu_1 x^1 + \nu_2 x^2 &= 0 \end{aligned}$$

This is very easy once we set things up correctly.

Let us set $e_1 = [1,0,0]$, $e_2 = [0,1,0]$, $e_3 = [0,0,1]$ so that $\{e_1, e_2, e_3\}$ is a basis of the vector space of all $[a,b,c]$. Let $\{e^1, e^2, e^3\}$ be the dual basis, $\Omega = e_1 \wedge e_2 \wedge e_3$ and $\Omega^* = e^1 \wedge e^2 \wedge e^3$. Now recall the general formula

$$\mu \wedge *z = \langle \mu, z \rangle \Omega^* \qquad \mu \in \Lambda^r(V^*) \quad z \in \Lambda^r(V)$$

Now we take $\mu \in \Lambda^1(V*)$, $\mu = \mu_i e^i$ where $\mu_i$ are from the triple $\{\mu_i\}$ in $\mu_i x^i = 0$. Forming $\nu$ in the same way from the triple $\{\nu_i\}$ in $\nu_i x^i = 0$ we set $x = *(\mu \wedge \nu) \in \Lambda^1(V)$. Then, using the general formula and for some $k \in \mathbb{Z}$ whose value we don't need,

$$\begin{aligned} \mu \wedge **(\mu \wedge \nu) &= \langle \mu, *(\mu \wedge \nu) \rangle \Omega^* \\ (-1)^k \mu \wedge \mu \wedge \nu &= \langle \mu, *(\mu \wedge \nu) \rangle \Omega^* \\ 0 &= \langle \mu, x \rangle \Omega^* \end{aligned}$$

Thus we have $\mu_i x^i = \langle \mu, x \rangle = 0$ for $x = *(\mu \wedge \nu)$. In a precisely similar way we have $\nu_i x^i = 0$. Thus $x = *(\mu \wedge \nu)$ is the point of intersection of the two lines.

From these results we can produce useful computational formulas for the intersections of two lies. Let the two lines, in projective form be

$$\mu = \mu_i e^i \qquad \nu = \nu_j e^j$$

and then the point of intersection $x = *(\mu \wedge \nu)$ will be calcuated thus:

$$\begin{aligned} \mu \wedge \nu &= (\mu_1\nu_2 - \mu_2\nu_1) e^1 \wedge e^2 + (\mu_2\nu_0 - \mu_0\nu_2) e^2 \wedge e^0 + (\mu_0\nu_1 - \mu_1\nu_0) e^0 \wedge e^1 \\ *(\mu \wedge \nu) &= (\mu_1\nu_2 - \mu_2\nu_1) e_0 + (\mu_2\nu_0 - \mu_0\nu_2) e_1 + (\mu_0\nu_1 - \mu_1\nu_0) e_2 \\ x &= (\mu_1\nu_2 - \mu_2\nu_1) e_0 + (\mu_2\nu_0 - \mu_0\nu_2) e_1 + (\mu_0\nu_1 - \mu_1\nu_0) e_2 \\ [x^0, x^1, x^2] &= [\mu_1\nu_2 - \mu_2\nu_1, \mu_2\nu_0 - \mu_0\nu_2, \mu_0\nu_1 - \mu_1\nu_0] \\ &= \left[ \begin{vmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{vmatrix}, \begin{vmatrix} \mu_2 & \mu_0 \\ \nu_2 & \nu_0 \end{vmatrix}, \begin{vmatrix} \mu_0 & \mu_1 \\ \nu_0 & \nu_1 \end{vmatrix} \right] \end{aligned}$$

If $\begin{vmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{vmatrix} \neq 0$ (meaning the lines are not parallel) then the affine coordinates of the intersection will be

$$\left( \frac{\begin{vmatrix} \mu_2 & \mu_0 \\ \nu_2 & \nu_0 \end{vmatrix}}{\begin{vmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{vmatrix}}, \frac{\begin{vmatrix} \mu_0 & \mu_1 \\ \nu_0 & \nu_1 \end{vmatrix}}{\begin{vmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{vmatrix}} \right)$$

This is just Cramer's rule in a thin disguise.

**Example** 1.

$$\begin{aligned} X^1 + X^2 &= 4 \qquad & -4x^0 + x^1 + x^2 &= 0 \\ X^1 - 2X^2 &= -5 \qquad & 5x^0 + x^1 - 2x^2 &= 0 \end{aligned}$$

Then using the above formulas

$$\begin{aligned} [x^0, x^1, x^2] &= \left[ \begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix}, \begin{vmatrix} 1 & -4 \\ -2 & 5 \end{vmatrix}, \begin{vmatrix} -4 & 1 \\ 5 & 1 \end{vmatrix} \right] \\ &= [-3, -3, -9] \sim [1, 1, 3] \\ (X^1, X^2) &= (1, 3) \end{aligned}$$

