Tensor Products, Wedge Products and Differential Forms

Phil Lucht
Rimrock Digital Technology, Salt Lake City, Utah 84103
last update: June 4, 2016
rimrock@xmission.com

Maple code is available upon request. Comments and errata are welcome.

The graphics look ratty in Windows Adobe PDF viewers when not scaled up, but look just fine in this excellent freeware viewer: http://www.tracker-software.com/product/pdf-xchange-viewer .

The table of contents has live links. Most PDF viewers provide these links as bookmarks on the left.

Overview and Summary

This monograph is meant as a user guide for both tensor products and wedge products. These objects are sometimes glossed over in literature that makes heavy use of them, the assumption being that everything is obvious and not worth describing too much. As we shall show, there is in fact quite a lot to be said about tensor and wedge products, and much of it is not particularly obvious.

Our final chapter discusses aspects of differential k-forms which inhabit the wedge product spaces, with an emphasis on the notion of pullbacks and integration on manifolds.

We attempt to include both the mathematical view and the engineering/physics view of things, but the emphasis is on the latter. The discussion is more about activities in the engine room and less about why the ship travels where it does.

The study of wedge products is known as the **exterior algebra** and is credited to Grassmann.

Maple is used as appropriate to do basic calculations. Covariant notation is used throughout.

Equations which are repeats of earlier ones are shown with italic equation numbers.

Here is a brief summary of our document which has ten Chapters and eight Appendices :

Chapter 1 surveys the **mathematician's** description of the **tensor product** as a quotient space, and then places the tensor product in the framework of category theory. This approach is resumed much later in Chapter 9 for the wedge product, after the reader is more familiar with that object.

Chapter 2 reviews **tensor algebra** and then introduces a meaning for the tensor product symbol ⊗ in terms of outer products of tensors. After a quick review of tensor expansions and projections, the last section introduces the notion of a **dual space** and includes the use of the **Dirac bra-ket** notation. The notion of a **tensor function** is introduced.

Chapter 3 discusses the theory of Chapter 1 versus the practicality of Chapter 2 in terms of outer products. It then derives the **Kronecker product** of two matrices in covariant notation. This topic is somewhat tangential to the main development, but is included since it is sometimes not explained very well in the literature. Maple is used to compute a few such Kronecker products.

Chapter 4 has four parts involving **products of two vectors** and their vector spaces: tensor product, dual tensor product, wedge product, and then dual wedge product. This chapter serves as an introduction to the four chapters which follow.

Chapters 5, 6, 7, 8 continue this order of presentation for products of k vectors and then for products of any number of general tensors. The order is: tensor product (Ch 5), dual tensor product (Ch 6), wedge product (Ch 7) and then dual wedge product (Ch 8). The chapters intentionally have a high degree of parallelism, though some details are omitted from the later chapters to reduce repetition. The dual tensor chapters involve tensor functions as the closure of tensor functionals onto a general set of vectors. The tensor-product tensor functions are multilinear, whereas the wedge-product ones are multilinear and totally antisymmetric. Alternate wedge product normalizations are discussed. The reader is warned that these four chapters (especially the last two) are exceedingly tedious because there is a huge amount of detail involved in laying out these subjects. The silver lining is that all notations are heavily exercised and many examples are provided.

Chapter 9 returns to the mathematician's world giving two descriptions of the wedge product in terms of quotient spaces.

Chapter 10 presents an outline of **differential k-forms** and **pullbacks** with an emphasis on underlying transformations. The contents of Chapter 2 on covariant tensor algebra and Chapter 8 on dual wedge products (exterior algebra) come into play. Various k-form facts are derived and cataloged. Manifolds are described without rigor, leading to a discussion of the integration of both functions and k-forms over manifolds. A special notation is used to distinguish dual space functionals like λ^i = <u^i| = dx^i from calculus differentials like dx^i, and no wedge product hats ^ are suppressed. The topics of boundaries ∂M and orientation are mentioned only in passing, with reference to other sources. Although the generalized Stokes' Theorem for differential forms on manifolds is not derived, it is nevertheless used. Our goal is to expose underpinning structures which are sometimes ignored. Multiindex and Dirac notations are used side-by-side with full index display and normal vector notation.

Appendix A explains our **permutation** notation and the powerful rearrangement theorems used in various proofs throughout the document. The Alt and Sym operator properties are presented in a generic permutation space, and then those generic results are applied to tensors and tensor functions. The permutation tensor ε is given honorable mention, and a few obscure theorems are proved.

Appendix B discusses the **direct sum** of vectors, vector spaces and operators in those spaces.

Appendix C shows that when one antisymmetrizes a product of tensors, pre-antisymmetrizing one or more of those tensors makes no difference in the result.

Appendix D shows how tensors and tensor functions are the same objects expressed in different bases.

Appendix E gives details of the "transformation kinematics package" for **x**' = **F**(**x**) translated to **x** = **φ**(**t**).

Appendix F derives the fact that det(R^TR) is the volume of an n-piped in R^m ($n \leq m$) , where R is a matrix whose column vectors span the n-piped. This result is then used to write an expression for the volume of the differential n-piped for the tangent space T_xM associated with a point **x** on a manifold.

Appendix G shows that det(R^TR) equals the sum of the squared full-width minors of R, and then relates this fact to the measures appearing in pulled-back differential n-form integrals.

Appendix H describes properties of the Hodge dual operator (called *). It then derives certain Hodge correspondences between differential forms and differential operators, and shows how the generalized Stokes' Theorem produces many integral theorems of analysis. The last section converts Maxwell's four partial differential equations to two differential form equations, one of which is dα = 0.

Notation

This list gives most of the symbols used in the document and should give the reader an idea of the general flavor of the presentation. Vectors are sometimes bolded, sometimes not. Vector functionals and tensors of rank 2 or greater are never bolded. In general, dual objects are given Greek or script symbols. Unfortunately certain symbols have several unrelated meanings.

$M^{\mathcal{T}}$, M^{T}	covariant transpose and matrix transpose of a matrix M, see 2.11 (f).
m x n	used to describe a matrix which has m rows and n columns
det(M)	determinant of the square matrix M
det[**a**,**b**,**c**...]	determinant of a square matrix whose columns are vectors **a**,**b**,**c**...
R^m	Euclidean space with m dimensions
rank	rank of a matrix; rank of a tensor or tensor function

⊗	tensor product of spaces or objects in those spaces
⊕	direct sum of spaces or objects in those spaces (App B)
x	Cartesian product, as in VxW with element (v,w)
^	wedge product of spaces or objects in those spaces
K	a real field (such as the reals, or such as binary {0,1})
s_i	scalars in K
≡	is defined as
*	simple multiplication; complex conjugation; the Hodge star operator; pullback function F*
g	metric tensor; generic function name
f	generic function name
o	function composition operator, as in h = (f o g)

F(**x**)	a general transformation **x**' = **F**(**x**)
φ(**t**)	a general transformation **x** = **φ**(**t**) (alternate notation to the above)
$R^i{}_j$	the differential of transformation **F** (or **φ**), down-tilt matrix element
R,S	differentials of transformations **F** and $\mathbf{F}^{-1}$ written as matrices
$\mathcal{R}$,$\mathcal{S}$	corresponding Dirac space operators, see 2.11 (g)
F*	pullback function for **x**' = **F**(**x**), defined (10.7.17)
φ*	pullback function for **x** = **φ**(**t**)

Alt	total antisymmetrization operator (App A), short for Alternating
Sym	total symmetrization operator

V	real vector space of dimension n	
a,**b**	vectors in V	$\lvert\mathbf{a}\rangle$ a Dirac ket
a • **b**	scalar product of two vectors (real)	$\langle\mathbf{a}\lvert\mathbf{b}\rangle = \langle\mathbf{b}\lvert\mathbf{a}\rangle$
V,**v**	vectors in V	$\lvert\mathbf{v}\rangle$
$\mathbf{u}_i$	axis-aligned basis vector in V	$\lvert\mathbf{u}_i\rangle$
$\mathbf{v}_i$	vector with label i in V	$\lvert\mathbf{v}_i\rangle$
v_i	covariant component i of the vector **v**	$\langle\mathbf{u}_i\lvert\mathbf{v}\rangle$

$\mathbf{e}_i$	tangent base vector in V	$\|\mathbf{e}_i>$
$\mathbf{e}^i$	dual of the above	$\|\mathbf{e}^i>$
W	real vector space of dimension n'	
$\mathbf{e'}_i$	basis vector in W	$\|\mathbf{e'}_i>$
$\mathbf{w}$	vector in W	$\|\mathbf{w}>$

$\mathbf{a} \otimes \mathbf{b}$	tensor product of two vectors = a pure element of the vector space $V^2 = V\otimes V$
$\mathbf{u}_i\otimes\mathbf{u}_j$	tensor product of two basis vectors = basis vector of $V^2 = V\otimes V$
$\mathbf{a} \wedge \mathbf{b}$	wedge product of two vectors = a pure element of the vector space $L^2 = V\wedge V \subset V^2$
$\mathbf{u}_i\wedge \mathbf{u}_j$	wedge product of two basis vectors = basis vector of $L^2 = V\wedge V \subset V^2$

T	tensor of rank k	$\|T>$	$T = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k}\ \mathbf{u}_{i_1}\otimes \mathbf{u}_{i_2} \ldots\ldots\otimes \mathbf{u}_{i_k}$
S	tensor of rank k'	$\|S>$	
R	tensor of rank k"	$\|R>$	

$T_\wedge$	general element of L^k	$T_\wedge = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1}\wedge u_{i_2}\wedge \ldots\ldots\wedge u_{i_k})$
$S_\wedge$	general element of $L^{k'}$	
$R_\wedge$	general element of $L^{k"}$	

V^*	dual space to V	
α,β	vector functionals in V^*	$\alpha = <\boldsymbol{\alpha}\|$, a Dirac bra
α_i	vector functional in V^* with label i	$\alpha_i = <\boldsymbol{\alpha}_i\|$
λ^i	basis vector in V^*	$\lambda^i = <\mathbf{e}^i\| = (\mathbf{e}^i)^T$
dx^i	cosmetic notation for λ^i, an example of a 1-form, used in Chapter 10	
$\mathrm{d}x^i$	a normal calculus differential	
∂_i	abbreviation for $\partial/\partial x^i$	

$dV = dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n$	differential volume form in R^n
$dA^i = {}^*dx^i$ and $d\mathbf{A} = {}^*d\mathbf{x}$	differential "area" form in terms of Hodge dual of vector (App H)

$\lambda^I = \lambda^{i_1} \otimes \lambda^{i_2} \ldots\ldots\otimes \lambda^{i_k}$	multiindex notation for a tensor product of basis functionals
$\lambda_\wedge{}^I = \lambda^{i_1} \wedge \lambda^{i_2} \ldots\ldots\wedge \lambda^{i_k}$	multiindex notation for a wedge product of basis functionals
$dx_\wedge{}^I = \lambda_\wedge{}^I = dx^{i_1} \wedge dx^{i_2} \ldots\ldots\wedge dx^{i_k}$	the above line in cosmetic notation

$\alpha \otimes \beta$	tensor product of two dual vectors = a pure element of the vector space $V^{*2} = V^*\otimes V^*$
$\lambda^i\otimes\lambda^j$	tensor product of two dual basis vectors of V^* = basis vector of $V^{*2} = V^*\otimes V^*$
$\alpha \wedge \beta$	wedge product of 2 dual vectors = a pure element of the vector space $\Lambda^2 = V^*\wedge V^* \subset V^{*2}$
$\lambda^i \wedge \lambda^j$	wedge product of dual basis vectors, basis element of $\Lambda^2 = V^*\wedge V^* \subset V^{*2}$

α,β	names used in Chapter 10 for differential k-forms
$d\alpha$	exterior derivative of a k-form α (Chapter 10)

$\mathcal{T}$	tensor functional of rank k	$<T\|$	$\mathcal{T} = \Sigma_{i_1 i_2 \ldots i_k} T_{i_1 i_2 \ldots i_k} \lambda^{i_1} \otimes \lambda^{i_2} \ldots\ldots \otimes \lambda^{i_k}$
$\mathcal{S}$	tensor functional of rank k'	$<S\|$	
$\mathcal{R}$	tensor functional of rank k"	$<R\|$	

$\mathcal{T}^\wedge$	general element of Λ^k	$\mathcal{T}^\wedge = \Sigma_{i_1 i_2 \ldots i_k} T_{i_1 i_2 \ldots i_k} (\lambda^{i_1} \wedge \lambda^{i_2} \ldots\ldots \wedge \lambda^{i_k})$
$\mathcal{S}^\wedge$	general element of $\Lambda^{k'}$	
$\mathcal{R}^\wedge$	general element of $\Lambda^{k''}$	

$\lambda^i(\mathbf{v})$	dual basis vector tensor function, rank-1	$<\mathbf{u}^i \|\mathbf{v}>$
$\alpha(\mathbf{v})$	general rank-1 tensor function	$<\boldsymbol{\alpha} \|\mathbf{v}>$
$\mathcal{T}(\mathbf{v}_1,\mathbf{v}_2)$	rank-2 tensor function	$<T \| \mathbf{v}_1, \mathbf{v}_2>$

I, J	multiindices	
Σ_I	$\Sigma_{i_1 i_2 \ldots i_k}$	symmetric sum
Σ'_I	$\Sigma_{i_1<i_2<\ldots<i_k}$	ordered sum (increasing)

V^k	vector space of rank-k tensors $\|T>$
V^{*k}	vector space of dual rank-k tensors = rank-k tensor functionals $<T\|$
V^{*k}_f	vector space of rank-k tensor functions $\mathcal{T}(v_I) = <T\| v_I>$ (k-multilinear)

L^k	vector space of totally antisymmetric rank-k tensors
Λ^k	vector space of totally antisymmetric dual rank-k tensors
Λ^k_f	vector space of totally antisymmetric rank-k tensor functions (k-multilinear)

T(V)	$V^0 \oplus V \oplus V^2 \oplus V^3 \oplus \ldots\ldots$	the tensor algebra
T(V*)	$V^{*0} \oplus V^* \oplus V^{*2} \oplus V^{*3} \oplus \ldots\ldots.$	dual tensor algebra
L(V)	$L^0 \oplus L^1 \oplus L^2 \oplus L^3 + \ldots.$	exterior tensor algebra
Λ(V)	$\Lambda^0 \oplus \Lambda^1 \oplus \Lambda^2 \oplus \Lambda^3 + \ldots.$	dual exterior tensor algebra

$\varepsilon_{i_1 i_2 \ldots i_k}$	rank-k permutation tensor
P	permutation operator
S(P)	swaps in a permutation
$(-1)^{S(P)}$	swap parity of a permutation

T_xM	tangent space at a point **x** on a manifold M

1. The Tensor Product

There are two theoretical paths leading to the tensor product. These are briefly summarized in a non-rigorous manner in Sections 1.1 and 1.2 below, after a comment on terminology.

Tensor Product vs Direct Product

The tensor product described below sometimes goes by other names.

In quantum mechanics, a system of two particles might be in a quantum state $|\psi_1> \otimes |\psi_2>$ which is an element of a tensor product space $V_1 \otimes V_2$ (as we shall describe below). Some quantum authors refer to this tensor product as a direct product (e.g. Shankar pp 248-250) while others call it a tensor product (e.g. Messiah p 252, 307). It happens that in quantum theory states like $|\psi_1>$ reside in a vector space which is also a Hilbert space. Similarly, when a quantum system has a symmetry, such as rotational invariance (e.g. an isolated atom), the quantum states can be classified into certain vector spaces associated with the matrix representations of the symmetry group, and the tensor products of these spaces are usually called direct products. For example, the rotation group has matrix representations called "j" = (n/2) for any integer n (matrices are n+1 x n+1), and one writes for example $j_1 \otimes j_2$ to indicate the "direct product" of two such spaces.

Sometimes the tensor product is called a tensor direct product, which phrase seems associated with the outer product componentization of the tensor product noted in Chapter 3.

Occasionally the raw Cartesian product (see below) is called a tensor product, but usually there is some additional structure involved.

Generally, the term direct product seems most suitable for the direct product of groups, rings, modules and related objects, whereas in the current document we are discussing the tensor product of vector spaces and of the tensors contained within those spaces.

Category Theory mentioned below attempts to put all these products into a uniform framework.

1.1 The Tensor Product as a Quotient Space

It does seem odd that one might think of a product V⊗W in terms of a quotient. We shall outline how this path goes in a series of steps. The key results are stated in Steps 7 and 9.

1. Cartesian Product. Start with the inert **Cartesian product** set VxW with elements (v,w), where in our application the sets V and W are vector spaces. This set VxW is "inert" in the sense that one has no instructions for what can be done with its elements.

2. Space F(VxW). We now endow VxW with an addition operator + and a scalar multiplication operator (indicated by juxtaposition) allowing us to form linear combinations of elements of VxW with scalar coefficients. Let's define F(VxW) to be a space which contains all such linear combinations. A typical element of this F(VxW) space might be $3(v_1,w_3) - 2.1(v_2,w_5)$. Of course (v_1,w_3) also lies in F(VxW) and one might call this a **pure** element, whereas $3(v_1,w_3) - 2.1(v_2,w_5)$ is a **mixed** element. Because the sum of two linear combinations is again a linear combination of the same form, the space F(VxW) is closed under addition.

3. Field K. We usually assume (as above) that the scalars are in the field R of real numbers, but to be more general one can assume the scalars are elements of some arbitrary field traditionally called K (though sometimes F or $\mathcal{F}$). In addition to the reals R, there are various fields having an infinite number of elements (like rational or complex numbers), and there are various fields of having a finite number of elements (the Galois Fields).

Footnote: Sometimes the space F(VxW) is described as a "free vector space" which is a set of functions f such that f: VxW → K. Usually such spaces are defined over a discrete set S, and it is not clear how this works when the set S is continuous, this being the case for S = VxW. Moreover, the functions f mapping to K cannot be identified with our linear combinations since for example (v_2,v_5) is not an element of K. We therefore refrain from giving F(VxW) this moniker and the reader should regard F(VxW) only as we have defined it above.

4. Equivalence Relations and Classes. Now define the following set of "equivalence relations"

$$\begin{aligned}
&(v_1+v_2, w) \sim (v_1,w) + (v_2,w) && \text{for all } v_1,v_2 \in V \text{ and all } w \in W \\
&(v, w_1+w_2) \sim (v,w_1) + (v,w_2) && \text{for all } v \in V \text{ and all } w_1, w_2 \in W \\
&s(v,w) \sim (sv,w) && \text{for all } v \in V \text{ and all } w \in W \text{ and all } s \in K \\
&s(v,w) \sim (v,sw) && \text{for all } v \in V \text{ and all } w \in W \text{ and all } s \in K
\end{aligned} \qquad (1.1.1)$$

where ~ means "is equivalent to" and s is a scalar in K. Rewrite these relations as

$$\begin{aligned}
&(v_1+v_2, w) - (v_1,w) - (v_2,w) \sim 0 \\
&(v, w_1+w_2) - (v,w_1) - (v,w_2) \sim 0 \\
&s(v,w) - (sv,w) \sim 0 \\
&s(v,w) - (v,sw) \sim 0 \,.
\end{aligned} \qquad (1.1.2)$$

We are declaring here that lots of linear combinations in F(VxW) are equivalent to 0. The reason we do this is to make our tensor product space (to be defined below) have "nice properties" (*i.e.*, it is then a vector space).

These linear combinations taken together define an "equivalence class" which is equivalent to 0. Call this class N (for null).

5. The Quotient F(VxW)/N. There then exists a space which we shall call F(VxW)/N, or F(VxW) "mod" N. This is a standard structure in equivalence class theory where one takes the quotient of one space S divided by another space of equivalent items in space S, often written S/~. The upshot is that the elements of the new quotient space F(VxW)/N consist of all linear combinations of F(VxW) *except that* any linear combination which has one of the four forms shown above is filtered out ("modded out") by setting it equal to 0.

Example:

$$\begin{aligned}
&3(v_3,w_4) + (v_2, w_1+w_2) - (v_2,w_1) - (v_2,w_2) = \text{an element of } F(VxW) \\
&3(v_3,w_4) = \text{the corresponding element of } F(VxW)/N.
\end{aligned} \qquad (1.1.3)$$

<u>6. The Space V⊗W.</u> We now give this space F(VxW)/N a new name:

$$F(VxW)/N = V\otimes W = \text{the tensor product space of V and W} . \qquad (1.1.4)$$

The elements of V⊗W are linear combinations of elements called v⊗w instead of (v,w) as a reminder that the equivalence class N must be respected. Whereas the comma in (v,w) was a mere separation operator, the ⊗ in v⊗w is regarded as a new "tensor product multiplication operator" with the properties listed below which, in effect, implement the equivalence relations stated above.