**Example** 2.

$$\begin{aligned} X^1 - 3X^2 &= 5 \qquad & -5x^0 + x^1 - 3x^2 &= 0 \\ -2X^1 + 6X^2 &= 11 \qquad & -11x^0 - 2x^1 + 6x^2 &= 0 \end{aligned}$$

Then using the above formulas

$$\begin{aligned} [x^0, x^1, x^2] &= \left[ \begin{vmatrix} 1 & -3 \\ -2 & 6 \end{vmatrix}, \begin{vmatrix} -3 & -5 \\ 6 & -11 \end{vmatrix}, \begin{vmatrix} -5 & 1 \\ -11 & -2 \end{vmatrix} \right] \\ &= [0, 63, 21] \sim [0, 3, 1] \end{aligned}$$

which is the point at $\infty$ on both lines. Notice how the projective techniques handle this case with no strain.

Now it is time to make use of the principle of duality. We want a formula for the line through two points which is dual to the question about the point on two lines which we just handled. Due to the algebraic duality between the two problems, the answers must come out in a similar way. Let the two points be $[x^0, x^1, x^2]$ and $[y^0, y^1, y^2]$. Then the solution must be

$$\ell = *(x \wedge y)$$

with coordinates

$$[\ell_0, \ell_1, \ell_2] = \left[ \begin{vmatrix} x^1 & y^1 \\ x^2 & y^2 \end{vmatrix}, \begin{vmatrix} x^2 & y^2 \\ x^0 & y^0 \end{vmatrix}, \begin{vmatrix} x^0 & y^0 \\ x^1 & y^1 \end{vmatrix} \right]$$

The really careful reader will note that in the determiant I have transposed the rows and columns, in keeping with the idea that points have vertical vectors and lines horizontal vectors, a convention which we have been ignoring in this section.

**Example** 1. line through $(1, 3)$ and $(5, -2)$. The projective corrdinate are $[1, 1, 3]$ and $[1, 5, -2]$. Then the line $\ell = *(x \wedge y)$ through these two points will have coordinates (in the standard dual basis)

$$\begin{aligned}
\left[\begin{vmatrix} 1 & 5 \\ 3 & -2 \end{vmatrix} \begin{vmatrix} 3 & -2 \\ 1 & 1 \end{vmatrix} \begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix}\right] &= [-17, 5, 4] \\
-17x^0 + 5x^1 + 4x^2 &= 0 \\
5X^1 + 4X^2 &= 17
\end{aligned}$$

**Example** 2. line through points with projective corrdinates $[0, 4, 3]$ and $[1, 5, -2]$. Then the line $\ell = *(x \wedge y)$ through these two points will have coordinates (in the standard dual basis)

$$\begin{aligned}
\left[\begin{vmatrix} 4 & 5 \\ 3 & -2 \end{vmatrix} \begin{vmatrix} 3 & -2 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 \\ 4 & 5 \end{vmatrix}\right] &= [-23, 3, -4] \\
-23x^0 + 3x^1 - 4x^2 &= 0 \\
3X^1 - 4X^2 &= 23
\end{aligned}$$

We end this section with a demonstration of the power of projective geometry to give stubstance to our intuition and instincts. Who has not looked at a hyperbola with it's asymptotes and not felt that the asymptotes are tangents to the hyperbola at $\infty$? [1] Projective Geometry allows us to justify this intuitive impression. We will will now look at this in detail. Let us use

$$\begin{aligned}
\frac{{X^1}^2}{2^2} - \frac{{X^2}^2}{3^2} &= 1 \qquad \text{Hyperbola} \\
\frac{X^1}{2} - \frac{X^2}{3} &= 0 \qquad \text{Asymptote}
\end{aligned}$$

Finding the projective forms of the curve by using $X^1 = \frac{x^1}{x^0}$ and $X^2 = \frac{x^2}{x^0}$ gives

$$\begin{aligned}
-{x^0}^2 + \frac{{x^1}^2}{2^2} - \frac{{x^2}^2}{3^2} &= 0 \\
\frac{x^1}{2} - \frac{x^2}{3} &= 0
\end{aligned}$$

The point at $\infty$ on the asymptote is

$$[x^0, x^1, x^2] \sim [0, \tfrac{1}{3}, \tfrac{1}{2}] \sim [0, 2, 3]$$

and *this point is also on the hyperbola* as we can verify by inputting the coordinates into the equation. Thus $[0, 2, 3]$ is a point of intersection of the hyperbola and the asymptote, as we suspected. Notice two things in passing. First, the intersection point is both in the "upper right" and "lower left" of the graph; they are the same point. (The intersection with the second asymptote is $[0, 2, -3]$

[1] The question is rhetorical; in actual fact not all that many people have this impression.

and is found in the "lower right" and "upper left.") In general a straight line in projective space will hit a conic twice; the fact that an asymptote has only one point of intersections suggests the hyperbola and asymptote are tangent. This can be verified by using a projective transformation. Imagine the the hyperbola and asymptotes drawn on a transparent sheet and use a flashlight to project the sheet on a wall. If one is clever one can arrange it so that the point $[0, 2, 3]$ projects into the point $[1, 1, 0]$. Analytically this can be done by the linear transformation[2]

$$\begin{aligned} x^0 &= \tilde{x}^2 \\ x^1 &= \tilde{x}^1 \\ x^2 &= \tilde{x}^0 \end{aligned}$$

Substituting these into the previous projective equations for the hyperbola and asymptote one gets

$$\begin{aligned} -\tilde{x}^{2\,2} + \frac{\tilde{x}^{1\,2}}{2^2} - \frac{\tilde{x}^{0\,2}}{3^2} &= 0 \\ \frac{\tilde{x}^1}{2} - \frac{\tilde{x}^0}{3} &= 0 \end{aligned}$$

Now we devide by $\tilde{x}^{0\,2}$ and $\tilde{x}^0$ to get the affine equations

$$\begin{aligned} -\tilde{X}^{2\,2} + \frac{\tilde{X}^{1\,2}}{2^2} &= \frac{1}{3^2} \\ \frac{\tilde{X}^1}{2} &= \frac{1}{3} \end{aligned}$$

which meet at the image $[\tilde{x}^0, \tilde{x}^1, \tilde{x}^2] = [3, 2, 0] \sim [1, \frac{2}{3}, 0]$ of the point of intersection $[x^0, x^1, x^2] = [0, 2, 3]$. On graphing, we see that the vertical line $\tilde{X}^1 = \frac{2}{3}$ is indeed tangent to the image of the hyperbola (which is another hyperbola) at the affine point $(\tilde{X}^1, \tilde{X}^2) = (\frac{2}{3}, 0)$. We have slid over a few details here, such as that linear transformations preserve tangency, but the basic idea should be clear.

Everything we have done in this section could be done just as well using the field $\mathbb{C}$ instead of $\mathbb{R}$ or in fact any field you like. The formal manipulations would be the same. Standard algebraic geometry is done over $\mathbb{C}$.

If you would like to try this, here is a little investigation you might try. The two circles

$$\begin{aligned} {X^1}^2 + {X^2}^2 &= a^2 \\ {X^1}^2 + {X^2}^2 &= b^2 \end{aligned}$$

when you have switched over to projective coordinates as before, both go through the *complex* points at $\infty$ with coordinates $[0, i, 1]$ and $[0, -i, 1]$. You can show

[2] In actual performance it might be necessary to do a sequence of projections on the wall.

that the circles are tangent to one another at each of these points by using the complex linear transformation

$$\begin{aligned} x^0 &= i\,\tilde{x}^2 \\ x^1 &= \tilde{x}^2 \\ x^2 &= i\,\tilde{x}^0 \end{aligned}$$

I ought to resist the temptation to mention one of the most beautiful results in mathematics; Bezout's theorem in $\mathbb{P}^2(\mathbb{C})$. This is

> **Bezout's Theorem** A curve of degree $m$ and a curve of degree $n$ intersect in exactly $mn$ points when points at $\infty$ (possibly complex) and multiplicity of intersection are taken into account.

For examples of multiplicity, consider two ellipses which usually intersect at 4 distinct points, but may intersect at two ordinary points and a point where the ellipses are tangent (point of multiplicity 2), or two distinct points where the ellipses are tangent at each of the two points (two points of multiplicity 2) or the ellipses do not intersect in the real plane (in which case they have complex intersections possibly with multiplicity. In all cases the total number of intersections, counting multiplicity is $2 \times 2 = 4$.

A line usually hits a parabola $1 \times 2 = 2$ times but the $X^1$-axis is tangent $X^2 = {X^1}^2$ so $(0, 0)$ is an intersection of multiplicity 2 and the vertical axis $X^1 = 0$ hits the parabola at $(0, 0)$ and also at the point $[0, 0, 1]$ at $\infty$. In all cases there are thus two points of intersection.

For Bezout's theorem to be true, it is critical to be working over the algebraically closed field $\mathbb{C}$. If you enjoyed this excursion there are many wonderful books you can consult to learn more. I have a particular fondness for [**?**] which deals with plane curves in $\mathbb{P}^2(\mathbb{C})$. If you wish to go for the full algebraic geometry experience two beautiful and classic texts are [Sharfarevich] (more concrete, at least to begin with) and [Hartshorne] (more abstract). Another possiblity is [Perrin] which is a shorter introduction and very student oriented but gets a long way.