<u>7. Practical Summary</u>. The **end result** of all this song and dance is the following:

The tensor product space V⊗W is the set of all linear combinations of elements (v,w) of the Cartesian product set VxW, written as v⊗w, where the following rules are *declared by fiat*:

$$\begin{aligned}
&(v_1+v_2) \otimes w = (v_1\otimes w) + (v_2\otimes w) && \text{for all } v_1,v_2 \in V \text{ and all } w \in W \\
&v \otimes (w_1+w_2) = (v\otimes w_1) + (v\otimes w_2) && \text{for all } v \in V \text{ and all } w_1, w_2 \in W \\
&s(v\otimes w) = (sv)\otimes w && \text{for all } v \in V \text{ and all } w \in W \text{ and all } s \in K \\
&s(v\otimes w) = v\otimes(sw) . && \text{for all } v \in V \text{ and all } w \in W \text{ and all } s \in K \qquad (1.1.5)
\end{aligned}$$

The first two rules state that ⊗ multiplication is distributive over addition (from right and left), while the last two rules state the scalars work in the expected manner.

If these rules were declared for a *function* f(v,w), they would appear as

$$\begin{aligned}
&f(v_1+v_2,w) = f(v_1,w) + f(v_2,w) \\
&f(v,w_1+w_2) = f(v,w_1) + f(v,w_1) \\
&s\, f(v,w) = f(sv,w) \\
&s\, f(v,w) = f(v,sw) \qquad (1.1.6)
\end{aligned}$$

Such a function would then be described as being **bilinear** because it is linear separately in each argument with the other argument held fixed. One can then regard the rules shown above for ⊗ as expressing bilinearity for the tensor product space V⊗W.

Usually the above scalar and distributive rules are combined into the slightly more compact form,

$$\begin{aligned}
&(s_1v_1+s_2v_2) \otimes w = s_1(v_1\otimes w) + s_2(v_2\otimes w) \\
&v \otimes (s_1w_1+s_2w_2) = s_1(v\otimes w_1) + s_2(v\otimes w_2) \qquad (1.1.7)
\end{aligned}$$

and similarly for a bilinear function,

$$\begin{aligned}
&f(s_1v_1+s_2v_2,w) = s_1f(v_1,w) + s_2f(v_2,w) \\
&f(v, s_1w_1+s_2w_2) = s_1f(v,w_1) + s_2f(v,w_2) . \qquad (1.1.8)
\end{aligned}$$

8. v⊗w does not commute. Whereas the + operation within V⊗W is commutative, it should be clear that the ⊗ operation is **not commutative**. If $v \in V$ and $w \in W$, then $v\otimes w \in V\otimes W$ whereas w⊗v is an element of a completely different space which is W⊗V. Even if W = V, one has $v\otimes v' \neq v'\otimes v$ if $v\neq v'$. The fact goes back to the original Cartesian product set VxV where one has $(v,v') \neq (v',v)$ if $v\neq v'$ because (v,v') is an ordered tuplet, not a set {v,v'}. If V = W = R, one would not identify the point (x,y) with the point (y,x) in RxR = R^2 if $x \neq y$. Another word for commutative is abelian.

9. V⊗W is a vector space. The space V⊗W is a vector space whose vectors are linear combinations of v⊗w. We shall now verify this to be the case. We already know V⊗W is closed under addition since F(VxW) has this property. The + inverse of v⊗w is (-1)(v⊗w). Addition is commutative and associative. Any element of the form 0⊗w or v⊗0 can be taken as the identity for addition (the "zero") since, for example, using the first rule of (1.1.5),

$$0\otimes w = (v - v) \otimes w = (v\otimes w) + ((-v)\otimes w) = (v\otimes w) - (v\otimes w) = 0 \ \ (0 \text{ in the space } V\otimes W) \ . \qquad (1.1.9)$$

There is a scalar multiplicative identity since all fields K have an identity "1": 1(v⊗w) = (v⊗w). "Vector multiplication" is distributive over scalar addition (here the "vector" is v⊗w),

$$(s_1 + s_2)(v\otimes w) = [(s_1+s_2)v]\otimes w = [s_1v+s_2v]\otimes w = (s_1v)\otimes w + (s_2v)\otimes w = s_1(v\otimes w) + s_2(v\otimes w). \qquad (1.1.10)$$

Multiplication by a scalar is distributive over "vector addition" :

$$s\,(v_1\otimes w_2 + v_3\otimes w_4) = s\,(v_1\otimes w_2) + s\,(v_3\otimes w_4) \ . \qquad (1.1.11)$$

This property we more or less add by fiat to the earlier properties. It is the only reasonable way to do things since elements of V⊗W are linear combinations of pure elements of the form v⊗w.

10. Basis of V⊗W and general elements of V⊗W. In the above verification that V⊗W is a vector space, we used only pure vectors of V⊗W, but *general* vectors of V⊗W are linear combinations of the pure vectors so we really should rehash the above for general vectors. To do this, we first note that, since V and W are vector spaces, each has a basis, and we call these bases $\{u_i\}$ for V and $\{u'_i\}$ for W. It is not hard to show that the set of elements of the form $u_j\otimes u'_j$ forms a basis for V⊗W, so a general vector T in V⊗W can be expressed as

$$T = \Sigma_{ij}\, T^{ij}\, (u_i\otimes u'_j) \ . \qquad \text{// coefficients } T^{ij} \in \text{ field K} \qquad (1.1.12)$$

The inverse element -T is pretty obvious. Addition T + T' is commutative and T + T' + T" is associative. The zero element is the same. Vector multiplication is still distributive over scalar addition,

$$(s_1 + s_2)T = (s_1 + s_2)[\, \Sigma_{ij}\, T^{ij}\, (u_i\otimes u'_j)] = \Sigma_{ij}\, T^{ij}\, [\, (s_1 + s_2)\, (u_i\otimes u'_j)]$$

$$= \Sigma_{ij}\, T^{ij}\, [\, s_1\, (u_i\otimes u'_j) + s_2(u_i\otimes u'_j)] = s_1\, [\, \Sigma_{ij}\, T^{ij}\, (u_i\otimes u'_j)] + s_2\, [\, \Sigma_{ij}\, T^{ij}\, (u_i\otimes u'_j)]$$

$$= s_1\, T + s_2\, T \ . \qquad (1.1.13)$$

In this manner, all the required properties of a vector space can be verified for general elements of V⊗W.

11. Vector vs Tensor. Since V⊗W is a vector space, it is proper to refer to its elements v⊗w (or linear combinations of same) as "vectors". On the other hand, we shall refer to v⊗w as a "tensor" (a cross tensor) in the tensor product space V⊗W. In particular, it is a "rank-2 tensor" composed from v and w which are vectors in their respective vector spaces V and W. The word vector must be evaluated in its context. The subject of rank-2 tensors is developed more in Chapter 4.

12. Dimension of V⊗W. As noted above, the basis of the vector space V⊗W consists of elements of the form $u_i \otimes u'_j$. If the dimensions of V and W are n and n', then i takes n values, j takes n' values, and the dimension of the vector space V⊗W is n*n' (= nn'), the product of the separate vector space dimensions:

$$\dim(V\otimes W) = n*n' \qquad \text{where} \quad n = \dim(V) \text{ and } n' = \dim(W) \tag{1.1.14}$$

13. Generalization. The above development is easily generalized to the tensor product of any finite number of vector spaces. One first defines F(V,W,....Z) as linear combinations of elements of the Cartesian product space VxWx..xZ , which elements have the form (v,w,...z). One then defines a large set of equivalence relations analogous to those described above. One ends up with a large set of linear combinations which are all equivalent to 0, and this defines the equivalence class N. One then creates F(V,W,....Z)/N as the space of linear combinations where any pieces which are equivalent to 0 are filtered out. One then defines

$$V\otimes W\otimes ...\otimes Z \equiv F(VxWx...xZ)/N = \text{the tensor product of spaces V and W and... and Z.} \tag{1.1.15}$$

The tensor product space V⊗W⊗...⊗Z is the set of all linear combinations of elements (v,w,...z) of the Cartesian product space VxWx...xZ, written as v⊗w⊗...⊗z, where the following rules are *declared by fiat*:

$$(v_1+v_2)\otimes w\otimes .... \otimes z = v_1\otimes w\otimes .... \otimes z + v_2\otimes w\otimes .... \otimes z$$

$$v\otimes(w_1+w_2)\otimes .... \otimes z = v\otimes w_1\otimes .... \otimes z + v\otimes w_2\otimes .... \otimes z , \qquad \text{etc.}$$

and

$$s(v\otimes w\otimes ...\otimes z) = (sv)\otimes w\otimes ...\otimes z = v\otimes(sw)\otimes ...\otimes z , \quad \text{etc.} \qquad s \in K \tag{1.1.16}$$

When these rules are written for a function f(v,w,....z) one has,

$$f(v_1+v_2,w,...z) = f(v_1,w,...z) + f(v_2,w,...z)$$
$$f(v,w_1+w_2,...z) = f(v,w_1,...z) + f(v,w_2,...z), \quad \text{etc}$$

$$s\, f(v,w,...z) = f(sv,w...z) = f(v,sw,...), \text{ etc.} \qquad s \in K \tag{1.1.17}$$

If there are k factors in the tensor product V⊗W⊗...⊗Z, then the function f has k arguments, and a function obeying all of the above rules is said to be **k-multilinear**. For k = 2 we have bilinear, for k = 3 we have trilinear, and so on. One can mix in the scalar rule by saying for example

$$f(s_1v_1+s_2v_2, w, ...z) = s_1f(v_1,w,...z) + s_2f(v_2,w,...z)$$
$$f(v, s_1w_1+s_2w_2, ...z) = s_1f(v,w_1,...z) + s_2f(v,w_1,...z), \text{ etc.} \qquad (1.1.18)$$

We can then regard the set of ⊗ rules shown above as describing k-multilinearity for the tensor product space V⊗W⊗...⊗Z. Written in the second form,

$$(s_1v_1+s_2v_2)\otimes w\otimes ..... \otimes z = s_1 (v_1\otimes w\otimes .... \otimes z) + s_2 (v_2\otimes w\otimes .... \otimes z)$$
$$v \otimes (s_1w_1+s_2w_2)\otimes ...\otimes z = s_1 (v\otimes w_1\otimes .... \otimes z) + s_2 (v\otimes w_2\otimes .... \otimes z). \text{ etc.} \qquad (1.1.19)$$

An alternate approach to developing the tensor product of three or more vector spaces is to inductively build up by grouping things. For example

V⊗W⊗X = (V⊗W) ⊗ X = the tensor product of two vector spaces, one of which is V⊗W

V⊗W⊗X⊗Y = (V⊗W⊗X)⊗Y = the tensor product of two vector spaces, one of which is V⊗W⊗X

The results are the same with either approach.

1.2 The Tensor Product in Category Theory

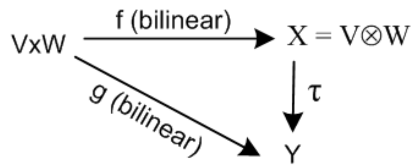
Category theory is an attempt to abstract the essence of algebraic structures which apply generally to objects like vector spaces, sets, rings, groups, modules and so on. One encounters certain category diagrams which must allow for flow through the diagram in all possible ways (the diagram must "commute"). A diagram consists of certain objects which are connected by arrows known as morphisms. For our application, these arrows are function mappings between spaces, and two sequential arrows in a path represent function composition in the sense f o g.

At a higher level, if the objects in the diagram are themselves categories, the morphism arrows are called functors. For example, for the category C of "all vector spaces over a field K" where the diagram arrows are linear maps, one can regard the equation V^2 = V⊗V as lying in the map CxC → C, and this map is then a functor, and the mapping is said to be functorial.

Category theory is a relatively recent addition to the mathematical house of many mansions. With precursor work done by Emily Noether (whose work shows up in a lot of places), category theory was developed in the early 1940's by Saunders Mac Lane (and others) who then summarized the theory in a text *Algebra* (1967) with coauthor Garrett Birkhoff. These same authors wrote the classic textbook *A Survey of Modern Algebra* (1941/1997) which is known to many students as "Birkhoff and Mac Lane".

We give here just an outline of this rather slippery tensor product development. It seems more of a fitting of our conclusions of Section 1.1 into category theory. The reader interested in more detail can look in *Algebra* or in Chapter 14 "Tensor Products" of Roman's text *Advanced Linear Algebra* (2008).

We start with the following triangle diagram (an example of a category diagram),



(1.2.1)

In this diagram VxW is the Cartesian product of two vector spaces V and W, exactly as in Section 1.1 above. Elements of VxW are (v,w). There are two mappings f: VxW → X and g: VxW → Y where X and Y are for the moment just spaces. They in turn are linked by a mapping traditionally called τ, so τ : X → Y. One says that "g can be factored through f".

The functions f and g are declared bilinear from the get-go. This is analogous to our declared equivalence relations in the approach of Section 1.1.

The set of all bilinear mappings f: VxW→X is called $\text{hom}_K(V,W; X)$ where K is the field of scalars. The letters hom stand for homomorphism ("same shape") which is a structure-preserving map. Linear maps (like τ discussed below) preserve vector space structure.

The triangle diagram must commute, so we must have g = τ o f (function composition).

The space X is our candidate space for the tensor product V⊗W space.

One needs to construct the function τ. To do so, use the fact that the diagram commutes to evaluate τ at the pure point v⊗w,

$$\tau(v\otimes w) = g(v,w). \tag{1.2.2}$$

Now "extend" τ so it applies to linear combinations of v⊗w elements by declaring that, for $s_i \in K$,

$$\tau\,(s_1\, v\otimes w + s_2\, v'\otimes w') = s_1\, \tau(v\otimes w) + s_1\, \tau(v'\otimes w') = s_1\, g(v,w) + s_2 g(v',w') \tag{1.2.3}$$

so τ is now a linear function τ: X→Y. It maps every element of X into an element of Y, and it is unique by its construction.

Once we have τ being a unique linear mapping, the "pair" (X, f:VxW→X) becomes a "universal pair". The idea here is that any alternate "pair" like (Y, g:VxW→Y) is equivalent to (X, f:VxW→X) up to the isomorphism implied by τ. In this sense, then, the mapping f:VxW → V⊗W is essentially unique -- it is "universal for bilinearity" -- so the tensor product mapping is well-defined. Function τ is called a mediating morphism, f is called the tensor map, and the elements of V⊗W are tensors.

From the top of the triangle one has

$$v\otimes w = f(v,w) \tag{1.2.4}$$

since f :VxW→X = V⊗W. Our "rules" of Section 1.1 for operator ⊗ now derive from the fact that f is a bilinear function:

$$(v_1+v_2) \otimes w = f(v_1+v_2,w) = f(v_1,w) + f(v_2,w) = (v_1\otimes w) + (v_2\otimes w)$$

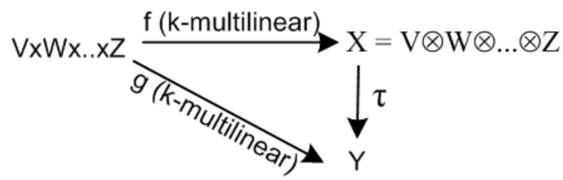
$$v \otimes (w_1+w_2) = f(v,w_1+w_2) = f(v,w_1) + f(v,w_2) = (v\otimes w_1) + (v\otimes w_2)$$

$$s(v\otimes w) = s\, f(v,w) = f(sv,w) = (sv)\otimes w \qquad s \in K$$

$$s(v\otimes w) = s\, f(v,w) = f(v,sw) = v\otimes(sw) \qquad s \in K \tag{1.2.5}$$

We then end up with the same space V⊗W and rules as in the previous quotient development, and we have extra assurance that V⊗W is a unique and well-defined object (it is universal).

The above scenario directly generalizes to the tensor product of k vector spaces with the following corresponding category diagram,



(1.2.6)

Lang (*Algebra*) for example shows the equivalent of this diagram on page 602 of his Chapter 16 (The Tensor Product). Note that Lang also wrote a different book *Linear Algebra*. The diagram above also appears in Roman's Chapter 14 on the Tensor Product, p 383.

2. A Review of Tensors in Covariant Notation

In Chapter 1 we generally avoided mentioning *components* of vectors and tensor products. But in many ways, "components" is what tensors are all about. Anyone who wants to use tensor analysis to actually do something practical is going to use tensor components. The whole notion of what it means to be a tensor of some rank requires components and component indices. Later when we deal with tensor functions, the components will morph into the vector arguments of multilinear functions.

Tensor analysis (algebra) provides some very heavy-duty machinery to handle manipulations of tensors and tensor components. A key idea is that a true tensor is something that transforms in a certain manner *relative to* some defined underlying transformation which below is called **x**' = **F**(**x**). In the following notes, we review this machinery.

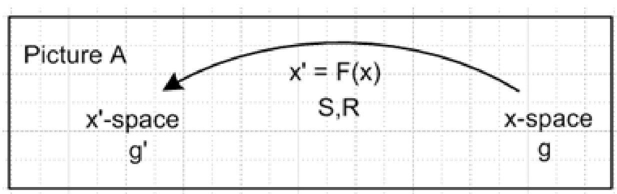
The review is based on our document on tensor analysis and curvilinear coordinates (*Tensor*, see Refs.) which follows the unusual path of developing tensor analysis in a "developmental notation" where all indices are down and covariant objects have overbars, then later this notation is converted to "standard notation" with the usual up and down indices. It is a large and complicated world, and below we report out only those facts which are useful for our efforts here.

Equation numbers referring to *Tensor* are followed by a prime ' .

Bolding Vectors. For the time being we shall display all vectors in bold font because we feel it helps the reader when dealing with covariant dot products and is compatible with *Tensor*. However, vector components are *not* bolded. Thus vector **V** will have components V^a and V_a. The exception is when vectors have extra labels, such as for the basis vectors $\mathbf{e}_n$. It's components are written $(\mathbf{e}_n)^a$ and $(\mathbf{e}_n)_a$. Eventually in Section 3.1 where we finally tie back to Sections 1.1 and 1.2 we shall quietly stop bolding vectors and will then be compatible with those earlier sections. Higher rank tensors are never bolded.

2.1 R, S and how tensors transform : Picture A

Tensor is in large part based on the following "picture",



(1.11)' (2.1.1)

Below we shall be thinking of x-space as a **vector space** V having a set of basis vectors $\{\mathbf{e}_i\}$ or $\{\mathbf{u}_i\}$. Then x'-space is a vector space V'. Below we shall use **V** as a prototype vector in space V, so $\mathbf{V} \in V$.

The two vector spaces V and V' have the same dimension N. In what follows, repeated indices are implicitly summed (**Einstein convention**) so for example $R^a{}_b V^b$ means $\Sigma_{b=1}{}^N R^a{}_b V^b$. Hanging indices like a in $R^a{}_b V^b$ can take any value in the range a = 1,2...N. The implied summation convention reduces symbol clutter especially when there are many summed indices in an equation. Sometimes however we will display sums for emphasis.

Figure (2.1.1) summarizes a generally non-linear transformation **x'** = **F**(**x**) between two spaces called x-space on the right (metric tensor g) and x'-space on the left (metric tensor g'). The coordinates of x-space are called **x**, and those of x'-space are called **x'**. Quantities in x'-space always have a prime, while those in x-space have no prime. A vector **V** in x-space has contravariant components V^a and covariant components V_a. The corresponding components V'^a and V'_a of **V'** in x'-space are these,

$$V'^a = R^a{}_b V^b \qquad R^a{}_b \equiv (\partial x'^a/\partial x^b) = \partial_b x'^a \qquad (7.5.3)' \ (7.5.2)'$$
$$V'_a = S^b{}_a V_b \qquad S^b{}_a \equiv (\partial x^b/\partial x'^a) = \partial'_a x^b \qquad (7.5.4)' \qquad (2.1.2)$$

As noted, primed equation numbers refer to *Tensor*.

The matrices R and S in (2.1.2) and Fig (2.1.1) are **linearizations** of the generally non-linear transformation **x'** = **F**(**x**) and **x** = $\mathbf{F}^{-1}(\mathbf{x'})$ in the close neighborhood of a selected point **x** in x-space and **x'** = **F**(**x**) in x'-space. Therefore, R and S are in general functions of **x**, though we suppress this dependence.