## 9.3 Weight Functions

We now want to look at the material in the previous section in a different way. A point in $\mathbb{P}^2(\mathbb{R})$ has coordinates $\lambda[x^0, x^1, x^2]$ where $\lambda \in \mathbb{R} - 0$. If we let $\lambda$ run through $\mathbb{R}$, we always get the same point in $\mathbb{P}^2(\mathbb{R})$ (unless $\lambda = 0$) but we get a *line* in 3-space. If $x^0 \neq 0$ there will be a unique point in 3-space on the line with $x^0 = 1$. Thus the plane in 3-space with $x^0 = 1$ is a model for the affine part of $\mathbb{P}^2(\mathbb{R})$. Each infinite point in $\mathbb{P}^2(\mathbb{R})$ has coordinates $\lambda[0, x^1, x^2]$ and this is a line parallel to $M = \{[x^0, x^1, x^2] \mid x^0 = 1\}$. Thus the points of $\mathbb{P}^2(\mathbb{R})$ are in one to one correspondence with the lines through the origin, and indeed $\mathbb{P}^2(\mathbb{R})$ is often defined as "the set of lines through the origin in $\mathbb{R}^3$," which is accurate but confusing to beginners.

With this picture, then, a line in $\mathbb{P}^2(\mathbb{R})$ (visualized as $M$) corresponds to the intersection of a plane through the origin in $\mathbb{R}^3$ with $M$, plus the point at $\infty$ corresponding to a line in the plane through the origin and parallel to $M$. Lines in $\mathbb{R}^3$ in the plane thus correspond one to one with the points on the line in $\mathbb{P}^2(\mathbb{R})$

Similarly a curve in $\mathbb{P}^2(\mathbb{R})$ will correspond to a set of lines going from the origin through the curve in M. If the curve is closed this will be conelike. There will be additional lines parallel to M if the curve has points at $\infty$.

Now the obvious question is "What is so special about the plane $M = \{[x^0, y^1, y^2] \mid x^0 = 1\}$?" The answer is "Nothing at all," as we are going to show. We are going to generalize slightly a construction of Grassmann (called "weighted points") by introducing a linear form for the weight function $\lambda(x)$, which will thus have the form

$$\lambda(x) = \lambda_0 x^0 + \lambda_1 x^1 + \lambda_2 x^2$$

This weight function will usually be kept constant during a discussion. The new model of $\mathbb{P}^2(\mathbb{R})$ will now be

$$M_1 = \{[x^0, y^1, y^2] \mid \lambda(x) = 1\}$$

and we will also use the notation

$$M_r = \{[x^0, y^1, y^2] \mid \lambda(x) = r\}$$

We could use any $M_r$ (plus points at $\infty$) as a model of $\mathbb{P}^2(\mathbb{R})$. However there is no advantage in this and we will stick to $M_1$ as the affine part of our model of $\mathbb{P}^2(\mathbb{R})$.

The points in $M_1$ correspond as before to lines through the origin in $\mathbb{R}^3$ and the points at $\infty$ correspond to lines through the organ parallel to $M_1$. Notice that the plane though the origin parallel to $M_1$, containing the lines that correspond to points at $\infty$, is $M_0$, and the triples in this plane satisfy $\lambda([x^0, y^1, y^2]) = 0$

Thus our new model of $\mathbb{P}^2(\mathbb{R})$ consists of two parts; the points of $M_1$ characterized by $\lambda(x) = 1$ and what we shall call *vectors* which are triples with

$x = [x^0, x^1, x^2]$ with $\lambda(x) = 0$. These correspond as usual to points at $\infty$ when we identify $[x^0, x^1, x^2]$ with $\lambda[x^0, x^1, x^2]$. However, in Grassmann's picture the points at $\infty$ fade out, and their role is replaced by the vectors. We can of course switch pictures in the wink of an eye whenever it is helpful. If we explicitly define the equivalence relation $\sim$ by

$$[y^0, y^1, y^2] \sim [x^0, x^1, x^2] \;\Leftrightarrow\; [y^0, y^1, y^2] = \lambda[x^0, x^1, x^2] \qquad \text{with } \lambda \neq 0$$

then we can write

$$\mathbb{P}^2(\mathbb{R}) = M_1 \cup (M_0/\sim)$$

I mention in passing that topologically $M_0/\sim$ is just the unit circle with antipodal points identified.

Now let us contrast the standard model of projective geometry with Grassmann's which we are about to study. In the standard model we use

$$\lambda([x^0, x^1, x^2]) = 1x^0 + 0x^1 + 0x^2$$

and the only use of the vectors is to determine points at $\infty$.

Grassmann's approach to the subject differs from the standard version in that

- **a.** $\lambda(x) = x^0 + x^1 + x^2$
- **b.** The value of $\lambda(x)$ is used for weights
- **c.** The infinite points are replaced by the vectors of $M_0$

Thus Grassmann's picture amounts to an *augmentation* of $\mathbb{P}^2(\mathbb{R})$ with additional structure, which makes it look a lot like $\mathbb{R}^3$ but interpreting the points in $\mathbb{R}^3$ in a different way, and the reinterpretation is a lot like $\mathbb{P}^2(\mathbb{R})$. Of course, $\mathbb{R}^3$ and $\mathbb{P}^2(\mathbb{R})$ are of different topological type, so if we want to go back and forth some fancy footwork is necessary.