R is sometimes call **the differential** of the transformation **x'** = **F**(**x**). We call it the R matrix. Many texts don't make up a symbol like R for the differential, and so tensor equations such as (2.1.8) below are strewn with partial derivatives of the form $R^a{}_b = \frac{\partial x'^a}{\partial x^b}$. This is useful in doing chain rules, but otherwise obscures how the indices work. We settled on symbols R and S after rejecting various reasonable alternatives. One should understand that $R^a{}_b$ is in general *not* a simple rotation matrix despite the letter R.

From the chain rule, one can see that the matrices R and S are inverses of each other,

$$S^a{}_b \, R^b{}_c = \delta^a{}_c \, . \qquad // \ SR = 1 \qquad (7.6.1)' \qquad (2.1.3)$$

If desired, the matrix S can be eliminated from the discussion by the fact that (reflect indices in a vertical line between the indices)

$$S^a{}_b = R_b{}^a$$
$$S_a{}^b = R^b{}_a \, . \qquad (7.5.13)' \qquad (2.1.4)$$

Then (2.1.2) can be then be written with only R's ,

$$V'^a = R^a{}_b V^b \qquad R^a{}_b \equiv (\partial x'^a/\partial x^b) = \partial_b x'^a \qquad // \ \mathbf{V'} = R\mathbf{V}$$
$$V'_a = R_a{}^b V_b \qquad R_a{}^b \equiv (\partial x^b/\partial x'^a) = \partial'_a x^b \, . \qquad (2.1.5)$$

Here is a table summarizing different forms of the differentials R and S :

$$S^a{}_b = R_b{}^a \qquad = (\partial x^a/\partial x'^b) = (\partial x'_b/\partial x_a) \qquad = \partial'_b x^a = \partial^a x'_b$$
$$S_{ab} = R_{ba} \qquad = (\partial x_a/\partial x'^b) = (\partial x'_b/\partial x^a) \qquad = \partial'_b x_a = \partial_a x'_b$$
$$S^{ab} = R^{ba} \qquad = (\partial x^a/\partial x'_b) = (\partial x'^b/\partial x_a) \qquad = \partial'^b x^a = \partial^a x'^b$$
$$S_a{}^b = R^b{}_a \qquad = (\partial x_a/\partial x'_b) = (\partial x'^b/\partial x^a) \qquad = \partial'^b x_a = \partial_a x'^b \, . \qquad (7.5.16)' \qquad (2.1.6)$$

Notice how an upper index in a derivative denominator acts as a lower index and vice versa. The fact that each item can be represented by two partial derivatives follows from (2.1.4), (2.1.2) and the raising and lowering operations described in Section 2.2 below.

Vectors which transform under (according to) transformation **F** as in (2.1.5) are called rank-1 tensors. Here is how the four forms of a rank-2 tensor M transform under **F**,

$$
\begin{aligned}
M'^{ab} &= R^{a}{}_{a'}\ R^{b}{}_{b'}\ M^{a'b'} && \text{// pure contravariant} \\
M'^{a}{}_{b} &= R^{a}{}_{a'}\ R_{b}{}^{b'}\ M^{a'}{}_{b'} && \text{// mixed} \\
M'_{a}{}^{b} &= R_{a}{}^{a'}\ R^{b}{}_{b'}\ M_{a'}{}^{b'} && \text{// mixed} \\
M'_{ab} &= R_{a}{}^{a'}\ R_{b}{}^{b'}\ M_{a'b'}\ . && \text{// pure covariant} \qquad (7.5.8)' \qquad (2.1.7)
\end{aligned}
$$

We sometimes refer to $R^{a}{}_{a'}$ as the "down-tilt" R matrix, and $R_{a}{}^{a'}$ as the "up-tilt" R matrix. One sees that a down-tilt R transforms each contravariant (up) index, while an up-tilt R transforms each covariant (down) index.

From the above, one can intuit the way an arbitrary **rank-n tensor** transforms under **F**. For example

$$T'^{abc}{}_{de} = R^{a}{}_{a'}\ R^{b}{}_{b'}\ R^{c}{}_{c'}\ R_{d}{}^{d'}\ R_{e}{}^{e'}\ T^{a'b'c'}{}_{d'e'} \quad . \qquad (7.10.1)' \qquad (2.1.8)$$

The various forms of the R matrix have these four **orthogonality rules**,

$$1: R_{b}{}^{a}\ R^{b}{}_{c} = \delta^{a}{}_{c} \qquad 2: R^{b}{}_{a}\ R_{b}{}^{c} = \delta_{a}{}^{c} \qquad \text{// sum is on 1st index}$$

$$3: R^{a}{}_{b}\ R_{c}{}^{b} = \delta^{a}{}_{c} \qquad 4: R_{a}{}^{b}\ R^{c}{}_{b} = \delta_{a}{}^{c}\ . \qquad \text{// sum is on 2nd index} \quad (7.6.4)' \qquad (2.1.9)$$

These are just renditions of (2.1.3) with (2.1.4) (that is to say, SR = RS = 1).

Using these rules, one can show (proof below) that the inverses of the vector transforms shown above in (2.1.5) are these

$$
\begin{aligned}
V^{a} &= R_{b}{}^{a} V'^{b} && (7.6.7)' \\
V_{a} &= R^{b}{}_{a} V'_{b} && (7.6.8)' \qquad (2.1.10)
\end{aligned}
$$

where the summation index b on R is *not abutted* against the following vector.

The inversion of any tensor equation can be obtained instantly using the following simple rule:

Inversion Rule: For each R, reflect the indices in a vertical line between the indices. (2.1.11)

<u>Examples</u>: $V'^{n} = R^{n}{}_{m}\ V^{m}$, inversion is: $V^{n} = R_{m}{}^{n}\ V'^{m}$ $\qquad R^{n}|_{m} \rightarrow R_{m}{}^{n}$
$V'^{n} = R^{nm}\ V_{m}$, inversion is: $V^{n} = R^{mn}\ V'_{m}$ $\qquad R^{n}|^{m} \rightarrow R^{mn}$

Proof : (2.1.8)#1

$V'^n = R^n{}_m V^m \Rightarrow R_n{}^i V'^n = R_n{}^i R^n{}_m V^m = \delta^i{}_m V^m = V^i \Rightarrow V^i = R_n{}^i V'^n \Rightarrow V^n = R_m{}^n V'^m$.

The proof of the 2nd example follows from application of the tilt reversal rule (2.9.1) to the first example.

General Proof: Recall from (2.1.4) that $S^a{}_b = R_b{}^a$. The vertical line reflection $R_a|^b \rightarrow R^b{}_a = S^a{}_b$ just changes R into S, and S is the inverse of R as in (2.1.3). So **V**' = R**V** gives **V** = S**V**' and similarly for higher tensor cases.

Exercise: Invert equation (2.1.7) which says $M'^{ab} = R^a{}_{a'}\ R^b{}_{b'}\ M^{a'b'}$:

Result: $M^{ab} = R_{a'}{}^a\ R_{b'}{}^b\ M'^{a'b'}$.

The canonical vectors are the differential distances dx^i in x-space and dx'^i in x'-space (near some point **x** and corresponding **x**'). From (2.1.5) one then has,

$$dx'^a = R^a{}_b dx^b \qquad // \ d\mathbf{x}' = Rd\mathbf{x}$$
$$dx'_a = R_a{}^b dx_b \ . \qquad (2.1.12)$$

From inversion rule (2.1.11) the inverses of (2.1.12) are,

$$dx^a = R_b{}^a dx'^b$$
$$dx_a = R^b{}_a dx'_b \ . \qquad (2.1.13)$$

The derivative operator $\partial_a \equiv \partial/\partial x^a$ transforms like any other covariant vector. It is in fact the canonical covariant vector for the transformation **F**. Thus, from (2.1.5),

$$\partial'^a = R^a{}_b \partial^b \qquad R^a{}_b \equiv (\partial x'^a/\partial x^b) = \partial_b x'^a \qquad \partial^a = \partial/\partial x_a \qquad \partial'^a = \partial/\partial x'_a$$
$$\partial'_a = R_a{}^b \partial_b \qquad R_a{}^b \equiv (\partial x^b/\partial x'^a) = \partial'_a x^b \qquad \partial_a = \partial/\partial x^a \qquad \partial'^a = \partial/\partial x'_a \ . \qquad (2.1.14)$$

For example, if φ(x) is a scalar field (rank-0 tensor), $\partial_a\varphi(x)$ transforms as a covariant vector under F, and $\partial^a\varphi(x)$ transforms as a contravariant vector. Derivatives of tensors above rank-0 are more complicated, see Comment 1 below.

The R matrix is really four matrices, and we have seen two of its forms above. The R object is not a tensor because, as $R^a{}_b \equiv (\partial x'^a/\partial x^b)$ suggests, R has one foot in x'-space and one foot in x-space. In fact, the first index of R is raised or lowered by g', while the second index is raised or lowered by g:

$$R^a{}_b = (\partial x'^a/\partial x^b)$$
$$R^{ab} = R^a{}_{b'}\ g^{b'b} = (\partial x'^a/\partial x_b) \qquad // \text{ g pulls up the second index of } R^a{}_b$$
$$R_{ab} = g'_{aa'} R^{a'}{}_b = (\partial x'_a/\partial x^b) \qquad // \text{ g' pulls down the first index of } R^a{}_b$$
$$R_a{}^b = g'_{aa'} R^{a'}{}_{b'}\ g^{b'b} = (\partial x'_a/\partial x_b) \ . \qquad // \text{ both actions at once} \qquad (7.5.9)' \qquad (2.1.15)$$

Tensor fields

We have suppressed the fact that *in general* everything above is a function of **x** (or equivalently **x**'). For example, when we compute $R^a{}_b \equiv (\partial x'^a/\partial x^b)$ we generally obtain $R^a{}_b(\mathbf{x})$. The transformation of a vector from x-space to x'-space was given in (2.1.5) as $V'^a = R^a{}_bV^b$. For general $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ this really a statement about the transformation of vector *fields*: $V'^a(\mathbf{x}') = R^a{}_b(\mathbf{x})V^b(\mathbf{x})$. The rank-2 tensor transformation in (2.1.7) really says $M'^{ab}(\mathbf{x}') = R^a{}_{a'}(\mathbf{x})\ R^b{}_{b'}(\mathbf{x})\ M^{a'b'}(\mathbf{x})$ and we are transforming a rank-2 **tensor field**.

In special relativity it happens that $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is linear so $x'^a = F^a{}_bx^b$ (usually written with non-bold 4-vectors and Greek indices like $x'^\mu = \Lambda^\mu{}_\nu x^\nu$). In this situation $R^a{}_b$ does not depend on x, and one can then have vectors which are not fields like $p'^\mu = \Lambda^\mu{}_\nu p^\nu$ (momentum of a point particle) *and* vectors that are fields like $A'^\mu(\mathbf{x}') = \Lambda^\mu{}_\nu A^\nu(\mathbf{x})$ (electromagnetic vector potential). Notice on the x'-space side of the equation that the vector field A'^μ is a function of the x'-space coordinate x', while on the x-space side A^μ has argument x. In continuum mechanics and general relativity, $R^a{}_b$ *is* a function of **x** so everything is a field.

In the following **examples**, matrix R is the linearization of $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. In the neighborhood of the point **x** one has $d\mathbf{x}' = R(\mathbf{x})d\mathbf{x}$ as a "linear fit" to the generally non-linear $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. If $\mathbf{F}(\mathbf{x})$ is a linear transformation, then $\mathbf{x}' = F\mathbf{x}$ so $d\mathbf{x}' = Fd\mathbf{x}$ and then R = F = independent of **x**.

(a) $T'^{a'b'}{}_{c'd'} = R^{a'}{}_aR^{b'}{}_bR_{c'}{}^cR_{d'}{}^d\,T^{ab}{}_{cd}$ — tensor (rank 4, mixed)

(b) $T'^{a'b'}{}_{c'd'}(\mathbf{x}') = R^{a'}{}_a(\mathbf{x})R^{b'}{}_b(\mathbf{x})R_{c'}{}^c(\mathbf{x})R_{d'}{}^d(\mathbf{x})T^{ab}{}_{cd}$ — tensor which is not a field in x-space

(c) $T'^{a'b'}{}_{c'd'}(\mathbf{x}') = R^{a'}{}_aR^{b'}{}_bR_{c'}{}^cR_{d'}{}^d\,T^{ab}{}_{cd}(\mathbf{x})$ — tensor field, linear $\mathbf{F}(\mathbf{x})$

(d) $T'^{a'b'}{}_{c'd'}(\mathbf{x}') = R^{a'}{}_a(\mathbf{x})R^{b'}{}_b(\mathbf{x})R_{c'}{}^c(\mathbf{x})R_{d'}{}^d(\mathbf{x})\,T^{ab}{}_{cd}(\mathbf{x})$ — general tensor field (2.1.16)

Item (a) is the generic form we use for a transformation of a rank-4 tensor. If $\mathbf{F}(\mathbf{x})$ is linear (as in special relativity), then $R^a{}_b$ does not depend on position, and it is possible for $T'^{ab}{}_{cd}$ and $T^{ab}{}_{cd}$ not to be fields.

Item (b) is for a non-linear $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ where $T^{ab}{}_{cd}$ is not a field. Obviously $T'^{a'b'}{}_{c'd'}$ must depend therefore on **x**, and hence **x**', and so it *is* a field. In this unusual situation, the entire dependence of T' on **x**' is induced by the non-linearity of the transformation.

Item (c) is more standard, where the transformation is linear and a tensor field is being transformed. This is the case in special relativity.

Item (d) is the same thing, but the transformation is non-linear so one has $R^i{}_j(\mathbf{x})$.

Comments:

1. For situations where $R^a{}_b$ is a function of **x**, it is easy to see why there is trouble with derivatives. One need only consider :

$$V'^a(x') = R^a{}_b(x)V^b(x) \quad \text{and} \quad \partial'^b = R^b{}_c(x)\partial^c$$

so

$V'^{a,b}(x') \equiv \partial'^b V'^a(x')$ // $V^{a,b}$ is just a new notation for a derivative

$$= (R^b{}_c(x)\partial^c)[R^a{}_d(x)V^d(x)] = R^b{}_c(x)\, R^a{}_d(x)\, (\partial^c V^d(x)) + R^b{}_c(x)(\partial^c R^a{}_d(x))\, V^d(x)\ . \qquad (2.1.17)$$

It is this second term that causes $\partial^c V^d$ *not* to transform as a rank-2 tensor. It only transforms as a tensor if it happens that **x**' = F(**x**) is linear so $R^a{}_b$ is constant (as in special relativity). This problem is remedied by introducing the **covariant derivative** $V^{d;c}$ as discussed in *Tensor* Appendix F, see for example (F.9.5). This new object then properly transforms as a rank-2 tensor,

$$V'^{b;a}(\mathbf{x}') = R^b{}_c(\mathbf{x})\, R^a{}_d(\mathbf{x})V^{d;c}(\mathbf{x})\ . \qquad (2.1.18)$$

2. We have chosen to write $(\partial x'^a/\partial x^b)$ as $R^a{}_b$ as a space-saving notation. This object is often called "**the differential**" of the transformation **x**' = **F**(**x**) at point **x**. *Tensor* deals only with transformations where **x** and **x**' have the same number of components N, but the idea $(\partial x'^a/\partial x^b)$ as $R^a{}_b$ generalizes beyond this restriction. Of course then the matrix $R^a{}_b$ is no longer square and the associated linear algebra is more complicated. This situation arises in Chapter 10 below.

3 Although we have used the letter R in $R^a{}_b$, one should *not* think that R is a rotation. It *could be* a rotation, but in general it is more complicated, involving both rotation and stretching. It *could be* a rotation which, although being a rotation, is a different rotation at every point in space, like $R_y(\theta(\mathbf{x}))$.

2.2 The metric tensors g and g' and the dot product

Within each space (x-space and x'-space in the (2.1.1) Picture A), the metric tensor lowers or raises vector indices,

$$V_a = g_{ab}V^b \qquad V'_a = g'_{ab}V'^b$$
$$V^a = g^{ab}\, V_b \qquad V'^a = g'^{ab}\, V'_b \ . \qquad (7.4.4)' \text{ in std notation} \qquad (2.2.1)$$

In the same way, the metric tensor lowers or raises any index on any tensor.

The contravariant and covariant metric tensors are inverses of each other,

$$g_{ab}g^{bc} = g_a{}^c = \delta_a{}^c = \delta_{a,c} \qquad \text{// note that } \delta_i{}^j = \delta^i{}_j = \delta_{i,j}\ . \qquad (2.2.2)$$

Here the g_{ab} lowers the first index on g^{bc} to make $g_a{}^c$ which is $\delta_a{}^c = \delta_{a,c}$ so $g_{dn}g_{up} = 1$.

The objects g_{ab}, $g_a{}^b$, $g^b{}_a$ and g^{ab} are true tensor objects whereas $\delta_a{}^c$ and $\delta_{a,c}$ are not. It just happens that the value of $g_a{}^c$ is $\delta_a{}^c$. In writing covariant equations, one should replace $\delta_a{}^c$ by $g_a{}^c$ before attempting to raise index a or lower index c.

The metric tensor is a rank-2 tensor like any other rank-2 tensor, and so, looking at the first and last lines of (2.1.7),

$$g'^{ab} = R^a{}_{a'} \; R^b{}_{b'} \; g^{a'b'}$$
$$g'_{ab} = R_a{}^{a'} \; R_b{}^{b'} \; g_{a'b'} \quad . \qquad (7.5.6)' \qquad (2.2.3)$$

Any metric tensor is symmetric,

$$g^{ab} = g^{ba} \qquad\qquad g_{ab} = g_{ba}$$
$$g'^{ab} = g'^{ba} \qquad\qquad g'_{ab} = g'_{ba} \quad . \qquad (5.4.3)' \text{ in std notation} \qquad (2.2.4)$$

The metric tensor defines a (covariant) dot product in each space

$$\mathbf{a} \bullet \mathbf{b} = g_{ij}a^ib^j = g^{ij}a_ib_j = a_ib^i = a^ib_i \qquad \text{x-space}$$

$$\mathbf{a}\, '\bullet\, \mathbf{b}' = g'_{ij}a'^ib'^j = g'^{ij}a'_ib'_j = a'_ib'^i = a'^ib'_i \qquad \text{x'-space} \quad . \qquad (2.2.5)$$

The dot product is a scalar (rank-0 tensor) so it must be the same in either space

$$\mathbf{a}\, '\bullet\, \mathbf{b}' = \mathbf{a} \bullet \mathbf{b} \; . \qquad (2.2.6)$$

An exception to this rule is noted for fluid flow, see *Tensor* end of Section 5.2.

When applied to the canonical differential vector dx^i we find

$$d\mathbf{x} \bullet d\mathbf{x} = g_{ij}dx^idx^j = ||\, d\mathbf{x}\, ||^2 \equiv (ds)^2 \qquad \text{x-space}$$

$$d\mathbf{x}' \bullet d\mathbf{x}' = g'_{ij}dx'^idx'^j = ||\, d\mathbf{x}'\, ||^2 \equiv (ds')^2 \qquad \text{x'-space} \qquad (2.2.7)$$

Thus from (2.2.6) ds = ds' (invariant distance). In special relativity, ds is called the proper time dτ.

The metric tensor gets its name from these last equations. The distance between two points **x** and **y** in a metric space is determined by a function called "the metric" d(**x**,**y**). A commonly used metric is d(**x**,**y**) = ||**x**-**y**|| where the metric is defined by the norm. The distance between two close points **x** and **x**+d**x** is then given by d(**x**,**x**+**dx**) = ||**x**+**dx**-**x**|| = || d**x** || = ds. Squaring, d^2(**x**,**x**+**dx**) = || d**x** $||^2 = g_{ij}x^ix^j$ which shows how the metric tensor g_{ij} describes the squared metric d(**x**,**x**+**dx**) in the metric space of interest. One might recall that in a raw vector space, there is no distance concept d($\mathbf{v}_1$,$\mathbf{v}_2$). A vector space with a metric and an inner product, such as that shown above as • is then a Hilbert Space. It happens that the same metric tensor g_{ij} has two roles to play: it determines differential distance, and it lowers an index. See Chapter 5 of *Tensor* for more details.

As shown in (2.1.16) (d), for a general transformation everything (including g) is a function of **x**. For example,

$V_a(\mathbf{x}) = g_{ab}(\mathbf{x})V^b(\mathbf{x})$ $\quad\quad V'_a(\mathbf{x}') = g'_{ab}(\mathbf{x}')V'^b(\mathbf{x}')$ $\quad\quad$ *(2.2.1)*

$g'_{ab}(\mathbf{x}') = R_a{}^{a'}(\mathbf{x})\ R_b{}^{b'}(\mathbf{x})\ g_{a'b'}(\mathbf{x})$. $\quad\quad$ *(2.2.3)* $\quad\quad$ (2.2.8)

2.3 The basis vectors $\mathbf{e}_n$ and $\mathbf{e}^n$

There are two sets of basis vectors called $\mathbf{e}_n$ and $\mathbf{e}^n$ which exist in x-space (vector space V). The integer n is a **label**, not a component index. These basis vectors are defined as

$\mathbf{e}_n = \partial\mathbf{x}/\partial x'^n = \partial'_n\mathbf{x}$ $\quad\quad$ tangent base vectors

$\mathbf{e}^n = \partial\mathbf{x}/\partial x'_n = \partial'^n\mathbf{x}$. $\quad\quad$ reciprocal base vectors $\quad\quad$ (7.13.5)' $\quad\quad$ (2.3.1)

where $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. The **tangent base vectors $\mathbf{e}_n$** are tangent to the "coordinate lines" in x-space. If in x'-space a particular x'^n is allowed to vary while all other x'^i are held fixed, the locus in x'-space is a line parallel to the n axis, while the mapping of that line in x-space is the (often curved) **coordinate line** associated with x'_n. Then $\mathbf{e}_n = \partial\mathbf{x}/\partial x'^n$ evaluated at a point **x** on that coordinate line is a vector in x-space tangent to that coordinate line. For coordinate line examples, see e.g. *Tensor* (1.12)', (1.13)' and (3.2.8)',.

Meanwhile, the reciprocal vectors $\mathbf{e}^n$ are "dual to" the $\mathbf{e}_n$ in that $\mathbf{e}^n \bullet \mathbf{e}_m = \delta^n{}_m$. In fact we have,

$\mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm}$ $\quad\quad$ $g'^{in}\ \mathbf{e}_n = \mathbf{e}^i$

$\mathbf{e}^n \bullet \mathbf{e}_m = \delta^n{}_m = g'^n{}_m$

$\mathbf{e}^n \bullet \mathbf{e}^m = g'^{nm}$ $\quad\quad$ $g'_{in}\ \mathbf{e}^n = \mathbf{e}_i$ $\quad\quad$ (7.18.1)' $\quad\quad$ (2.3.2)

The equations on the left imply those on the right which show how to raise and lower basis vector labels. Although the $\mathbf{e}_n$ and $\mathbf{e}^n$ are vectors in x-space, it is the metric tensor of x'-space which raises and lowers.

For "duality" see text above *Tensor* (6.2.8)'. It is shown there that a unique dual basis $\mathbf{b}^n$ always exists for any given basis $\mathbf{b}_n$.

For a general transformation **F**, the tangent base vectors are functions of location and should be written $\mathbf{e}_n(\mathbf{x})$, and of course the same is true for $\mathbf{e}^n(\mathbf{x})$. Looking at (2.3.2), we see that $\mathbf{e}^n(\mathbf{x}) \bullet \mathbf{e}_m(\mathbf{x}) = \delta^n{}_m$ manages to be valid at every point in x-space. On the other hand, $g'_{nm}(\mathbf{x}) = \mathbf{e}_n(\mathbf{x}) \bullet \mathbf{e}_m(\mathbf{x})$ shows that the metric tensor is also a function of **x**.

<u>Example</u>: In polar coordinates (r,θ) one has $\mathbf{e}_r = \hat{\mathbf{r}}$ and $\mathbf{e}_\theta = r\,\hat{\boldsymbol{\theta}}$, both of which obviously depend on location in space. The metric tensor is $g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$ and also depends on spatial location through r. The coordinate lines for θ are circles whose tangents are $\mathbf{e}_\theta$, while those for r are rays whose tangents are $\mathbf{e}_r$.

These basis vectors are like any other vectors in x-space, and so they have contravariant and covariant components such as $(\mathbf{e}^n)^i$ and $(\mathbf{e}^n)_i$. Going back to our definition (2.3.1) we see that

$$\mathbf{e}_n \equiv \partial\mathbf{x}/\partial x'^n \qquad \Rightarrow \qquad (\mathbf{e}_n)^i = \partial x^i/\partial x'^n = R_n{}^i . \qquad \text{// from (2.1.5)} \qquad (2.3.3)$$

Looking at $(\mathbf{e}_n)^i = R_n{}^i$ we make these observations:

(1) the n label of $(\mathbf{e}_n)^i$ goes up and down with g' as shown in (2.3.2).
(2) the n index of $R_n{}^i$ also goes up and down with g' as shown in (2.1.15).
(3) the i index of $(\mathbf{e}_n)^i$ goes up and down with g as shown in (2.2.1).
(4) the i index of $R_n{}^i$ also goes up and down with g as shown in (2.1.15).

Therefore the equation $(\mathbf{e}_n)^i = R_n{}^i$ is "covariant", even though it is not a tensor equation (since R is not a tensor), so we can raise and lower indices at will on both sides. Therefore

$$(\mathbf{e}_n)^i = R_n{}^i \qquad (\mathbf{e}_n)_i = R_{ni} \qquad (\mathbf{e}^n)_i = R^n{}_i \qquad (\mathbf{e}^n)^i = R^{ni} \quad . \qquad (2.3.4)$$

The $\mathbf{e}_n$ and $\mathbf{e}^n$ also satisfy a "completeness relation",

$$(\mathbf{e}^n)_a(\mathbf{e}_n)^b = \delta_a{}^b \qquad (2.3.5)$$

where the implied sum is over the label n. From (2.3.4) this says $R^n{}_a\, R_n{}^b = \delta_a{}^b$ which in fact is just orthogonality rule #2 in (2.1.9). This completeness relation is *different* from the "orthogonality relation" $\mathbf{e}^n \bullet \mathbf{e}_m = \delta^n{}_m$ of (2.3.2) written as $(\mathbf{e}^n)_i\, (\mathbf{e}_m)^i = \delta^n{}_m$. Here the implied sum is over the component index i.

Recall from (2.3.2) that

$$\mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm} \quad . \qquad \textit{(2.3.2)}$$

There are two cases that are often of interest:

$\mathbf{e}_n \bullet \mathbf{e}_m = f_n\, \delta_{n,m}$ the $\{\mathbf{e}_n\}$ form an **orthogonal basis** for V

$\mathbf{e}_n \bullet \mathbf{e}_m = \delta_{n,m}$ the $\{\mathbf{e}_n\}$ form an **orthonormal basis** for V . (2.3.6)

Remember that the vectors $\mathbf{e}_n$ exist in x-space, despite the fact that the metric tensor g'_{nm} is for x'-space in our transformation Picture A.

2.4 The basis vectors $\mathbf{u}_n$ and $\mathbf{u}^n$

One can also define a set of "axis-aligned" basis vectors in x-space as follows

$$(\mathbf{u}_n)^i = \delta_n{}^i = g_n{}^i \qquad (\mathbf{u}^n)_i = \delta^n{}_i = g^n{}_i \qquad (7.18.3)' \qquad (2.4.1)$$

which can be compared with (2.3.4) for $\mathbf{e_n}$ and $\mathbf{e^n}$. Relations involving the u basis vectors are

$$\mathbf{u_n} \bullet \mathbf{u_m} = g_{nm} \qquad g^{in}\ \mathbf{u_n} = \mathbf{u^i}$$
$$\mathbf{u^n} \bullet \mathbf{u_m} = \delta^n{}_m = g^n{}_m$$
$$\mathbf{u^n} \bullet \mathbf{u^m} = g^{nm} \qquad g^{in}\ \mathbf{u_n} = \mathbf{u^i}\ . \qquad (7.18.3)' \qquad (2.4.2)$$

Notice the similarity to the relations (2.3.2) for the $\mathbf{e_n}$ and $\mathbf{e^n}$. The $\mathbf{u^n}$ are dual to the $\mathbf{u_n}$. Whereas the $\mathbf{e_n}$ and $\mathbf{e^n}$ involve the x'-space metric tensor g', the $\mathbf{u_n}$ and $\mathbf{u^n}$ involve the x-space metric tensor g.

We can easily calculate from (2.3.4) and (2.4.1) that

$$\mathbf{e^n} \bullet \mathbf{u_m} = (\mathbf{e^n})_i(\mathbf{u_m})^i = R^n{}_i\ \delta_m{}^i = R^n{}_m\ . \qquad (2.4.3)$$

According to (2.3.2), g' raises and lowers the label n on $\mathbf{e^n}$. According to (2.1.15), g' raises and lowers the first index n on $R^n{}_m$. Similarly, the label on $\mathbf{u_m}$ and the second index of $R^n{}_m$ are raised and lowered by g. Thus our equation (2.4.3) is "covariant" (even though it is not a true tensor equation), so we can at once write out all four forms of the dot products between the e and u basis vectors on the left below,

$$\mathbf{e^n} \bullet \mathbf{u_m} = R^n{}_m \qquad \Leftrightarrow \qquad \mathbf{e^n} = R^n{}_m\ \mathbf{u^m} = R^{nm}\ \mathbf{u_m}$$
$$\mathbf{e_n} \bullet \mathbf{u_m} = R_{nm} \qquad \Leftrightarrow \qquad \mathbf{e_n} = R_{nm}\ \mathbf{u^m} = R_n{}^m\ \mathbf{u_m}$$
$$\mathbf{e_n} \bullet \mathbf{u^m} = R_n{}^m \qquad \Leftrightarrow \qquad \mathbf{u^n} = R_m{}^n\ \mathbf{e^m} = R^{mn}\ \mathbf{e_m}$$
$$\mathbf{e^n} \bullet \mathbf{u^m} = R^{nm} \qquad \Leftrightarrow \qquad \mathbf{u_n} = R_{mn}\ \mathbf{e^m} = R^m{}_n\ \mathbf{e_m}\ . \qquad (2.4.4)$$

Each column implies the other. For example, for the third line,

$\Rightarrow$: $\mathbf{u^n} = \Sigma_m (\mathbf{e_m} \bullet \mathbf{u^n})\ \mathbf{e^m} = \Sigma_m\ R_m{}^n\ \mathbf{e^m}$ // $(\mathbf{e_m} \bullet \mathbf{u^n})$ is coefficient of the $\mathbf{e^m}$ expansion of $\mathbf{u^n}$

$\Leftarrow$: $\mathbf{u^n} = \Sigma_m\ R_m{}^n\ \mathbf{e^m} \quad \Rightarrow \quad \mathbf{e_k} \bullet \mathbf{u^n} = \Sigma_m R_m{}^n\ \mathbf{e_k} \bullet \mathbf{e^m} = \Sigma_m R_m{}^n\ \delta_k{}^m = R_k{}^n\ .$ (2.4.5)

The equations on the right of (2.4.4) show that the R matrix is the "basis change matrix" relating the two different bases e and u of x-space. These right-side equations are "vector sum equations" in which no component indices appear.

The basis $\mathbf{u_i}$ has a special place among possible bases for x-space. Above we call it an "axis-aligned" basis since $(\mathbf{u_n})^i = \delta_n{}^i$ so for example in R^2 we would have $\mathbf{u_1}^* = (1,0)$ and $\mathbf{u_2}^* = (0,1)$. Here the little asterisk is a notation to show we are talking about contravariant components. To say that $(\mathbf{u_n})^i = \delta_n{}^i$ does not say that x-space is Cartesian, since x-space could have any metric tensor g_{ij}.

What is really being said by the statement $(\mathbf{u_n})^i = \delta_n{}^i$ is that the *components* of vectors (and higher tensors) are being defined in a specific way.

If $\mathbf{v}$ is a vector, then v^i has the following meaning: $v^i = \mathbf{u^i} \bullet \mathbf{v}$. Another way to say this is that the v^i are the components of $\mathbf{v}$ when $\mathbf{v}$ is expanded in the u basis: $\mathbf{v} = \Sigma_i\ v^i\mathbf{u_i}$. In particular, since $\mathbf{u_n}$ is itself a vector, we have $\mathbf{u_n} = \Sigma_i\ (\mathbf{u_n})^i\mathbf{u_i} = \Sigma_i\ \delta_n{}^i\mathbf{u_i} = \mathbf{u_n}$. So the components of vectors are always defined *relative to* the u basis when u is selected as the "axis aligned basis" with $(\mathbf{u_n})^i = \delta_n{}^i$. The component indices on all tensors of Chapter 2 are referred to the u basis, examples being the i and j of the metric tensor g_{ij} and of any tensor M_{ij} or M_{ijk}.

We shall see below that one could expand vector $\mathbf{v}$ on the $\mathbf{e_n}$ basis, but then one gets $\mathbf{v} = \Sigma_i v'^i \mathbf{e_i}$ where $v'^i = \mathbf{e^i} \bullet \mathbf{v}$. The coefficient v'^i here is of course different from v^i since the bases are different. In fact $v'^i = R^i{}_j v^j$.

2.5 The basis vectors $\mathbf{e'_n}$ and $\mathbf{u'_n}$ and a summary

Basis vectors $\mathbf{e'_n}$ and $\mathbf{u'_n}$ are just mappings of $\mathbf{e_n}$ and $\mathbf{u_n}$ from x-space to x'-space,

$$\mathbf{u'_n} = R\,\mathbf{u_n} \qquad \mathbf{u_n} = S\,\mathbf{u'_n}$$
$$\mathbf{u'^n} = R\,\mathbf{u^n} \qquad \mathbf{u^n} = S\,\mathbf{u'^n}$$

$$\mathbf{e'_n} = R\,\mathbf{e_n} \qquad \mathbf{e_n} = S\,\mathbf{e'_n}$$
$$\mathbf{e'^n} = R\,\mathbf{e^n} \qquad \mathbf{e^n} = S\,\mathbf{e'^n} \quad . \tag{2.5.1}$$

Each of these mappings is like any vector mapping $\mathbf{V}' = R\mathbf{V}$. The primed basis vectors exist in x'-space, whereas the unprimed ones exist in x-space.

The components of these vectors are easily computed,

$$(\mathbf{u'_n})^i = R^i{}_j\,(\mathbf{u_n})^j = R^i{}_j\,\delta_n{}^j = R^i{}_n \qquad // \text{ (2.4.1)}$$

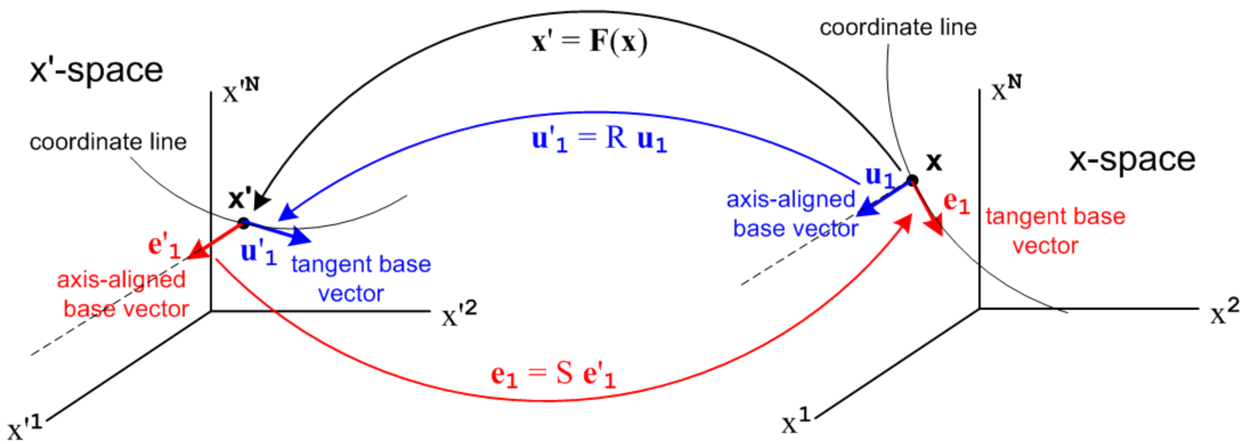
$$(\mathbf{e'_n})^i = R^i{}_j\,(\mathbf{e_n})^j = R^i{}_j\,R_n{}^j = \delta^i{}_n = g'^i{}_n \qquad // \text{ (2.3.4) and (2.1.9) \#3} \quad . \tag{2.5.2}$$

Whereas the $\mathbf{u_n}$ were the axis-aligned basis vectors in x-space, we see that the $\mathbf{e'_n}$ are the axis-aligned basis vectors in x'-space.

Since $(\mathbf{u'_n})^i = R^i{}_n$ from (2.5.2), and since $R^i{}_n = (\partial x'^i/\partial x^n) = \partial_n x'^i$ from (2.1.6), we can write $\mathbf{u'_n} = \partial_n \mathbf{x}'$ which is similar to the corresponding $\mathbf{e_n} = \partial'_n \mathbf{x}$ of (2.3.1) and shows that the $\mathbf{u'_n}$ really are tangent base vectors for the inverse transformation. We summarize :

$(\mathbf{u_n})^i = \delta_n{}^i$	axis-aligned basis vectors in x-space
$\mathbf{e_n} = \partial'_n \mathbf{x}$	tangent base vectors in x-space
$(\mathbf{e'_n})^i = \delta_n{}^i$	axis-aligned basis vectors in x'-space
$\mathbf{u'_n} = \partial_n \mathbf{x}'$	inverse tangent base vectors in x'-space . (2.5.3)

The situation is depicted in this drawing,



(2.5.4)

The figure shows just one basis vector of each type. Here red $\mathbf{e}_1$ is the tangent base vector for coordinate x'^1, whereas blue $\mathbf{u'}_1$ is the tangent base vector of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ for coordinate x^1. In each case a light black curve represents a piece of a coordinate line (curve) whose tangent is the tangent base vector.

The associated dot products are obtained from (2.4.2) and (2.3.2),

$$\begin{array}{lll}
\mathbf{u'}_n \bullet \mathbf{u'}_m = g_{nm} & // = \mathbf{u}_n \bullet \mathbf{u}_m & g^{in}\ \mathbf{u'}_n = \mathbf{u'}^i \\
\mathbf{u'}^n \bullet \mathbf{u'}_m = g^n{}_m & & \\
\mathbf{u'}^n \bullet \mathbf{u'}^m = g^{nm} & & g_{in}\ \mathbf{u'}^n = \mathbf{u'}_i
\end{array} \qquad (2.5.5)$$

$$\begin{array}{lll}
\mathbf{e'}_n \bullet \mathbf{e'}_m = g'_{nm} & // = \mathbf{e}_n \bullet \mathbf{e}_m & g'^{in}\ \mathbf{e'}_n = \mathbf{e'}^i \\
\mathbf{e'}^n \bullet \mathbf{e'}_m = g'^n{}_m & & \\
\mathbf{e'}^n \bullet \mathbf{e'}^m = g'^{nm} & & g'_{in}\ \mathbf{u'}^n = \mathbf{u'}_i
\end{array} \qquad (2.5.6)$$

and the basis vectors can therefore be raised and lowered as shown on the right by an appropriate metric tensor.

So far we have computed these basis vector components,

$$\begin{array}{llll}
(\mathbf{u}_n)^i = g_n{}^i = \delta_n{}^i & \textit{(2.4.1)} & (\mathbf{e}_n)^i = R_n{}^i & \textit{(2.3.4)} \\
(\mathbf{u'}_n)^i = R^i{}_n & \textit{(2.5.2)} & (\mathbf{e'}_n)^i = g'^i{}_n = \delta^i{}_n & \textit{(2.5.2)}\ .
\end{array} \qquad (2.5.7)$$

By raising n, lowering i, or doing both, one arrives at 12 more equations to get this full set of 16,

1 $(\mathbf{u}_n)^i = g_n{}^i = \delta_n{}^i$ $\qquad (\mathbf{e}_n)^i = R_n{}^i$
2 $(\mathbf{u'}_n)^i = R^i{}_n$ $\qquad (\mathbf{e'}_n)^i = g'^i{}_n = \delta^i{}_n$

3 $(\mathbf{u}^n)^i = g^{ni}$ $\qquad (\mathbf{e}^n)^i = R^{ni}$
4 $(\mathbf{u'}^n)^i = R^{in}$ $\qquad (\mathbf{e'}^n)^i = g'^{in}$

5 $(\mathbf{u}_n)_i = g_{ni}$ $\quad$ $(\mathbf{e}_n)_i = R_{ni}$
6 $(\mathbf{u}'_n)_i = R_{in}$ $\quad$ $(\mathbf{e}'_n)_i = g'_{in}$

7 $(\mathbf{u}^n)_i = g^n{}_i$ $\quad$ $(\mathbf{e}^n)_i = R^n{}_i$
8 $(\mathbf{u}'^n)_i = R_i{}^n$ $\quad$ $(\mathbf{e}'^n)_i = g'_i{}^n$. (2.5.8)

The reason one is allowed to do this follows from the label raising and lowering relations shown on the right side of (2.3.2), (2.4.2),(2.5.5), (2.5.6), and finally from (2.1.15) concerning indices on $R^i{}_j$.