We are doing all our development in $\mathbb{P}^2(\mathbb{R})$ and $\mathbb{R}^3$, but it is important to understand that we could just as well be doing it in $\mathbb{P}^n(\mathbb{R})$ and $\mathbb{R}^{n+1}$. The extension is so obvious and trivial that it is not worth losing the visualization possible here in order to deal with the more general situation.

The next phase of Grassmann's development is to introduce the *weighted point*. Each triple $x = [x^0, x^1, x^2]$ with $\lambda(x) \neq 0$ has a weight given by $\lambda(x)$ and is associated with a point $\tilde{x}$ of $M_1$ given by $\tilde{x} = \frac{1}{\lambda(x)}[x^0, x^1, x^2]$. By the construction, $\lambda(\tilde{x}) = 1$ so $\tilde{x}$ is in $M_1$. Thus we say $x$ is *associated* with the *unit point* $\tilde{x}$ and that $x$ is $\tilde{x}$ with weight $\lambda(x)$. and we can write $x = \lambda(x)\tilde{x}$. Such multiples of unit points are called weighted points.

If $\lambda(x) = 0$ then Grassmann calls x a *free vector*. (There is another kind of vector called a *line bound vector* which we will discuss later.) Notice that the point $[0, 0, 0]$ counts as a vector in Grassmann's system although it plays no role whatever in the standard model of $\mathbb{P}^2(\mathbb{R})$. The weighted points and the free vectors together constitute all triples $[x^0, x^1, x^2]$ and they thus form a vector space, with addition and scalar multiplication as in $\mathbb{R}^3$. In fact they *are* $\mathbb{R}^3$ for all practical purposes, but we interpret the triples differently.

We now wish to discuss the points on a line corresponding to $\mu \in \mathbb{R}^{3*}$, with $\mu \neq \lambda$ and we assume the line goes through the distinct points $x, y \in M_1$. We have $\mu(x) = \mu(y) = 0$ and also $\lambda(x) = \lambda(y) = 1$. Let $\sigma, \tau \in (R)$ and

$$z_1 = \sigma x + \tau y$$

$z_1$ is not a unit point; in fact its weight is

$$\lambda(z_1) = \sigma\lambda(x) + \tau\lambda(y) = \sigma + \tau$$

Set

$$z = \frac{\sigma x + \tau y}{\sigma + \tau} = \frac{1}{\sigma + \tau} \cdot (\sigma x + \tau y)$$

and now we have $\lambda(z) = 1$; hence $z \in M_1$. Also, $z$ is on the line connecting $x$ and $y$ because

$$\mu(z) = \sigma\mu(x) + \tau\mu(y) = \sigma \cdot 0 + \tau \cdot 0 = 0$$

Although we have no metric, it still makes sense to ask the relation between the segments ($x$ to $z$) and ($z$ to $y$). This we now examine.

First note $z - x$ is a vector since

$$\lambda(z - x) = \lambda(z) - \lambda(x) = 1 - 1 = 0$$

This vector is parallel to $M_1$ and also to the line $\mu$. Similarly, $y - x$ is parallel to the line to we must have

$$\begin{aligned}
z - x &= \kappa(y - x) \\
\frac{\sigma x + \tau y}{\sigma + \tau} - x &= \kappa(y - x) \\
\frac{1}{\sigma + \tau}\left[\sigma x + \tau y - (\sigma + \tau)x\right] &= \kappa(y - x) \\
\frac{\tau}{\sigma + \tau}(y - x) &= \kappa(y - x) \\
\frac{\tau}{\sigma + \tau} &= \kappa \\
z - x &= \frac{\tau}{\sigma + \tau}(y - x)
\end{aligned}$$

Similarly

$$y - z = \frac{\sigma}{\sigma + \tau}(y - x)$$

Suppose now that we want $z$ to lie $\frac{1}{3}$ of the way from $x$ to $y$. Even without a metric we can express this as

$$z - x = \tfrac{1}{3}(y - x)$$

For this we need $\tau/(\sigma + \tau) = \frac{1}{3}$ which gives

$$z = \frac{\sigma x + \tau y}{\sigma + \tau} = \frac{2\tau x + \tau y}{2\tau + \tau} = \frac{2x + y}{3} = \tfrac{2}{3}x + \tfrac{1}{3}y$$

In words, if we want $z$ to be $\frac{1}{3}$ of the way from $x$ to $y$, we weight $x$ twice as heavily as $y$ and make sure the weights add up to 1; that is $\frac{2}{3}$ and $\frac{1}{3}$.

Notice that the ratio of $\frac{\tau}{\sigma+\tau}$ and $\frac{\sigma}{\sigma+\tau}$ is just $\frac{\tau}{\sigma}$

A better way to think of the previous example is as follows; we wish $z$ to divide the segment from $x$ to $y$ in the ratio 1 to 2. This gives preliminary values $\sigma_1 = 2$ and $\tau_1 = 1$ and then they must be scaled by their sum to $\sigma = \frac{2}{3}$ and $\tau = \frac{1}{3}$

More generally we wish $z$ to divide the segment from $x$ to $y$ in the ratio $\tau$ to $\sigma$. Then it is only necessary to rescale $\tau$ and $\sigma$ by their sum and form $z = \sigma x + \tau y$. Note carefully the order, which might seem a little counterintuitive.

Grassmann's explanation of this runs along the following lines. If we have unit points $x$ and $y$ we think of these points as having 2 kilogram and 1 kg weights attached. Then $z$ is the balance point of the segment. This physical interpretation can be very useful for many purposes, both for intuitive understanding and practical calculation.

Grassmann's original explanation of these ideas mixed up the previous theory of weighted points and the wedge product. I have separated them as they do have separate origins. We will combine them usefully later.

To proceed further, it is useful to introduce a notation which might at first seem somewhat alarming. Suppose we have two parallel vector $u$ and $v$ and that

$$v = \kappa u$$

Then I propose to write

$$\kappa = \frac{v}{u}$$

The division notation for vectors is to be used *only when the vectors are parallel* in which case it makes perfect sense. We *never* use this notation with vectors that are not parallel. The purpose of this is to eliminate very clumsy circumlocutions.

We are now going to use Grassmann's theory of weighted points to do some geometry. We start with a simple result; The medians of a triangle meet at a point two thirds of the way along each median.

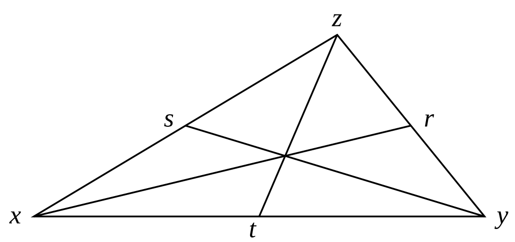



Figure 9.1: Medians meet at a point

Let $x$, $y$ and $z$ be the three vertices of the triangle visualized in $M_1$ so that they are unit points. Let $r$,$s$ and $t$ be the midpoints of the sides $xy$, $yz$ and $zx$

(using $x$, $y$ and $z$ here as mere labels, not products). Then

$$r = \frac{x+y}{2} \qquad s = \frac{y+z}{2} \qquad z = \frac{z+x}{2}$$

We claim the three medians meet at a point $w$ and additionally that

$$w = \frac{x+y+z}{3}$$

We need to show that $w$ is a linear combination of $x$ and $s$ and that it divides the segment $xs$ in the ration 2 to 1. But $2s = y + z$ so

$$w = \frac{x+2s}{3} = \tfrac{1}{3}x + \tfrac{2}{3}s$$

which is exactly the requirement for $w$ to divide the segment in the ratio 2 to 1. Similarly $w$ will be a linear combination or $y$ and $t$ and of $z$ and $r$, and divide the segments in the proper ratio. This completes the proof.

You might wonder if there is a way to prove the similar theorem about the angle bisectors meeting at a point. This can be done but it requires *weighted lines*, not weighted points. Since the two theories are identical due to duality, there is no difficulty.

The preceding theorem is a special case of the theorem of Ceva[3]: If we connect the vertices of a triangle to the opposite sides, the three lines will be concurrent if and only if the ratios of the divisions of the sides multiply to 1. (We remind the reader of the terminology that three lines are *concurrent* if they all meet at a point.)

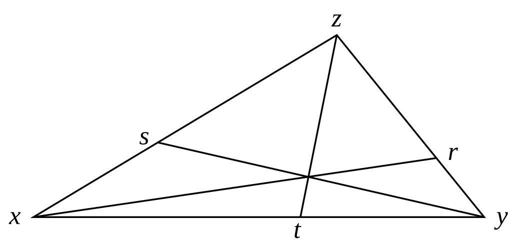



Figure 9.2: Theorem of Ceva

More explicitly using unit points $x$,$y$ and $z$,

if $r$ divides segment $yz$ in the ratio $\frac{\alpha}{\beta}$

and $s$ divides segment $zx$ in the ratio $\frac{\gamma}{\delta}$

and $t$ divides segment $xy$ in the ratio $\frac{\epsilon}{\zeta}$

Then segments $xr$,$ys$ and $zt$ are concurrent $\Leftrightarrow \frac{\alpha}{\beta}\frac{\gamma}{\delta}\frac{\epsilon}{\zeta} = 1$

[3]Ceva is pronounced Chayva approximately

We will prove the $\Rightarrow$ direction. The converse is left to the reader. Without loss of generality we may assume that the coefficients have been normalized: $\alpha+\beta=\gamma+\delta=\epsilon+\zeta=1$. Then "$r$ divides segment $yz$ in the ratio $\frac{\alpha}{\beta}$" means

$$\begin{aligned}
\frac{z-r}{r-y} &= \frac{\alpha}{\beta} \\
\beta z-\beta r &= \alpha r-\alpha y \\
\alpha y+\beta z &= \alpha r+\beta r \;=\; (\alpha+\beta)r \;=\; r
\end{aligned}$$