Since $(\mathbf{u}'_n)^i = R^i{}_n$, the contravariant $\mathbf{u}'_n$ vectors are the columns of $R^*{}_*$.

Since $(\mathbf{e}^n)_i = R^n{}_i$, the covariant $\mathbf{e}^n$ vectors are the rows of $R^*{}_*$:

$$R^*{}_* = [\mathbf{u}'_1{}^*, \mathbf{u}'_2{}^*, \ldots \mathbf{u}'_N{}^*] = \begin{bmatrix} \mathbf{e}^1{}_* \\ \mathbf{e}^2{}_* \\ \ldots \\ \mathbf{e}^3{}_* \end{bmatrix} \qquad (2.5.9)$$

Using the fact that the dot product is a scalar, we can instantly obtain the following vector sum equations from (2.4.4),

$$\begin{aligned}
\mathbf{e}'^n \bullet \mathbf{u}'_m &= R^n{}_m &\Leftrightarrow&\quad \mathbf{e}'^n = R^n{}_m\, \mathbf{u}'^m = R^{nm}\, \mathbf{u}'_m \\
\mathbf{e}'_n \bullet \mathbf{u}'_m &= R_{nm} &\Leftrightarrow&\quad \mathbf{e}'_n = R_{nm}\, \mathbf{u}'^m = R_n{}^m\, \mathbf{u}'_m \\
\mathbf{e}'_n \bullet \mathbf{u}'^m &= R_n{}^m &\Leftrightarrow&\quad \mathbf{u}'^n = R_m{}^n\, \mathbf{e}'^m = R^{mn}\, \mathbf{e}'_m \\
\mathbf{e}'^n \bullet \mathbf{u}'^m &= R^{nm} &\Leftrightarrow&\quad \mathbf{u}'_n = R_{mn}\, \mathbf{e}'^m = R^m{}_n\, \mathbf{e}'_m \ .
\end{aligned} \qquad (2.5.10)$$

again showing that the R matrix acts as a basis-change matrix this time in x'-space.

Fact: One can treat the $\mathbf{e}_n$ as an *arbitrary* set of basis vectors . (2.5.11)

Above we started off by assuming some arbitrary transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ and then the $\mathbf{e}_n(\mathbf{x})$ are the tangent base vectors for this transformation $\mathbf{F}$.

From a different viewpoint, one can assume some arbitrary expressions for the $\mathbf{e}_n(\mathbf{x})$ and try to find a corresponding $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ for which those $\mathbf{e}_n(\mathbf{x})$ are the tangent base vectors. Given the functions $\mathbf{e}_n(\mathbf{x})$, one would know the matrix of functions $R_n{}^i(\mathbf{x})$ from (2.5.8) item 1. One could then attempt to integrate (2.1.12) which says $dx'_a = R_a{}^b(\mathbf{x})dx_b$ to find $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Let's assume this is all doable so $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ can always be found. From this point of view, one can regard the above equations concerning the $\mathbf{e}_n(\mathbf{x})$ to apply to an *arbitrary set* of basis functions $\mathbf{e}_n(\mathbf{x})$. Of course they have to be linearly independent at each value of $\mathbf{x}$. The next section provides a very simple example.

2.6 How to compute a viable x' = F(x) from a set of constant basis vectors $\mathbf{e}_n$

As a first step, select g = 1 so that x-space is the usual Cartesian space and $\mathbf{u}_n = \mathbf{u}^n$ are the usual orthonormal axis-aligned unit vectors of Cartesian space.

Suppose we are handed a set of constant-in-space basis vectors $\mathbf{e}_n$ specified by their components relative to the x-space axes. From (2.5.8) item 1,

$$(\mathbf{e}_m)^n = R_m{}^n \tag{2.6.1}$$

so we know the matrix $R_a{}^b$ (the $\mathbf{e}_m{}^*$ are the rows of the up-tilt R matrix).

What is the simplest way to fit this scenario into the tensor environment of Picture A in (2.1.1)?

We try a linear transformation of the form $\mathbf{x}' = \mathbf{F}(\mathbf{x}) = F\mathbf{x}$ where F is a constant matrix (independent of $\mathbf{x}$). Since $x'^a = F^a{}_b x^b$ we find $R^a{}_b \equiv (\partial x'^a/\partial x^b) = F^a{}_b$, and then of course $R_a{}^b = F_a{}^b$. So we have found a linear transformation F that works: $F_a{}^b = R_a{}^b$. Then the $\mathbf{e}_n$ are the tangent base vectors for this transformation F, and $\mathbf{e}^n$ are the reciprocal base vectors (the dual vectors of $\mathbf{e}_n$).

The metric tensor g'_{nm} can be computed from the dot product (2.3.2),

$$g'_{nm} = \mathbf{e}_n \bullet \mathbf{e}_m = (\mathbf{e}_n)^i (\mathbf{e}_m)_i = (\mathbf{e}_n)^i (\mathbf{e}_m)^i = R_n{}^i R_m{}^i . \tag{2.6.2}$$

Since g = 1, the second index on R is allowed to move up and down "for free".

This g'_{nm} can then be inverted to determine g'^{nm}. The reciprocal base vectors are then given by (2.3.2),

$$\mathbf{e}^n = g'^{nm}\, \mathbf{e}_m \tag{2.6.3}$$

with components

$$(\mathbf{e}^n)^i = g'^{nm} (\mathbf{e}_m)^i = g'^{nm} R_m{}^i . \tag{2.6.4}$$

We then rewrite the above equations as

$$(\mathbf{e}_m)_n = R_{mn} \qquad \text{// the } \mathbf{e}_m \text{ are the rows of matrix R, (2.6.1) lower n}$$

$$g'_{nm} = R_{ni}R_{mi} = R_{ni}R^T{}_{im} = (RR^T)_{nm} \qquad \Rightarrow \qquad g' = RR^T \qquad \text{// (2.6.2) lower i}$$

$$\mathbf{e}^n = g'^{nm}\, \mathbf{e}_m = h_{nm}\, \mathbf{e}_m \qquad \text{// where we define } h_{nm} \equiv g^{nm} \quad \text{// (2.3.2)}$$

$$(\mathbf{e}^n)_i = h_{nm}R_{mi} = (hR)_{ni}. \qquad \text{// the } \mathbf{e}^n \text{ are the rows of matrix (hR), (2.6.4)} \tag{2.6.5}$$

<u>Exercise</u>: You are handed these three constant vectors $\mathbf{e}_n$ in a 3-dimensional Cartesian x-space,

$\mathbf{e}_1 = (2,-1,3)$ // for example, $(\mathbf{e}_1)_2 = (\mathbf{e}_1)^2 = -1$
$\mathbf{e}_2 = (-1,2,4)$
$\mathbf{e}_3 = (1,3,2)$ // $\mathbf{e}_3 = 1\,\mathbf{u}_1 + 3\,\mathbf{u}_2 + 2\,\mathbf{u}_3$

so

$$R_{**} = \begin{pmatrix} 2 & -1 & 3 \\ -1 & 2 & 4 \\ 1 & 3 & 2 \end{pmatrix} \qquad // = F; \text{ "the } \mathbf{e_m} \text{ are the rows of matrix R" from (2.6.5)} \qquad (2.6.6)$$

Use Maple to compute g, h and then (hR). Here we are just implementing the equations in (2.6.5).

First enter the three $\mathbf{e_n}$ vectors and construct matrix R_{**},

```
e[1] := [2,-1,3];
```

$$e_1 := [2, -1, 3]$$

```
e[2] := [-1,2,4];
```

$$e_2 := [-1, 2, 4]$$

```
e[3] := [1,3,2];
```

$$e_3 := [1, 3, 2]$$

```
R := matrix(3,3,[e[1],e[2],e[3]]);
```

$$R := \begin{bmatrix} 2 & -1 & 3 \\ -1 & 2 & 4 \\ 1 & 3 & 2 \end{bmatrix}$$

Then compute the covariant metric tensor $g' = RR^T$,

```
gp := evalm(R &* transpose(R));
```

$$gp := \begin{bmatrix} 14 & 8 & 5 \\ 8 & 21 & 13 \\ 5 & 13 & 14 \end{bmatrix}$$

From this compute the contravariant metric tensor $h = g'^{-1}$ and then matrix (hR)

```
h := inverse(gp);
```

$$h := \begin{bmatrix} \frac{125}{1369} & \frac{-47}{1369} & \frac{-1}{1369} \\ \frac{-47}{1369} & \frac{171}{1369} & \frac{-142}{1369} \\ \frac{-1}{1369} & \frac{-142}{1369} & \frac{230}{1369} \end{bmatrix}$$

```
hR := evalm(h &* R);
```

$$hR := \begin{bmatrix} \frac{8}{37} & \frac{-6}{37} & \frac{5}{37} \\ \frac{-11}{37} & \frac{-1}{37} & \frac{7}{37} \\ \frac{10}{37} & \frac{11}{37} & \frac{-3}{37} \end{bmatrix}$$

The rows of (hR) are the vectors $\mathbf{e^n}$ (called $\mathbf{E_n}$ in the code),

```
for k from 1 to 3 do E[k] := row(hR,k) od;
```