Similarly

$$\begin{aligned}
s &= \gamma z+\delta x \\
t &= \epsilon x+\zeta y
\end{aligned}$$

Since w is linearly dependent on $x$ and $r$, there must be constants $\eta_1$ and $\eta_1$ so that

$$w \;=\; \theta_1 x+\eta_1 r$$

Similarly, since the lines are assume concurrent,

$$\begin{aligned}
w &= \theta_2 y+\eta_2 s \\
w &= \theta_3 z+\eta_3 t
\end{aligned}$$

Then we have

$$\begin{aligned}
w &= \theta_1 x+\eta_1 r \;=\; \theta_2 y+\eta_2 s \;=\; \theta_3 z+\eta_3 t \\
&= \theta_1 x+\eta_1(\alpha y+\beta z) \;=\; \theta_2 y+\eta_2(\gamma z+\delta x) \;=\; \theta_3 z+\eta_3(\epsilon x+\zeta y) \\
&= \theta_1 x+\eta_1\alpha y+\eta_1\beta z \;=\; \theta_2 y+\eta_2\gamma z+\eta_2\delta x \;=\; \theta_3 z+\eta_3\epsilon x+\eta_3\zeta y
\end{aligned}$$

Since $x$, $y$, $z$ form a triangle, they must be linearly independent, so we have

$$\begin{aligned}
\theta_1 &= \eta_2\delta \;=\; \eta_3\epsilon \\
\eta_1\alpha &= \theta_2 \;=\; \eta_3\zeta \\
\eta_1\beta &= \eta_2\gamma \;=\; \theta_3
\end{aligned}$$

Thus

$$\frac{\alpha}{\beta}=\frac{\eta_1\alpha}{\eta_1\beta}=\frac{\theta_2}{\theta_3} \qquad \frac{\gamma}{\delta}=\frac{\eta_2\gamma}{\eta_2\delta}=\frac{\theta_3}{\theta_1} \qquad \frac{\epsilon}{\zeta}=\frac{\eta_3\epsilon}{\eta_3\zeta}=\frac{\theta_1}{\theta_2}$$

so

$$\frac{\alpha}{\beta}\cdot\frac{\gamma}{\delta}\cdot\frac{\epsilon}{\zeta}=\frac{\theta_2}{\theta_3}\cdot\frac{\theta_3}{\theta_1}\cdot\frac{\theta_1}{\theta_2}=1$$

as required. Note how the use of Grassmann's weighted points reduces the geometry to trivial algebraic calculations, the geometric details of which we do not need to consider, just as in ordinary analytic geometry.

Up to this point we have been emotionally locating the point $z$ *between* $x$ and $y$ which then presumes that $\sigma$ and $\tau$ are positive. However, if we attach a positive weight to $x$ and a negative weight ($\tau < 0$, point being pulled up) to $y$ then the balance point will be outside the segment $xy$ on the same side as $x$ is from $y$ as shown in the figure. This figure illustrates

$$z = \frac{4}{3}x + (-\frac{1}{3})y$$

Then, as usual, we have

$$\frac{y-z}{z-x} = \frac{\frac{4}{3}}{-\frac{1}{3}} = -4$$

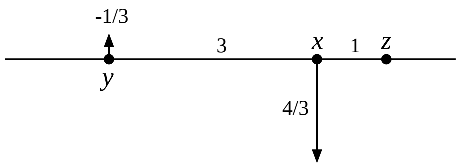



Figure 9.3: Point with negative weight

Let's analyze this in terms of lever arms(torque). If we scale $z-x$ to be 1, then $z-y=4$. The mass times the lever arm from $z$ is $-\frac{1}{3}\cdot 4 = -\frac{4}{3}$ for $y$ and $\frac{4}{3}\cdot 1 = \frac{4}{3}$ for $x$. Thus the total lever arm is thus

$$\sum m_i(\text{distance}_i) = -\frac{4}{3} + \frac{4}{3} = 0$$

as it should be if z is the balance point. Of course, there is no real "distance" here; we are just counterfeiting distance with the position coordinates.

In general there is little need to think explicitly of these matters (lever arms) as the algebra will take care of itself, was we will see in the following theorem of Menelaos[4]

**Theorem of Menelaos** Let $x, y, z$ be unit points forming a triangle and let $t$ be on line $xy$, $r$ on line $yz$ and $s$ on line $zx$. Then the points $r, s, t$ are collinear if and only if the product of the ratios of the divisions of the the segments $xy$,$yz$ and $zx$ equals -1.

We will prove the direction $\Rightarrow$ and leave the converse to the reader. When doing this the trick is to prove the $r$ is a linear combination of $s$ and $t$.

Suppose the three points are collinear. The ratios are

$$\begin{aligned} \frac{z-r}{r-y} &= \frac{\alpha}{\beta} & \alpha+\beta &= 1 \\ \frac{x-s}{s-z} &= \frac{\gamma}{\delta} & \gamma+\delta &= 1 \\ \frac{y-t}{t-x} &= \frac{\epsilon}{\zeta} & \epsilon+\zeta &= 1 \end{aligned}$$

[4]Menelaos is pronounced in Greek with the accent on the second e and the la sounded as a separate syllable: Me-nè-la-os. The spelling Menelaus is a Roman barbarism.

and these lead, as in the Theorem of Ceva, to

$$\begin{aligned} r &= \alpha y + \beta z \\ s &= \gamma z + \delta x \\ t &= \epsilon x + \zeta y \end{aligned}$$

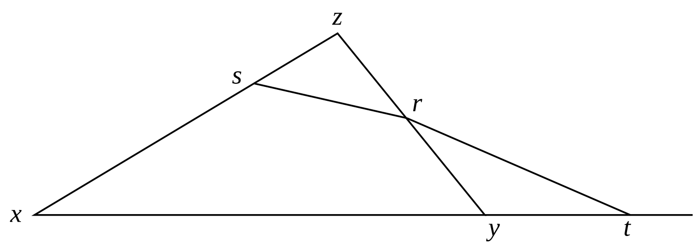



Figure 9.4: Theorem of Menelaos

Since the three points $r, s, t$ are assumed collinear, they are linearly dependent, so their are constants $\theta_1, \theta_2, \theta_3$ for which

$$\begin{aligned} \theta_1 r + \theta_2 s + \theta_3 t &= 0 \\ \theta_1(\alpha y + \beta z) + \theta_2(\gamma z + \delta x) + \theta_3(\epsilon x + \zeta y) &= 0 \\ (\theta_2\delta + \theta_3\epsilon)x + (\theta_1\alpha + \theta_3\zeta)y + (\theta_1\beta + \theta_2\gamma)z &= 0 \end{aligned}$$

By the linear dependence of $x, y$ and $z$ we now have

$$\theta_2\delta + \theta_3\epsilon = \theta_1\alpha + \theta_3\zeta = \theta_1\beta + \theta_2\gamma = 0$$

thus

$$\frac{\epsilon}{\delta} = -\frac{\theta_2}{\theta_3} \qquad \frac{\alpha}{\zeta} = -\frac{\theta_3}{\theta_1} \qquad \frac{\gamma}{\beta} = -\frac{\theta_1}{\theta_2}$$

$$\frac{\alpha}{\beta}\cdot\frac{\gamma}{\delta}\cdot\frac{\epsilon}{\zeta} = \frac{\alpha}{\zeta}\cdot\frac{\gamma}{\beta}\cdot\frac{\epsilon}{\delta} = (-\frac{\theta_3}{\theta_1})(-\frac{\theta_1}{\theta_2})(-\frac{\theta_2}{\theta_3}) = -1$$

as required.

There is the possibility that one of $r$,$s$ or $t$ might wander off to $\infty$, as $t$ would if $r = (y+z)/2$ and $s = (z+x)/2$. Nevertheless, properly interpreted the theorem is still true. The best way to do this would be to reinterpret ratios $\alpha/\beta$ as points in $[\beta, \alpha] \in \mathbb{P}^1(\mathbb{R})$ and work from there. However, this would have added another level of complication to the exposition, so I refrained, since this is not a book on projective geometry but only an illustration of how nicely Grassmann's weighted points perform there.

Once more I remind the reader that the notation

$$\frac{z-r}{r-y} = \frac{\alpha}{\beta}$$

where $z - r$ and $r - y$ are *vectors* is only a symbolic way of writing

$$z - r = \frac{\alpha}{\beta}(r - y)$$

the latter being a legitimate equation between vectors. In particular $z - r$ and $r - y$ are not distances since we have introduced no metric, although when we stay on a single line through the origin they do perform rather like signed distances.

# Bibliography


[Gelfand] Gelfand, I.M, LECTURES ON lINEAR ALGERBRA New York. Wiley, Interscience. 1963

[Grassmann 44] Grassmann, H.G, DIE LINEALE AUSDEHNUNGSLEHRE Leipzig: Wiegand 1844

[Grassmann 62] Grassmann, H.G, DIE LINEALE AUSDEHNUNGSLEHRE Berlin: Enslin. 1862

[Hartshorne] Hartshorne, Robin, ALGEBRAIC GEOMETRY, Springer, Berlin 1977

[Malcev] Malcev, A.I, FOUNDATIONS OF LINEAR ALGEBRA, 1963

[Perrin] Perrin, Daniel, GÉOMÉTRY, ALGÉBRIQUE Paris 2001

[Sharfarevich] Shafarevich, Igor R., ALGEBRAIC GEOMETRY, 2nd Editon Springer, Berlin 1994

[Sternberg] Sternberg, Schlomo LECTURES ON DIFFERENTIAL GEOMETRY