$$E_1 := \left[\frac{8}{37}, \frac{-6}{37}, \frac{5}{37}\right]$$

$$E_2 := \left[\frac{-11}{37}, \frac{-1}{37}, \frac{7}{37}\right]$$

$$E_3 := \left[\frac{10}{37}, \frac{11}{37}, \frac{-3}{37}\right]$$

As a check, we verify that $\mathbf{e}^n \bullet \mathbf{e}_m = \delta^n{}_m$:

```
check := matrix(3,3):
for n from 1 to 3 do
  for m from 1 to 3 do
    check[n,m] := sum(E[n][j]*e[m][j],j=1..3);
  od
od;
print(check);
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.6.7}$$

2.7 Expansions of vectors onto basis vectors

Given vector **V** in x-space and the corresponding vector **V'** in x'-space, one may write the following expansions (*Tensor* (7.13.10,11)),

1	$\mathbf{V} = \Sigma_n V^n \mathbf{u}_n$	where $V^n = \mathbf{u}^n \bullet \mathbf{V}$	axis-aligned basis
2	$\mathbf{V} = \Sigma_n V_n \mathbf{u}^n$	where $V_n = \mathbf{u}_n \bullet \mathbf{V}$	axis-aligned basis
3	$\mathbf{V} = \Sigma_n V'^n \mathbf{e}_n$	where $V'^n = \mathbf{e}^n \bullet \mathbf{V}$	tangent base vector basis
4	$\mathbf{V} = \Sigma_n V'_n \mathbf{e}^n$	where $V'^n = \mathbf{e}_n \bullet \mathbf{V}$	tangent base vector basis
5	$\mathbf{V}' = \Sigma_n V'^n \mathbf{e}'_n$	where $V'^n = \mathbf{e}'^n \bullet \mathbf{V}'$	axis-aligned basis
6	$\mathbf{V}' = \Sigma_n V'_n \mathbf{e}'^n$	where $V'_n = \mathbf{e}'_n \bullet \mathbf{V}'$	axis-aligned basis
7	$\mathbf{V}' = \Sigma_n V^n \mathbf{u}'_n$	where $V^n = \mathbf{u}'^n \bullet \mathbf{V}'$	tangent base vector basis
8	$\mathbf{V}' = \Sigma_n V_n \mathbf{u}'^n$	where $V_n = \mathbf{u}'_n \bullet \mathbf{V}'$	tangent base vector basis . (2.7.1)

Notice that expansions 7,8 are obtained from 1,2 by applying the R matrix, since $\mathbf{V}' = R\mathbf{V}$ and $\mathbf{u}'_n = R\mathbf{u}_n$. Similarly, expansions 5,6 are obtained from 3,4

Any expansion can be directly verified by dotting the left column into an appropriate basis vector. For example, for expansion 5, using $\mathbf{e}'^m \bullet \mathbf{e}'_n = \delta^m{}_n$,

$$\mathbf{V}' = \Sigma_n V'^n \mathbf{e}'_n$$

$$\mathbf{e}'^m \bullet \mathbf{V}' = \mathbf{e}'^m \bullet (\Sigma_n V'^n \mathbf{e}'_n) = \Sigma_n V'^n \ \mathbf{e}'^m \bullet \mathbf{e}'_n = \Sigma_n V'^n \delta^m{}_n = V'^m \quad . \tag{2.7.2}$$

Notice that each set of coefficients appears twice on the right in (2.7.1), once for **V** and once for **V'**. This duplication arises because $\mathbf{a} \bullet \mathbf{b} = \mathbf{a}' \bullet \mathbf{b}'$ for any pair of vectors in either space. For example,

$$V^n = \mathbf{u}^n \bullet \mathbf{V} = \mathbf{u}'^n \bullet \mathbf{V}' \qquad \text{appears in lines 1 and 7 .} \tag{2.7.3}$$

Looking a bit ahead, we shall be extending the notion of a vector expansion to that of tensors of any rank, and the meaning of component indices is still governed by Fact (2.7.1). First, for expansions on the axis-aligned $\mathbf{u_n}$ we write,

$\mathbf{V} = \Sigma_n V^n \mathbf{u_n}$ rank-1 tensor, (2.7.1) line 1

$M = \Sigma_{nm} M^{nm} \mathbf{u_n}\otimes\mathbf{u_m}$ rank-2 tensor

$$T = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (\mathbf{u_{i_1}}\otimes \mathbf{u_{i_2}} \ldots\ldots\otimes \mathbf{u_{i_k}}) . \qquad \text{rank-k tensor} \qquad (2.7.4)$$

As discussed in (2.5.11) we may regard the basis vectors $\mathbf{e_n}$ as being an arbitrary basis. The primed tensor components are then the components of the x'-space version of the tensor under the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ generated by those arbitrary $\mathbf{e_n}$. The corresponding expansions are then,

$\mathbf{V} = \Sigma_n V'^n \mathbf{e_n}$ rank-1 tensor, (2.7.1) line 3

$M = \Sigma_{nm} M'^{nm} \mathbf{e_n}\otimes\mathbf{e_m}$ rank-2 tensor

$$T = \Sigma_{i_1 i_2 \ldots i_k} T'^{i_1 i_2 \ldots i_k} (\mathbf{e_{i_1}}\otimes \mathbf{e_{i_2}} \ldots\ldots\otimes \mathbf{e_{i_k}}) . \qquad \text{rank-k tensor} \qquad (2.7.5)$$

In the above, any pair of tilted matching indices can be tilted the other way. The meaning of the $\otimes$ symbol is discussed below, but we use it just momentarily here.

For an expansion on a mixed basis like $\mathbf{e_n}\otimes\mathbf{u_m}$, we will use the following notation

$$M = \Sigma_{nm} [M^{(e,u)}]^{nm} \mathbf{e_n}\otimes\mathbf{u_m} . \qquad (2.7.6)$$

Using this same notation we could write,

$$M = \Sigma_{nm} M'^{nm} \mathbf{e_n}\otimes\mathbf{e_m} = \Sigma_{nm} [M^{(e,e)}]^{nm} \mathbf{e_n}\otimes\mathbf{e_m} \equiv \Sigma_{nm} [M^{(e)}]^{nm} \mathbf{e_n}\otimes\mathbf{e_m} . \qquad (2.7.7)$$

<u>Exercise</u>**:** Consider these two expansions shown above of the rank-2 tensor M,

$$M = \Sigma_{ab} M^{ab} \mathbf{u_a}\otimes\mathbf{u_b} \qquad // (2.7.4) \qquad (2.7.8)$$

$$M = \Sigma_{ab} M'^{ab} \mathbf{e_a}\otimes\mathbf{e_b} . \qquad // (2.7.5) \qquad (2.7.9)$$

Verify that the coefficients M'^{ab} and M^{ab} are related as expected.

Use the result in (2.4.4) line 2 that $\mathbf{e_m} = R_m{}^i \mathbf{u_i}$ to get for (2.7.8),

$$M = M^{ab}\ \mathbf{u_a}\otimes\mathbf{u_b} \qquad\qquad // \text{ all implied sums}$$

$$= M^{ab}\ (R^i{}_a\ \mathbf{e_i}) \otimes (R^j{}_b\ \mathbf{e_j})$$

$$= (R^i{}_a R^j{}_b M^{ab})\ \mathbf{e_i}\otimes \mathbf{e_j}$$

so comparing to (2.7.9) one concludes that

$$M'^{ab} = R^i{}_a R^j{}_b M^{ab}$$

which is the correct statement that M transforms as a rank-2 tensor as shown in (2.1.7).

Confusion about vectors and scalars

The following issue is a subtle one that is worth nailing down early on because it can lead to confusions and seeming paradoxes. Consider this fact,

$$\mathbf{u}^n \bullet \mathbf{V} = \ \mathbf{u}'^n \bullet \mathbf{V}' \ . \qquad (2.7.10)$$

Being the dot product of two vectors, this object transforms as a scalar under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ as shown in (2.2.6). We might express this fact by writing

$$s^{(n)} = \mathbf{u}^n \bullet \mathbf{V} \qquad s'^{(n)} = \mathbf{u}'^n \bullet \mathbf{V}' \qquad s^{(n)} = s'^{(n)} \qquad n = 1,2....N\ . \qquad (2.7.11)$$

What we have here is a set of N scalars, $s^{(n)}$ for n = 1,2..N, where n is just a label. One would never claim that this set of scalars $s^{(n)}$ transforms as a vector, which would require that $s'^{(n)} = R^n{}_m\ s^{(m)}$. We don't have such a relation; what we have is $s'^{(n)} = s^{(n)}$.

Now it happens that $s^{(n)} = s'^{(n)} = V^n$ where V^n is the component of a vector. Notice that :

$V'^n = R^n{}_m V^m$	true : the V^n transform as a vector
$V'^n = V^n$	false
$s^{(n)} = s'^{(n)}$	true : each $s^{(n)} = \mathbf{u}^n \bullet \mathbf{V}$ transforms as a scalar
$s'^{(n)} = R^n{}_m\, s^{(m)}$	false (2.7.12)

We can summarize this discussion as follows:

Fact: Just because the scalars $s^{(n)} = \mathbf{u}^n \bullet \mathbf{V}$ take the *values* V^n does not mean that the scalars $s^{(n)}$ transform as vectors, nor does it mean that the vector components V^n transform as scalars. (2.7.13)

2.8 The Outer Product of Tensors and Use of ⊗

Consider two vectors which transform in the usual rank-1 tensor manner relative to some underlying transformation **x'** = **F(x)** (for which R is the linearization),

$$a'^i = \Sigma_j\ R^i{}_j\ a^j$$
$$b'^k = \Sigma_m\ R^k{}_m\ b^m \qquad \text{from (2.1.5)} \qquad (2.8.1)$$

where we temporarily show the summation symbols. Multiplying these equations together gives

$$(a'^i)(b'^k) = (\Sigma_j R^i{}_j a^j)(\Sigma_m R^k{}_m b^m) = \Sigma_{jm}\ R^i{}_j R^k{}_m (a^j b^m)$$

and then hiding the sums again,

$$(a'^i b'^k) = R^i{}_j R^k{}_m (a^j b^m)\ . \qquad (2.8.2)$$

Looking at the first line of (2.1.7), we see that this object is transforming as a rank-2 tensor, therefore it *is* a rank-2 tensor, and we can write it as

$$M'^{ik} = R^i{}_j R^k{}_m M^{jm} \qquad \text{where} \quad M^{ij} \equiv a^i b^j\ . \qquad (2.8.3)$$

The rank-2 tensor $M^{ij} = a^i b^j$ is said to be the **outer product** of two rank-1 tensors (vectors).

This idea can be generalized *ad infinitum*. For example, if K is a rank-2 tensor and **v** is a vector, then

$$M^{ijk} = K^{ij} v^k \qquad (2.8.4)$$

is a rank-3 tensor because it transforms as one using the same argument shown above. Next, consider,

$$M^{abcde} = K^{ab} K^{cd}\, v^e\ . \qquad (2.8.5)$$

If K is a rank-2 tensor and **v** is a rank-1 tensor, then M is a rank-5 tensor. Of course since this is a "true tensor equation" (a covariant one), indices may be shuffled any way one wants, such as

$$M^a{}_{bc}{}^d{}_e = K^a{}_b K_c{}^d\, v_e\ . \qquad (2.8.6)$$

Just imagine applying g_{**} several times to both sides of (2.8.5) to get (2.8.6).

There are so many possibilities for creating outer product tensors that one sometimes forgets that not all tensors can be "factored" into products of lower rank tensors.

<u>The ⊗ Symbol Appears</u> subtitle: "the rabbit goes into the hat"

In Sections 1.1 and 1.2 we had v⊗w being an element of a "tensor product space" V⊗W and we described two approaches to the development of the meaning of the symbol ⊗ : quotient space and category theory.

Here we provide a third approach to the meaning of ⊗ which is equivalent to that of the first two approaches. This third approach is geared to dealing with tensor components so there are lots of indices floating around, whereas in Sections 1.1 and 1.2 components were not even mentioned.

Recall our previous two equations

$$M^{abcde} = K^{ab}K^{cd}v^{e} \quad . \qquad \textit{(2.8.5)}$$
$$M^{a}{}_{bc}{}^{d}{}_{e} = K^{a}{}_{b}K_{c}{}^{d}v_{e} \quad . \qquad \textit{(2.8.6)}$$

In order to display the outer product as a unified entity, we had to make up a new symbol M to represent the outer product tensor. We can avoid having to do this by writing M = K⊗K⊗**v**, so that the ⊗ symbol in our "third approach" is just a way to *name* an outer product tensor. The above equations are then

$$(K\otimes K\otimes \mathbf{v})^{abcde} = K^{ab}K^{cd}v^{e} \quad . \qquad (2.8.7)$$
$$(K\otimes K\otimes \mathbf{v})^{a}{}_{bc}{}^{d}{}_{e} = K^{a}{}_{b}K_{c}{}^{d}v_{e} \quad . \qquad (2.8.8)$$

Here K and **v** are tensors, they are not spaces, so this is more like **v**⊗**w** than V⊗W. In fact, as a special case we can use this idea to *name* the outer product of two vectors to be rank-2 tensor **a**⊗**b**,

$$(\mathbf{a}\otimes\mathbf{b})^{ij} = a^{i}b^{j} \qquad \mathbf{a},\mathbf{b} \in V. \qquad (2.8.9)$$

Notice that ⊗ is a non-commuting operator: **a**⊗**b** ≠ **b**⊗**a** .

In our "third approach" the ⊗ symbol exists only within the context of V⊗V, since vectors **a** and **b** both belong to the x-space of Picture A (2.1.1) which we identify with vector space V. However, one can extend this meaning of ⊗ to apply more generally as the outer product of vectors in different vector spaces,

$$(\mathbf{v}\otimes\mathbf{w})^{ij} = v^{i}w^{j} \qquad \mathbf{v} \in V \quad \mathbf{w} \in W \,. \qquad (2.8.10)$$

If we try to fit this into our notion of tensor transformations, we would need two copies of Picture A, one for U→V and the other for X→W with vector transformations

$$v^{(V)i} = R^{(V)i}{}_{j}v^{(U)j} \qquad R^{(V)i}{}_{j} = \text{linearization of some transformation } \mathbf{x}' = F^{(V)}(\mathbf{x})$$
$$w^{(W)i} = R^{(W)i}{}_{j}w^{(X)j} \,. \qquad R^{(W)i}{}_{j} = \text{linearization of some transformation } \mathbf{y}' = F^{(W)}(y) \,. \qquad (2.8.11)$$

Then the transformation of the outer product "tensor" would be written as,

$$(v^{(V)i}w^{(W)i}) = R^{(V)i}{}_{a}R^{(W)j}{}_{b}\,(v^{(U)a}w^{(X)b})$$

or

$$[(\mathbf{v}\otimes\mathbf{w})^{(V,W)}]^{ij} = R^{(V)i}{}_{a}R^{(W)j}{}_{b}\,[(\mathbf{v}\otimes\mathbf{w})^{(U,X)}]^{ab} \,. \qquad U\otimes X \rightarrow V\otimes W \qquad (2.8.12)$$

One might refer to $(\mathbf{v}\otimes\mathbf{w})^{(V,W)}$ as a "cross-space rank-2 tensor" (**cross tensor**). Normally the word "tensor" is used when W = V. Then the above reads,

$$[(\mathbf{v}\otimes\mathbf{w})^{(V,V)}]^{ij} = R^{(V)i}{}_a R^{(V)j}{}_b [(\mathbf{v}\otimes\mathbf{w})^{(U,U)}]^{ab} \qquad U\otimes U \rightarrow V\otimes V$$

or

$$(\mathbf{v}\otimes\mathbf{w})'^{\,ij} = R^i{}_a R^j{}_b (\mathbf{v}\otimes\mathbf{w})^{ab} . \qquad \text{// Picture A (2.1.1)} \qquad (2.8.13)$$

The outer product thus has the same form in x'-space and in x-space, being a rank-2 tensor,

$$(\mathbf{a}\otimes\mathbf{b})^{ij} = a^i b^j \qquad \mathbf{a},\mathbf{b} \in V.$$

$$(\mathbf{a}\otimes\mathbf{b})'^{ij} = a'^i b'^j \qquad \mathbf{a'},\mathbf{b'} \in V' . \qquad (2.8.14)$$

As a final outer product example, consider the outer product of three vectors,

$$M^{ijk} = a^i b^j c^k \qquad \mathbf{a},\mathbf{b},\mathbf{c} \in V , \qquad (2.8.15)$$

Using our ⊗ naming method for outer products, this becomes

$$(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c})^{ijk} = a^i b^j c^k \qquad (2.8.16)$$

with this obvious extension to the outer product of any number of vectors

$$(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes....)^{ijk....} = a^i b^j c^k.... \qquad (2.8.17)$$

<u>Associativity of ⊗</u>

The outer product operator ⊗ as defined here is an **associative** operator, because multiplication of real numbers is associative. Consider for example,

$$(A\otimes B\otimes\mathbf{v})^{abcde} = A^{ab}B^{cd}v^e$$

(multiplication of reals is associative)

$$[(A\otimes B)\otimes\mathbf{v}]^{abcde} = [(A\otimes B)^{abcd}]\, v^e = [\, A^{ab}B^{cd}]\, v^e = A^{ab}B^{cd}v^e \ . \qquad (2.8.18)$$

Adding the parentheses on the second line in (A⊗B)⊗**v** does not alter the value of the components. This is true for the tensor product of any number of tensors,

$$(T_1\otimes T_2\otimes T_3....\otimes T_N)^{I_1 I_2 I_3 ... I_N} = T_1{}^{I_1} T_2{}^{I_2} T_3{}^{I_3} .... T_N{}^{I_N} \qquad (2.8.19)$$

where each I_i represents a set of indices to go with T_i. For example,

$$(T_1\otimes(T_2\otimes T_3)....\otimes T_N)^{I_1 I_2 I_3 ... I_N} = T_1{}^{I_1}(T_2\otimes T_3)^{I_2 I_3} ... T_N{}^{I_N}$$

$$= T_1{}^{I_1}[\, T_2{}^{I_2} T_3{}^{I_3}] \ ... T_N{}^{I_N} = T_1{}^{I_1} T_2{}^{I_2} T_3{}^{I_3} .... T_N{}^{I_N}$$

$$= (T_1\otimes T_2\otimes T_3....\otimes T_N)^{I_1 I_2 I_3 ... I_N} \ . \qquad (2.8.20)$$

Therefore we have,

Fact: The ⊗ operator is associative for any tensor product, so parentheses can be added anywhere in a tensor product. (2.8.21)

The associativity of the product of a set of real numbers along with the outer product definition of ⊗ is what causes the ⊗ operator to be associative. With the abstract ⊗ definitions of Chapter 1, associativity of ⊗ is added by fiat as an axiom.

2.9 The Inner Product (Contraction) of Tensors

When any tensor structure contains a pair of implicitly summed indices which are "tilted", one says that those indices are **contracted**. It is easy to show that, due to the orthogonality rules (2.1.9), such internal index contractions behave as a scalar, which is to say, behave as if they weren't there at all with respect to a transformation. A proof appears in *Tensor* (7.12.2). Such contractions in a tensor structure reduce the rank of the tensor by two, resulting in an **inner product**. The contracting sum must occur only on a "tilted pair" of indices.

Tilt Reversal Rule: Any such tilted index pair can have its tilt reversed "for free". (2.9.1)

Proof: Using (2.2.1) and (2.2.2),

$$[\text{-----}^{a}\text{---------}_{a}\text{----}] = g^{ab}\, g_{ac}\, [\text{-----}_{b}\text{---------}^{c}\text{----}] = g^{ba}\, g_{ac}\, [\text{-----}_{b}\text{---------}^{c}\text{----}]$$

$$= \delta^{b}{}_{c}\, [\text{-----}_{b}\text{---------}^{c}\text{----}] = [\text{-----}_{b}\text{---------}^{b}\text{----}] = [\text{-----}_{a}\text{---------}^{a}\text{----}]$$

where dashes indicate up or down tensor indices we don't care about. This "tilt reversal rule" applies to any contracted index within a tensor expression. It applies as well in other cases where g raises and lowers things so the above proof still works. The classic example involves expansions of the form (2.5.1)

$$\mathbf{V} = V'_{n}\, \mathbf{e}^{n} = V'^{n}\, \mathbf{e}_{n} \ . \qquad (2.9.2)$$

The tilt can be reversed even though the n on $\mathbf{e}^{n}$ is a label and not a tensor index. The reason is that

$$\mathbf{e}^{n} = g'^{ni}\, \mathbf{e}_{i} \qquad \mathbf{e}_{n} = g'_{ni}\, \mathbf{e}^{i} \qquad \textit{(2.3.2)}$$

$$V'_{n} = g'_{nb} V'^{b} \qquad V'^{n} = g'^{nb}\, V'_{b} \ . \qquad \textit{(2.2.1)}$$

The standard first example of an inner product is the inner product of two vectors. Consider,

$M^{ij} = a^{i}b^{j}$ = a rank-2 tensor, which we now contract to form,

$$s = M^{i}{}_{i} = a^{i}b_{i} = a_{i}b^{i} = \text{a rank-0 tensor (a scalar)} \ . \qquad (2.9.3)$$

Using our notation (2.2.5) this is written

$$s = \mathbf{a} \bullet \mathbf{b} \qquad // <\mathbf{a}\,|\,\mathbf{b}> \text{ in Dirac notation} \tag{2.9.4}$$

which is an "inner product" of two vectors. This is of course the inner product / scalar product / dot product which makes our vector space V be a Hilbert space.

In this example, creating an "inner product" of the two vectors a^i and b^j which has rank-0 goes in the opposite direction of the "outer product" that creates $M^{ij} = M^{ij} = a^i b^j$ of rank-2.

The term "contraction" is more often applied to reducing the rank of tensors than is "inner product", and perhaps it is best to reserve the term "inner product" for the above dot product of two vectors.

Here are other examples of rank reduction by contraction. Define

$$M^{abcd} \equiv K^{ab}Q^{cd} \quad = \text{rank-4 tensor} \tag{2.9.5}$$

$$T^{ac} \equiv M^{abc}{}_b \;=\; K^{ab}Q^c{}_b \;= \text{rank-2 tensor} . \tag{2.9.6}$$

In this last example, contraction on the b index happens to occur between the two rank-2 tensors from which M was constructed as an outer product. One more step,

$$S \equiv T^a{}_a = K^{ab}Q_{ab} \;= \text{rank-0 tensor (scalar)} . \tag{2.9.7}$$

Using the ⊗ notation introduced in the previous section, we can write the inner product s = **a** • **b** as a contraction of the outer product **a**⊗**b**

$$s = (\mathbf{a}\otimes\mathbf{b})^i{}_i \;=\; (\mathbf{a}\otimes\mathbf{b})_i{}^i \qquad \text{and} \qquad ||a||^2 \equiv a^i a_i \;=\; (\mathbf{a}\otimes\mathbf{a})^i{}_i . \tag{2.9.8}$$

Similarly (2.9.5,6,7) can be written

$$(K\otimes Q)^{abcd} = K^{ab}Q^{cd} \quad = \text{rank-4 tensor} \tag{2.9.9}$$

$$T^{ac} = (K\otimes Q)^{abc}{}_b \;= \text{rank-2 tensor} \tag{2.9.10}$$

$$S = T^a{}_a = (K\otimes Q)^{ab}{}_{ab} \;= \text{rank-0 tensor (scalar)} . \tag{2.9.11}$$

Dot products in spaces V⊗V, V⊗W, V⊗V⊗V and V⊗W⊗X

Recall that (2.2.5) defines the (covariant) dot product of two vectors in V

$$\mathbf{a} \bullet \mathbf{b} = g_{ij}a^i b^j = g^{ij}a_i b_j = a_i b^i = a^i b_i . \qquad \text{x-space} = V \qquad \textit{(2.2.5)}$$

It is possible to define an inner product operator • for use between two elements of V⊗V :

$$\begin{aligned}(\mathbf{a}\otimes\mathbf{b}) \bullet (\mathbf{c}\otimes\mathbf{d}) &\equiv \Sigma_{ij}(\mathbf{a}\otimes\mathbf{b})^{ij}(\mathbf{c}\otimes\mathbf{d})_{ij} \qquad (\mathbf{a}\otimes\mathbf{b}), (\mathbf{c}\otimes\mathbf{d}) \in V\otimes V \\ &= \Sigma_{ij}a^i b^j c_i d_j .\end{aligned} \tag{2.9.12}$$

With this definition, we have a tiny theorem:

Theorem: $(\mathbf{a}\otimes\mathbf{b}) \bullet (\mathbf{c}\otimes\mathbf{d}) = (\mathbf{a}\bullet\mathbf{c})(\mathbf{b}\bullet\mathbf{d})$ $\quad\quad$ $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d} \in V$ (2.9.13)

Proof: $(\mathbf{a}\bullet\mathbf{c})(\mathbf{b}\bullet\mathbf{d}) = (\Sigma_i\, a^i c_i)(\Sigma_j\, b^j d_j) = \Sigma_{ij}\, a^i c_i b^j d_j = \Sigma_{ij}\, a^i b^j\, c_i d_j$

$= \Sigma_{ij}(\mathbf{a}\otimes\mathbf{b})^{ij}(\mathbf{c}\otimes\mathbf{d})_{ij} = (\mathbf{a}\otimes\mathbf{b}) \bullet (\mathbf{c}\otimes\mathbf{d})$.

Suppose dim(V) = n and dim(W) = n'. Then we can extend the above theorem to V⊗W in this way. First define the dot product as,

$$(\mathbf{a}\otimes\mathbf{b'}) \bullet (\mathbf{c}\otimes\mathbf{d'}) \equiv \Sigma_{i=1}{}^{n}\Sigma_{j=1}{}^{n'}(\mathbf{a}\otimes\mathbf{b'})^{ij}(\mathbf{c}\otimes\mathbf{d'})_{ij} \quad\quad (\mathbf{v}\otimes\mathbf{w}),(\mathbf{v'}\otimes\mathbf{w'}) \in V\otimes W$$
$$= \Sigma_{ij} a^i b'^j c^i d'^j \quad . \quad\quad (2.9.14)$$

The corresponding Theorem is then

Theorem: $(\mathbf{a}\otimes\mathbf{b'}) \bullet (\mathbf{c}\otimes\mathbf{d'}) = (\mathbf{a} \bullet \mathbf{c})(\mathbf{b'} \bullet \mathbf{d'})$ $\quad\quad$ $\mathbf{a},\mathbf{c} \in V \quad \mathbf{b'},\mathbf{d'} \in W$ (2.9.15)

Proof: $(\mathbf{a}\bullet\mathbf{c})(\mathbf{b'}\bullet\mathbf{d'}) = (\Sigma_{i=1}{}^{n}\, a^i c_i)(\Sigma_{j=1}{}^{n}\, b'^j d'_j) = \Sigma_{ij} a^i c_i b'^j d'_j = \Sigma_{ij} a^i b'^j c_i d'_j$

$= \Sigma_{ij}\, (\mathbf{a}\otimes\mathbf{b'})^{ij}\, (\mathbf{c}\otimes\mathbf{d'})_{ij} = (\mathbf{a}\otimes\mathbf{b'}) \bullet (\mathbf{c}\otimes\mathbf{d'})$.

In a similar fashion one can show using (2.8.17) that with the following definition,

$$(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}) \bullet (\mathbf{d}\otimes\mathbf{e}\otimes\mathbf{f}) \equiv \Sigma_{ijk}\, (\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c})^{ijk}\, (\mathbf{d}\otimes\mathbf{e}\otimes\mathbf{f})_{ijk} \quad\quad V\otimes V\otimes V \quad\quad (2.9.16)$$

one obtains

Theorem: $(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}) \bullet (\mathbf{d}\otimes\mathbf{e}\otimes\mathbf{f}) = (\mathbf{a} \bullet \mathbf{d})(\mathbf{b} \bullet \mathbf{e})(\mathbf{c} \bullet \mathbf{f})$ $\quad\quad$ all vectors $\in V$ (2.9.17)

with a similar extension to V⊗W⊗X,

Theorem: $(\mathbf{a}\otimes\mathbf{b'}\otimes\mathbf{c''}) \bullet (\mathbf{d}\otimes\mathbf{e'}\otimes\mathbf{f''}) = (\mathbf{a} \bullet \mathbf{d})(\mathbf{b'} \bullet \mathbf{e'})(\mathbf{c''} \bullet \mathbf{f''})$ $\quad$ $\mathbf{a},\mathbf{d} \in V;\ \mathbf{b'},\mathbf{e'} \in W;\ \mathbf{c''},\mathbf{f''} \in X$
(2.9.18)

2.10 Tensor Expansions

Having a name for the outer product of two vectors allows us to write expansions of tensors of rank greater than 1 in a compact notation. The template is the vector expansion from (2.7.1) line 3,

$$\mathbf{V} = \Sigma_a V'^a \mathbf{e}_a \ . \tag{2.10.1}$$

(a) Rank-2 Tensor Expansion and Projection

As shown in (2.7.5) and (2.7.9), one can expand a rank-2 tensor on the tangent base vectors as follows,

$$M = \Sigma_{ab} M'^{ab} \mathbf{e}_a \otimes \mathbf{e}_b \ . \tag{2.10.2}$$

To verify that this is the correct expansion, we take the ij components of both sides,

$$[M]'^{ij} = [\Sigma_{ab} M'^{ab} \mathbf{e}_a \otimes \mathbf{e}_b]'^{ij}$$

$$= \Sigma_{ab} M'^{ab} (\mathbf{e}_a \otimes \mathbf{e}_b)'^{ij}$$

$$= \Sigma_{ab} M'^{ab} (\mathbf{e}'_a)^i (\mathbf{e}'_b)^j \qquad // \text{(2.8.14), outer product in x'-space}$$

$$= \Sigma_{ab} M'^{ab} \delta_a{}^i \delta_b{}^j \qquad // \text{(2.5.8) line 2 used twice}$$

$$= M'^{ij} \tag{2.10.3a}$$

so the expansion is correct. Similarly,

$$M = \Sigma_{ab} M^{ab} \mathbf{u}_a \otimes \mathbf{u}_b$$

$$[M]^{ij} = [\Sigma_{ab} M^{ab} \mathbf{u}_a \otimes \mathbf{u}_b]^{ij}$$

$$= \Sigma_{ab} M^{ab} (\mathbf{u}_a \otimes \mathbf{u}_b)^{ij}$$

$$= \Sigma_{ab} M^{ab} (\mathbf{u}_a)^i (\mathbf{u}_b)^j \qquad // \text{(2.8.14), outer product in x-space}$$

$$= \Sigma_{ab} M^{ab} \delta_a{}^i \delta_b{}^j \qquad // \text{(2.5.8) line 1 used twice}$$

$$= M^{ij} \ . \tag{2.10.3b}$$

A convenient notational method for projecting out the coefficients of any tensor expansion is the use of tensor-product-space dot products defined in Section 2.9. To demonstrate, we use a tensor expansion in V⊗W where the basis vectors are $\mathbf{e_n}$ and $\mathbf{e'_n}$ for V and W, using notation of (2.7.6),

$$M = \Sigma_{ab} [M^{(e,e')}]^{ab}\ \mathbf{e_a}\otimes\mathbf{e'_b} . \qquad\qquad M \in V\otimes W . \qquad\qquad (2.10.4)$$

The appropriate projector is $(\mathbf{e^i}\otimes\mathbf{e'^j})$, which is just the expansion's basis $\mathbf{e_a}\otimes\mathbf{e'_b}$ with up/down toggled on the indices, and dummy labels like i,j selected. Using this projector one finds,

$$(\mathbf{e^i}\otimes\mathbf{e'^j}) \bullet M = \Sigma_{ab}\ [M^{(e,e')}]^{ab}\ (\mathbf{e^i}\otimes\mathbf{e'^j}) \bullet (\mathbf{e_a}\otimes\mathbf{e'_b})$$

$$= \Sigma_{ab}\ [M^{(e,e')}]^{ab}\ (\mathbf{e^i} \bullet \mathbf{e_a})(\mathbf{e'^j} \bullet \mathbf{e'_b}) \qquad // \text{ theorem (2.9.15)}$$

$$= \Sigma_{ab}\ [M^{(e,e')}]^{ab}\ \delta^i{}_a\, \delta^j{}_b \qquad // \text{ dual pairs as in (2.3.2)}$$

$$= [M^{(e,e')}]^{ij} \qquad\qquad (2.10.5)$$

and indeed, the coefficient is duly projected out of M. Here is a more complicated example where M is now a rank-3 tensor, and where we use a perverse mixed basis,

$$M = \Sigma_{abc} [M^{(e,u',e")}]_{ab}{}^c\ \mathbf{e^a} \otimes \mathbf{u'_b} \otimes \mathbf{e"_c} \qquad\qquad M \in V\otimes W\otimes X. \qquad\qquad (2.10.6)$$

The projector is $(\mathbf{e_i} \otimes \mathbf{u'^j} \otimes \mathbf{e"^k})$ and we use it to project out the coefficient in (2.10.6) :

$$(\mathbf{e_i} \otimes \mathbf{u'^j} \otimes \mathbf{e"^k}) \bullet M = (\mathbf{e_i} \otimes \mathbf{u'^j} \otimes \mathbf{e"^k}) \bullet \Sigma_{ab} [M^{(e,u',e")}]_{ab}{}^c\ \mathbf{e^a} \otimes \mathbf{u'_b} \otimes \mathbf{e"_c}$$

$$= \Sigma_{ab} [M^{(e,u',e")}]_{ab}{}^c\ (\mathbf{e_i} \otimes \mathbf{u'^j} \otimes \mathbf{e"^k}) \bullet (\mathbf{e^a} \otimes \mathbf{u'_b} \otimes \mathbf{e"_c})$$

$$= \Sigma_{ab} [M^{(e,u',e")}]_{ab}{}^c\ (\mathbf{e_i} \bullet \mathbf{e^a})(\mathbf{u'^j} \bullet \mathbf{u'_b})(\mathbf{e"^k} \bullet \mathbf{e"_c}) \qquad // \text{ theorem (2.9.18)}$$

$$= \Sigma_{ab} [M^{(e,u',e")}]_{ab}{}^c\ \delta_i{}^a\, \delta^j{}_b\, \delta^k{}_c \qquad // \text{ each pair is dual as in (2.3.2)}$$

$$= [M^{(e,u',e")}]_{ij}{}^k \quad . \qquad\qquad (2.10.7)$$

(b) Rank-k Tensor Expansions and Projections

A rank-k tensor T in V^k has this expansion on the $\mathbf{e_r}$ basis,

$$T = \Sigma_{i_1 i_2 \ldots i_k}\ T'^{i_1 i_2 \ldots i_k} (\mathbf{e_{i_1}} \otimes \mathbf{e_{i_2}} \ldots\ldots \otimes \mathbf{e_{i_k}}) . \qquad\qquad (2.10.8)$$

To verify, we take components of both sides,

$$[T]^{j_1 j_2 ... j_k} = \Sigma_{i_1 i_2 .... i_k} \ T'^{i_1 i_2 .... i_k} (\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} ..... \otimes \mathbf{e}_{i_k})^{j_1 j_2 ... j_k}$$

$$= \Sigma_{i_1 i_2 .... i_k} \ T'^{i_1 i_2 .... i_k} (\mathbf{e}_{i_1})^{j_1} (\mathbf{e}_{i_2})^{j_2} .....(\mathbf{e}_{i_k})^{j_k} \qquad // (2.8.17)$$

$$= \Sigma_{i_1 i_2 .... i_k} \ T'^{i_1 i_2 .... i_k} R_{i_1}{}^{j_1} R_{i_2}{}^{j_2} .....R_{i_k}{}^{j_k} \qquad // (2.5.8) \text{ line 1}$$

$$= \Sigma_{i_1 i_2 .... i_k} \ [\, R_{i_1}{}^{j_1} R_{i_2}{}^{j_2} .....R_{i_k}{}^{j_k} \ T'^{i_1 i_2 .... i_k} \,]$$

$$= T^{j_1 j_2 ... j_k} \ . \qquad (2.10.9)$$

To get the last step, we use the inversion rule (2.1.11) applied to the known tensor transformation

$$T'^{j_1 j_2 ... j_k} = R^{j_1}{}_{i_1} R^{j_2}{}_{i_2} .....R^{j_k}{}_{i_k} \ T^{i_1 i_2 .... i_k} \quad . \qquad (2.10.10)$$

The coefficients $T'^{i_1 i_2 .... i_k}$ can be projected out from T as in (2.10.5),

$$(\mathbf{e}^{i_1} \otimes \mathbf{e}^{i_2} \otimes ... \otimes \mathbf{e}^{i_k}) \bullet T = T'^{i_1 i_2 ... i_k} \qquad (2.10.11)$$

with an appropriate generalization of the dot product • to the space $V^k = V \otimes V ... \otimes V$,

$$(\mathbf{v}_1 \otimes \mathbf{v}_2 ... \otimes \mathbf{v}_k) \bullet (\mathbf{w}_1 \otimes \mathbf{w}_2 ... \otimes \mathbf{w}_k)$$

$$\equiv \Sigma_{i_1 i_2 .... i_k} \ (\mathbf{v}_1 \otimes \mathbf{v}_2 ... \otimes \mathbf{v}_k)^{i_1 i_2 .... i_k} \ (\mathbf{w}_1 \otimes \mathbf{w}_2 ... \otimes \mathbf{w}_k)_{i_1 i_2 .... i_k}$$

$$= \Sigma_{i_1 i_2 .... i_k} \ (\mathbf{v}_1)^{i_1} (\mathbf{v}_2)^{i_2} ... (\mathbf{v}_k)^{i_k} \ \ (\mathbf{w}_1)_{i_1} (\mathbf{w}_2)_{i_2} ... (\mathbf{w}_k)_{i_k} \qquad // \text{ outer products}$$

$$= (\mathbf{v}_1 \bullet \mathbf{w}_1) (\mathbf{v}_2 \bullet \mathbf{w}_2) .... (\mathbf{v}_k \bullet \mathbf{w}_k) \ . \qquad (2.10.12)$$

Using the notion of a multiindex I (an *ordinary* multiindex),

$$I \equiv i_1, i_2, ...i_k \qquad // \text{ each } i_r \text{ ranges } 1,2....n \qquad n = \dim(V) \qquad (2.10.13)$$

and a shorthand notation for the basis vectors

$$e_I \equiv \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} ..... \otimes \mathbf{e}_{i_k} \qquad e^I \equiv \mathbf{e}^{i_1} \otimes \mathbf{e}^{i_2} ..... \otimes \mathbf{e}^{i_k} \qquad (2.10.14)$$

the expansion (2.10.8) can be stated in the following compact form,

$$T = \Sigma_I \, T'^I \, e_I \qquad \textit{(2.10.8)} \qquad (2.10.15)$$

and the coefficients T'^I can be projected out according to (2.10.11),

$$e^I \bullet T = T'^I \ . \qquad \textit{(2.10.11)} \qquad (2.10.16)$$

With no comments, we now repeat the above set of steps for the expansion of T on the $\mathbf{u_r}$ basis:

$$T = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (\mathbf{u}_{i_1} \otimes \mathbf{u}_{i_2} \ldots\ldots \otimes \mathbf{u}_{i_k}) \qquad (2.10.17)$$

$$[T]^{j_1 j_2 \ldots j_k} = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (\mathbf{u}_{i_1} \otimes \mathbf{u}_{i_2} \ldots\ldots \otimes \mathbf{u}_{i_k})^{j_1 j_2 \ldots j_k}$$

$$= \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (\mathbf{u}_{i_1})^{j_1} (\mathbf{u}_{i_2})^{j_2} \ldots\ldots (\mathbf{u}_{i_k})^{j_k} \qquad // (2.8.17)$$

$$= \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} \delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} \ldots\ldots \delta_{i_k}{}^{j_k} \qquad // (2.5.8) \text{ line 1}$$

$$= T^{j_1 j_2 \ldots j_k} \qquad (2.10.18)$$

$$(\mathbf{u}^{i_1} \otimes \mathbf{u}^{i_2} \otimes \ldots \otimes \mathbf{u}^{i_k}) \bullet T = T^{i_1 i_2 \ldots i_k} \qquad (2.10.19)$$

$$u_I \equiv \mathbf{u}_{i_1} \otimes \mathbf{u}_{i_2} \ldots\ldots \otimes \mathbf{u}_{i_k} \qquad u^I \equiv \mathbf{u}^{i_1} \otimes \mathbf{u}^{i_2} \ldots\ldots \otimes \mathbf{u}^{i_k} \qquad (2.10.20)$$

$$T = \Sigma_I T^I u_I \qquad (2.10.21)$$

$$u^I \bullet T = T^I \ . \qquad (2.10.22)$$

2.11 Dual Spaces and Tensor Functions

We denote dual-space vectors and tensors by Greek or script font letters.

The dual space V* is by definition the space of **linear functionals over V**. If $\alpha \in V^*$, we can then write

$$\alpha : V \rightarrow K \qquad \alpha(\mathbf{v}) = k \in K \qquad (2.11.1)$$

where K is any field (but we always use the reals). Since α is a linear *functional*, α(**v**) is a linear *function*.

In normal calculus, if f: V → R, one refers to f as a function, and f(**v**) as that function evaluated at some point in V, though loosely speaking f(**v**) is also called a function. To emphasize the distinction, we shall refer to f as a "functional" and f(**v**) as a "function".

Comments: Much of the rest of this section will be repeated in later Chapters. We have found that the notations involved can be a *major* stumbling block, and feel it is important to exercise the notation in many ways to make the reader (and author) feel comfortable with it. As with most endeavors, it is a matter of practice. We also try to explain *why* certain notations are used.

(a) The Dual Space V* in Matrix and Dirac Notation

For every column vector $\mathbf{v}$ in V, there exists a row vector $\mathbf{v}^{\mathbf{T}}$ such that $(\mathbf{v}^{\mathbf{T}})_i = v_i$. For example, for N=2,

$$\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix} = |\mathbf{v}> \qquad\qquad \mathbf{v}^{\mathbf{T}} = (a, b) = <\mathbf{v}| \quad . \tag{2.11.a.1}$$

Here we have snuck in the Dirac bra-ket notation where the ket $|\mathbf{v}>$ is a column vector and the bra $<\mathbf{v}|$ is the corresponding row vector. The notation $\mathbf{v}^{\mathbf{T}}$ means that the row vector is the Transpose of the column vector.

We now have multiple ways to write the dot (inner, scalar) products of Section 2.9 :

$$\mathbf{v} \bullet \mathbf{v}' = \mathbf{v}^{\mathbf{T}}\mathbf{v}' = (a, b)\begin{pmatrix} a' \\ b' \end{pmatrix} = aa'+bb' = <\mathbf{v}\,|\,\mathbf{v}'> \quad . \tag{2.11.a.2}$$

Because our vectors have real components, the above can also be written

$$\mathbf{v}' \bullet \mathbf{v} = \mathbf{v}'^{\mathbf{T}}\mathbf{v} = (a', b')\begin{pmatrix} a \\ b \end{pmatrix} = aa'+bb' = <\mathbf{v}'\,|\,\mathbf{v}> \quad . \tag{2.11.a.3}$$

We mention real components only because in the Dirac notation one has $<\mathbf{a}|\mathbf{b}> = <\mathbf{b}|\mathbf{a}>^*$ where * means complex conjugation, so if this scalar product is real, then $<\mathbf{a}|\mathbf{b}> = <\mathbf{b}|\mathbf{a}>$.

We regard $\mathbf{v}$ or $|\mathbf{v}>$ as being a vector in the vector space V, while $\mathbf{v}^{\mathbf{T}}$ or $<\mathbf{v}|$ (the row vector) is a vector in the **dual space** V*. This is really a simple concept. Sometimes the dual-space vector $\mathbf{v}^{\mathbf{T}} = <\mathbf{v}|$ is referred to as the **covector** of $\mathbf{v} = |\,\mathbf{v}>$.

Suppose $\boldsymbol{\alpha}^{\mathbf{T}} = <\boldsymbol{\alpha}\,|$ is a vector in the dual space V* . One can regard this dual-space vector $<\boldsymbol{\alpha}\,|$ as being a functional which acts on vectors in the space V. Then,

$$\alpha \;=\; <\boldsymbol{\alpha}\,| \;=\text{functional}$$

$$\alpha(\mathbf{v}) = \;<\boldsymbol{\alpha}\,|\,\mathbf{v}> = \boldsymbol{\alpha} \bullet \mathbf{v} = \text{function} = \text{a scalar number} \tag{2.11.a.4}$$

$$\alpha : V \rightarrow K . \qquad\qquad // \text{ K = any real field, such as the real numbers}$$

Just as the space of column vectors V is a linear space (a vector space), so also the dual space of row vectors V* is a linear space, so we know that the functional $<\boldsymbol{\alpha}\,|$ is a linear functional. That in turn implies that the function $\alpha(\mathbf{v})$ is a linear function, which we now show directly:

$$\alpha(s\mathbf{v}) = \boldsymbol{\alpha} \bullet (s\mathbf{v}) = s\,(\,\boldsymbol{\alpha} \bullet \mathbf{v}\,) = s\,\alpha(\mathbf{v})$$

$$\alpha(\mathbf{v}+\mathbf{v}') = \boldsymbol{\alpha} \bullet (\mathbf{v}+\mathbf{v}') = \boldsymbol{\alpha} \bullet \mathbf{v} + \boldsymbol{\alpha} \bullet \mathbf{v}' = \alpha(\mathbf{v}) + \alpha(\mathbf{v}') \;. \tag{2.11.a.5}$$

(b) Functional notation

We have now a slight notational conundrum. We like to write a scalar-valued function F($\mathbf{v}$) in non-bold font, whereas a vector-valued function would be $\mathbf{F}(\mathbf{v})$. Thus we have written $\alpha(\mathbf{v})$ above with a non-bold α, since $\alpha(\mathbf{v})$ is a scalar-valued function. On the other hand, $\alpha(\mathbf{v})$ is really a function of the vector $\boldsymbol{\alpha}$, so it seems misleading to refer to it as $\alpha(\mathbf{v})$, and we ought to call it $\boldsymbol{\alpha}(\mathbf{v})$ so then $\boldsymbol{\alpha}(\mathbf{v}) = <\boldsymbol{\alpha} \mid \mathbf{v}>$ has everything bolded on both sides. But then the functional would have to be called $\boldsymbol{\alpha} = <\boldsymbol{\alpha} \mid$. But this contradicts our notation earlier that $\boldsymbol{\alpha}$ is a vector, $\boldsymbol{\alpha}^{\mathrm{T}}$ is the transpose, and we should write $\boldsymbol{\alpha}^{\mathrm{T}} = <\boldsymbol{\alpha} \mid$. If we use $\alpha = <\boldsymbol{\alpha} \mid$ then we avoid that contradiction. This is what authors end up doing, writing a functional as a scalar entity which for us means an unbolded entity. A possible solution would be to say,

$$f_{\boldsymbol{\alpha}} = <\boldsymbol{\alpha} \mid = \text{functional}$$

$$f_{\boldsymbol{\alpha}}(\mathbf{v}) = <\boldsymbol{\alpha} \mid \mathbf{v}> = \boldsymbol{\alpha} \bullet \mathbf{v} = \text{function} \qquad (2.11.b.1)$$

where f is non-bold, and the subscript label $\boldsymbol{\alpha}$ is bold, but then we have introduced a new symbol f which seems superfluous. So the conclusion is this: $\alpha = <\boldsymbol{\alpha}| = $ functional, $\alpha(\mathbf{v}) = <\boldsymbol{\alpha} \mid \mathbf{v}> = $ function, and one must understand that $\alpha(\mathbf{v})$ is a function of the vector quantity $\boldsymbol{\alpha}$. Obviously there is a unique functional $\alpha(\mathbf{v})$ for each vector $\boldsymbol{\alpha}$ in V (and thus for each $\boldsymbol{\alpha}^{\mathrm{T}}$ in V*). The spaces V and V* have the same dimension n and are isomorphic to each other in the sense just noted.

(c) Basis vectors for the dual space V*

Now recall that our axis-aligned x-space basis vectors $\mathbf{u}_i$ have dual basis vectors $\mathbf{u}^i$ where $\mathbf{u}^i \bullet \mathbf{u}_j = \delta^i{}_j$ which is the idea of orthogonality in the covariant world (which might be non-Cartesian). In our notations above,

$$\begin{aligned} &\mathbf{u}^i \bullet \mathbf{u}_j = (\mathbf{u}^i)^{\mathrm{T}} \mathbf{u}_j = <\mathbf{u}^i \mid \mathbf{u}_j> = \delta^i{}_j \\ = \ &\mathbf{u}_i \bullet \mathbf{u}^j = (\mathbf{u}_i)^{\mathrm{T}} \mathbf{u}^j = <\mathbf{u}_i \mid \mathbf{u}^j> = \delta_i{}^j = \delta_{i,j} \ . \end{aligned} \qquad (2.11.c.1)$$

Since $\mathbf{u}_i$ and $\mathbf{u}^i$ are in general different column vectors in V, $(\mathbf{u}_i)^{\mathrm{T}}$ and $(\mathbf{u}^i)^{\mathrm{T}}$ are different row vectors in the dual space V*.

Just as the column vectors $|\mathbf{u}_i>$ and $|\mathbf{u}^i>$ form two distinct bases for V, the row vectors $<\mathbf{u}^i \mid$ and $<\mathbf{u}_i \mid$ form two distinct bases for V*. Certainly dim(V) = dim(V*).

<u>Definition of λ^i</u>

Above we discussed $\alpha = <\boldsymbol{\alpha}|$ as a vector functional, and $\alpha(\mathbf{v}) = <\boldsymbol{\alpha}|\mathbf{v}>$ as the corresponding scalar function. Whereas $<\boldsymbol{\alpha}|$ is some general vector in V*, we now consider in its place a basis vector $<\mathbf{u}^i|$ in V*. With what notation shall we represent this functional? In analogy with α and $\alpha(\mathbf{v})$ we could use u^i and $u^i(\mathbf{v})$ where the u^i is unbolded to indicate a scalar function. Or we could use $f_{\mathbf{u}^i} = <\mathbf{u}^i|$ and $f_{\mathbf{u}^i}(\mathbf{v}) = <\mathbf{u}^i| \mathbf{v}>$. The first notation is not uncommon (see wiki dual space where $\mathbf{u}^i = \mathbf{e}^i$), while the latter notation is unpleasant. Other common notations are $v^{*i}(\mathbf{v})$ or $e^{*i}(\mathbf{v})$ which for us would be $u^{*i}(\mathbf{v})$.

We shall use the following notation,

$\lambda^i \equiv \langle \mathbf{u}^i|$ basis functional in V* // $\lambda^i = (\mathbf{u}^i)^T$

so

$\lambda^i(\mathbf{v}) = \langle \mathbf{u}^i|\mathbf{v}\rangle$ basis function in V^*_f // $\lambda^i(\mathbf{v}) = (\mathbf{u}^i)^T\mathbf{v}$. (2.11.c.2)

The λ is unbolded, consistent with α(**v**). λ is a Greek letter consistent with our plan to use Greek or script letters for dual space objects. The index on λ^i is up, matching the index on $\mathbf{u}^i$ in $\langle \mathbf{u}^i|$. Notice that

$$\lambda^i(\mathbf{u}_j) = \langle \mathbf{u}^i|\mathbf{u}_j\rangle = \delta^i{}_j \ . \qquad (2.11.c.3)$$

We think of $\lambda^i \equiv \langle \mathbf{u}^i|$ as being in the dual space V* while $\lambda^i(\mathbf{v}) = \langle \mathbf{u}^i|\mathbf{v}\rangle$ lies in a directly corresponding space of functions which we call V^*_f. This is our first example of what we shall call a "tensor function".

Comment: Lang [1999] uses λ_i(p 130) for the his dual space basis functionals. Sjamaar used symbol λ_i in his 2006 notes (p 84), but changed to β_i in his 2015 update (p 91). Spivak uses φ_i (p 76). Wiki (dual basis) uses basis vectors v_i instead of e_i so their λ^i is called v^i. Wiki (dual space) uses e^i while Lang [2002] uses f^i (p 143). There seems to be no standard notation as in physics where **F** = m**a** is universally recognized as Newton's Second Law which would be hard to identify if written **G** = n**b**. Probably u^i or u^i (unbolded) is the most logical choice if the V basis vectors are $\mathbf{u}_i$, but it is so easy to confuse functional u^i with the vector $\mathbf{u}^i$ (especially when we drop our bolding of vectors starting in Chapter 3) that we shall stick with λ^i.

Eq. (2.7.1) line 1 gives the expansion of a vector **v** onto the $\mathbf{u}_i$

$\mathbf{v} = \Sigma_i v^i \mathbf{u}_i$ where $v^i = \mathbf{u}^i \bullet \mathbf{v}$

or

$|\mathbf{v}\rangle = \Sigma_i v^i |\mathbf{u}_i\rangle$ where $v^i = \langle \mathbf{u}^i| \mathbf{v} \rangle = \mathbf{u}^i \bullet \mathbf{v}$. (2.11.c.4)

Notice therefore that

$$\lambda^i(\mathbf{v}) = \langle \mathbf{u}^i|\mathbf{v}\rangle = \mathbf{u}^i \bullet \mathbf{v} = v^i \ . \qquad (2.11.c.5)$$

The function $\lambda^i(\mathbf{v})$ is sometimes called "the i^{th} coordinate function" since it projects out the i^{th} component the vector **v**. As summarized in (2.7.13), since each dot product $\mathbf{u}^i \bullet \mathbf{v}$ is a scalar, the functions $\lambda^i(\mathbf{v})$ i = 1..N transform as scalars despite the fact that the values of these scalars are the components of the vector v^i.

Transposing (2.11.c.4) produces a vector functional expansion in V*,

$$\mathbf{v}^T = \Sigma_i v^i (\mathbf{u}_i)^T \quad \text{or} \quad \langle \mathbf{v}| = \Sigma_i v^i \langle \mathbf{u}_i| = \Sigma_i v_i \langle \mathbf{u}^i| \ . \qquad (2.11.c.6)$$

Using Greek letters for dual space objects we write this as

$$\alpha = \langle \boldsymbol{\alpha}| = \Sigma_i \ \alpha_i \langle \mathbf{u}^i| = \Sigma_i \ \alpha_i \lambda^i \ . \qquad (2.11.c.7)$$

Then,

$$\alpha = \Sigma_i \alpha_i \lambda^i \qquad \text{functional} \qquad (2.11.c.8)$$

$$\alpha(\mathbf{v}) = \Sigma_i \alpha_i \lambda^i(\mathbf{v}) = \Sigma_i \alpha_i v^i = \boldsymbol{\alpha} \bullet \mathbf{v} \qquad \text{function} \qquad (2.11.c.9)$$

or in bra-ket notation,

$$\langle \boldsymbol{\alpha}| = \Sigma_i \alpha_i \langle \mathbf{u}^i| \qquad \text{functional}$$

$$\alpha(\mathbf{v}) = \Sigma_i \alpha_i \langle \mathbf{u}^i|\mathbf{v}\rangle = \Sigma_i \alpha_i v^i = \boldsymbol{\alpha} \bullet \mathbf{v} = \langle \boldsymbol{\alpha} |\mathbf{v}\rangle \qquad \text{function} \qquad (2.11.c.10)$$

and we replicate the result (2.11.a.4).

Definition of λ'^i

We have defined $\lambda^i \equiv \langle \mathbf{u}^i|$ as a notation for a certain basis functional in dual x-space. We would like to somehow define an object λ'^i which is a basis functional in dual x'-space. How should this be done?

One might intuitively feel that one should set $\lambda'^i \equiv \langle \mathbf{u}'^i|$. Or one might think that once λ^i is defined as above, then the meaning of λ'^i is forced upon us by some equation like $\lambda'^i = R^i{}_j \lambda^j$. Both these notions are *not* what we want to do. We are not forced to say $\lambda'^i \equiv \langle \mathbf{u}'^i|$ just because $\lambda^i \equiv \langle \mathbf{u}^i|$ since we are making two separate definitions. And $\lambda'^i = R^i{}_j \lambda^j$ is complete nonsense for the following reason. The N functionals λ^i for i = 1..N are each vectors in V*, so $\{\lambda^i\}$ is a set of vectors, not a set of numbers, whereas when one tries to write $\lambda'^i = R^i{}_j \lambda^j$ one is implying that λ^j is a set of numbers which form a vector.

Recall that the $\mathbf{u}_i$ are the "axis-aligned" basis vectors in x-space since $(\mathbf{u}_i)^j = (\mathbf{u}^i)_j = \delta_{i,j}$.
Recall that the $\mathbf{e}'_i$ are the "axis-aligned" basis vectors in x'-space since $(\mathbf{e}'_i)^j = (\mathbf{e}'^i)_j = \delta_{i,j}$.

This suggests that the proper definition of λ'^i is the following:

$$\lambda'^i \equiv \langle \mathbf{e}'^i| \ . \qquad (2.11.c.11)$$

One then finds that, for $\mathbf{v}'$ a vector in x'-space,

$$\lambda'^i(\mathbf{v}') \equiv \langle \mathbf{e}'^i| \mathbf{v}'\rangle = v'^i \qquad (2.11.c.12)$$

which is then analogous to

$$\lambda^i(\mathbf{v}) \equiv \langle \mathbf{u}^i| \mathbf{v}'\rangle = v^i \ . \qquad \textit{(2.11.c.5)}$$

In both cases then λ^i and λ'^i are the "i^{th} coordinate functions", projecting out the i^{th} coordinate from a vector.

Since $v'^i = R^i{}_j v^j$ we can certainly write

$$\lambda'^i(\mathbf{v}') = R^i{}_j \ \lambda^j(\mathbf{v}) \qquad (2.11.c.13)$$

as a statement relating two vectors of scalars. Notice this does not say $\lambda'^i = R^i{}_j\lambda^j$ which we already noted above does not even make sense. If we display the fact that $R^i{}_j$ in general is $R^i{}_j(\mathbf{x})$ then

$$\lambda'^i(\mathbf{v}') = R^i{}_j(\mathbf{x}) \ \lambda^j(\mathbf{v}) \ . \qquad (2.11.c.14)$$

Since this does not fit into any of the molds shown in (2.1.16), one cannot quite claim that $\lambda^j(\mathbf{v})$ transforms as a vector field, but the transformation is similar.

We can study (2.11.c.13) in Dirac notation as follows (see below for Dirac notation details),

$$\lambda'^i(\mathbf{v}') = \langle \mathbf{e}'^i|\mathbf{v}'\rangle = \langle \mathbf{e}^i|\mathbf{v}\rangle = \langle \mathbf{e}^i|\ \mathit{1}\ |\ \mathbf{v}\rangle = \langle \mathbf{e}^i|\mathbf{u}_j\rangle\langle\ \mathbf{u}^j|\ \mathbf{v}\rangle = \mathbf{e}^i\bullet \mathbf{u}_j \ \lambda^j(\mathbf{v}) = R^i{}_j\lambda^j(\mathbf{v}) \qquad (2.11.c.15)$$

where the last step comes from (2.4.3).

Once one has a functional λ'^i, one can define a general rank-1 functional α' in dual x'-space as follows:

$$\alpha' = \Sigma_i\alpha'_i\lambda'^i \qquad \text{functional in } V'^* \qquad (2.11.c.16)$$

$$\alpha'(\mathbf{v}') = \Sigma_i\alpha'_i\lambda'^i(\mathbf{v}') = \Sigma_i\alpha'_i v'^i = \boldsymbol{\alpha}' \bullet \mathbf{v}' \qquad \text{function in } V'^*_f \ . \qquad (2.11.c.17)$$

It then follows that

$$\alpha'(\mathbf{v}') = \boldsymbol{\alpha}' \bullet \mathbf{v}' = \boldsymbol{\alpha} \bullet \mathbf{v} = \alpha(\mathbf{v}) \qquad (2.11.c.18)$$

and in some sense one could say that α(**v**) transforms as a scalar field, where **v** plays the role normally occupied by the position vector **x**. On the other hand, the vector |**v**> and the dual vector (functional) α = <**α**| transform as vectors and so α is a vector functional.

We now define α(**v**) to be a **"rank-1 tensor function"**. Spivak would call it a "1-tensor". We have this seeming contradiction that α(**v**) is a rank-1 tensor function, yet that function transforms as a rank-0 scalar. The rank-1 description really applies to the functional α = <**α**| which is in fact a vector and transforms as a vector. When this is closed with the ket |**v**> one obtains the scalar object α(**v**) = <**α** | **v**>. (2.11.c.19)

<u>Vector space names:</u> V, V^*, V^*_f and V', V'^*, V'^*_f (2.11.c.20)

Here we have associated vector space names V and V* with x-space in Picture A (2.1.1), while V' and V'* are associated with x'-space. All these spaces have the same dimension n and all are isomorphic. There is a 1-to-1 relationship between V and dual space V* as noted above, and there is a 1-to-1 relationship between V and V' since for every vector **v** in x-space there is a unique corresponding vector **v**' = R**v** in x'-space. We refer to V* as dual x-space and V'* as dual x'-space. Associated with the dual space V* of functionals is the space V^*_f of corresponding functions, and similarly for V'* and V'^*_f.

(d) Rank-2 functionals and tensor functions

A rank-2 tensor may be represented as

$$T = \Sigma_{ab}\ T^{ab}\ \mathbf{u}_a \otimes \mathbf{u}_b \qquad (2.11.d.1)$$

$$\mathbf{V} = \Sigma_a\ V^a\ \mathbf{u}_a$$

where on the second line for comparison we show a general rank-1 tensor (vector). In Dirac notation, we write

$$|\ \mathbf{u}_a, \mathbf{u}_b> \ \equiv\ |\mathbf{u}_a> \otimes\ |\mathbf{u}_b> \qquad \leftrightarrow \qquad \mathbf{u}_a \otimes \mathbf{u}_b \qquad (2.11.d.2)$$

which represents any of the n^2 basis vectors of the tensor product space $V^2 = V\otimes V$. We could write this as $|\ \mathbf{u}_a, \mathbf{u}_b>_2\ \equiv\ |\mathbf{u}_a>_1 \otimes |\mathbf{u}_b>_1$ to distinguish the fact that some kets are in V^1 and others in V^2, but the contents of the ket usually make it obvious to which vector space a ket belongs. In Dirac notation, the tensor T is written

$$|T> = \ \Sigma_{ab}T^{ab}\ |\mathbf{u}_a> \otimes |\mathbf{u}_b> \ =\ \Sigma_{ab}T^{ab}\ |\ \mathbf{u}_a, \mathbf{u}_b> \qquad (2.11.d.3)$$

and this is a general element of the space V^2. The corresponding **rank-2 linear functional** in the dual space V^{*2} is given by

$$<T|\ =\ \Sigma_{ab}T_{ab} <\mathbf{u}^a\ | \otimes <\mathbf{u}^b| \ =\ \Sigma_{ab}T_{ab} <\mathbf{u}^a, \mathbf{u}^b|\ . \qquad (2.11.d.4)$$

This is done in analogy with the vector case

$$|\mathbf{V}> = \ \Sigma_a\ V^a\ |\mathbf{u}_a>$$

$$<\mathbf{V}| = \Sigma_a\ V_a<\mathbf{u}^a\ | \quad . \qquad (2.11.d.5)$$

We are careful to have the index "tilt" have the form of a contraction, even though we are not really contracting indices on a tensor. The rank-2 functional $<T|$ is linear in both V* spaces of V*⊗V*, so it is called a **bilinear** functional. If we let (subscript 1 and 2 are labels of two vectors, not components of **v**)

$$|\ \mathbf{v}_1, \mathbf{v}_2> \ \equiv\ |\mathbf{v}_1> \otimes |\mathbf{v}_2> \qquad (2.11.d.6)$$

represent an arbitrary (but pure) element of $V^2 = V\otimes V$, then we may construct

$\mathcal{T} = <\text{T}|$ rank-2 tensor functional

$\mathcal{T}(\mathbf{v_1},\mathbf{v_2}) = <\text{T} | \mathbf{v_1}, \mathbf{v_2}>$ rank-2 tensor function (a Spivak "2-tensor") . (2.11.d.7)

It follows that

$$\mathcal{T}(\mathbf{v_1},\mathbf{v_2}) = <\text{T} | \mathbf{v_1}, \mathbf{v_2}> = \Sigma_{ab}T_{ab} < \mathbf{u}^a, \mathbf{u}^b| \mathbf{v_1}, \mathbf{v_2}>$$

$$= \Sigma_{ab}T_{ab} <\mathbf{u}^a| \mathbf{v_1}> <\mathbf{u}^b| \mathbf{v_2}>$$

$$= \Sigma_{ab}T_{ab} (\mathbf{v_1})^a (\mathbf{v_2})^b \qquad (2.11.d.8)$$

where we have used the fact that the scalar product for elements of V^{*2} with elements of V^2 is the product of two V*-with-V scalar products, as seen for example in (2.9.13). In the last line above we see that the tensor function $\mathcal{T}(\mathbf{v_1},\mathbf{v_2})$ is the contraction of a rank-2 tensor with two rank-1 tensors, and so is a scalar. Thus,

$$\mathcal{T}'(\mathbf{v'_1},\mathbf{v'_2}) = \mathcal{T}(\mathbf{v_1},\mathbf{v_2}) \qquad (2.11.d.9)$$

and a rank-2 tensor function transforms as a "scalar field of two arguments". The "rank-2" description applies to the tensor functional <T| , and when this is closed with an element of V^2 the result is a scalar.

Note from (2.11.d.8) and (2.4.1) that $(\mathbf{u_i})^a = \delta_i^{\ a}$,

$$\mathcal{T}(\mathbf{u_i},\mathbf{u_j}) = \Sigma_{ab}T_{ab} (\mathbf{u_i})^a (\mathbf{u_j})^b = \Sigma_{ab}T_{ab} \delta_i^{\ a} \delta_j^{\ b} = T_{ij} \qquad (2.11.d.10)$$

so the tensor function evaluated at the basis vectors gives a corresponding element of the tensor.

Using $\lambda^i = < \mathbf{u}^i|$ as defined above in (2.11.c.2), we can rewrite (2.11.d.4)

$$<\text{T}| = \Sigma_{ab}T_{ab} <\mathbf{u}^a | \otimes <\mathbf{u}^b|$$

as rank-2 tensor functional

$$\mathcal{T} = \Sigma_{ab}T_{ab} \lambda^a \otimes \lambda^b \qquad (2.11.d.11)$$

which is analogous to

$$<\boldsymbol{\alpha} | = \Sigma_a \alpha_a <\mathbf{u}^a|$$

rank-1 tensor functional

$$\alpha = \Sigma_a \alpha_a \lambda^a . \qquad (2.11.d.12)$$

We continue to use script or Greek fonts to represent functionals, such as α and $\mathcal{T}$.

Taking the special case of a rank-2 functional which is just $\lambda^a \otimes \lambda^b$ we construct the following rank-2 tensor function,

$$(\lambda^{\mathbf{a}} \otimes \lambda^{\mathbf{b}})(\mathbf{v_1},\mathbf{v_2}) = \langle\, \mathbf{u^a}, \mathbf{u^b} |\, \mathbf{v_1}, \mathbf{v_2} \rangle = \langle \mathbf{u^a} |\, \mathbf{v_1} \rangle \langle \mathbf{u^b} |\, \mathbf{v_2} \rangle = (\mathbf{v_1})^{\mathbf{a}}\ (\mathbf{v_2})^{\mathbf{b}}$$

$$= \lambda^{\mathbf{a}}(\mathbf{v_1})\, \lambda^{\mathbf{b}}(\mathbf{v_2}) \ . \qquad (2.11.d.13)$$

This function is manifestly linear in both arguments, since $\lambda^{\mathbf{a}}(\mathbf{v_1})$ is linear, so it is a bilinear function. For example,

$$(\lambda^{\mathbf{a}} \otimes \lambda^{\mathbf{b}})(\mathbf{v_1}+\mathbf{v_1}',\mathbf{v_2}) = (\mathbf{v_1}+\mathbf{v_1}')^{\mathbf{a}}\ (\mathbf{v_2})^{\mathbf{b}} = (\mathbf{v_1})^{\mathbf{a}}\ (\mathbf{v_2})^{\mathbf{b}} + (\mathbf{v'_1})^{\mathbf{a}}\ (\mathbf{v_2})^{\mathbf{b}}$$

$$= (\lambda^{\mathbf{a}} \otimes \lambda^{\mathbf{b}})(\mathbf{v_1},\mathbf{v_2}) + (\lambda^{\mathbf{a}} \otimes \lambda^{\mathbf{b}})(\mathbf{v'_1},\mathbf{v_2}) \ . \qquad (2.11.d.14)$$

Whereas $\langle T|$ shown above is a *general* rank-2 tensor functional, we can consider the special case of a *pure* rank-2 functional formed from two vector functionals $\alpha = \langle \boldsymbol{\alpha}|$ and $\beta = \langle \boldsymbol{\beta}|$. In that case one finds,

$(\alpha \otimes \beta) = \langle \boldsymbol{\alpha}| \otimes \langle \boldsymbol{\beta}| = \langle \boldsymbol{\alpha}, \boldsymbol{\beta} |$ rank-2 functional

$(\alpha \otimes \beta)(\mathbf{v_1},\mathbf{v_2}) = \langle \boldsymbol{\alpha}, \boldsymbol{\beta} \,|\, \mathbf{v_1}, \mathbf{v_2} \rangle$
$= \langle \boldsymbol{\alpha} \,|\, \mathbf{v_1} \rangle \ \langle \boldsymbol{\beta} \,|\, \mathbf{v_2} \rangle = \alpha(\mathbf{v_1})\beta(\mathbf{v_2})$ rank-2 tensor function

$(\alpha \otimes \beta)(\mathbf{u_i},\mathbf{u_j}) = \alpha(\mathbf{u_i})\beta(\mathbf{u_j}) = \alpha_i\beta_j = (\alpha \otimes \beta)_{ij}$ rank-2 tensor (2.11.d.15)

where the very last item is $\alpha_i\beta_j$ expressed in the tensor product notation of (2.8.9). Once again, evaluation of a tensor function at two basis vectors creates an element of the tensor.

<u>Comment on vertical bars in the Dirac Notation</u>

Let $|a\rangle$ be a vector in V, and $\langle b|$ a vector in the dual space V*. Notice that

$$\langle b| \ \ |a\rangle = \langle b||a\rangle = \langle b|a\rangle \ . \qquad (2.11.d.16)$$

The official notation for the scalar product is $\langle b \,|\, a\rangle$ not $\langle b \,||\, a\rangle$ so one replaces the || with | . The same replacement is made for example doing a scalar product between elements of V^{*2} and V^2

$$\langle a| \otimes \langle b| \quad |c\rangle \otimes |d\rangle = \langle a,b||c,d\rangle = \langle a,b \,|\, c,d\rangle$$

or

$$\langle a| \otimes \langle b| \quad |c\rangle \otimes |d\rangle = (\ \langle a| \ \ |c\rangle\ )\ (\ \langle b| \ \ |d\rangle\ ) = \langle a|c\rangle\langle b|d\rangle \ . \qquad (2.11.d.17)$$

(e) Rank-k functionals and tensor functions

It is a simple matter to generalize from k = 2 to k = k, so the vector space is V^k and the dual space is V^{*k},

$$\mathcal{V}^k \equiv VxVx....xV \qquad \text{k factors} \qquad \text{// Cartesian product of k spaces}$$

$$V^k \equiv V\otimes V\otimes....\otimes V \qquad \text{k factors} \qquad \text{// tensor product of k vector spaces}$$

$$V^{*k} \equiv V^*\otimes V^*\otimes....\otimes V^* \qquad \text{k factors} \qquad \text{// tensor product of k dual spaces} \quad . \tag{2.11.e.1}$$

We then have as a most general element of V^k (a rank-k tensor),

$$T = \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} (\mathbf{u}_{i_1}\otimes \mathbf{u}_{i_2} .....\otimes \mathbf{u}_{i_k}) \ . \qquad T = \Sigma_I T^I u_I \tag{2.11.e.2}$$

with

$$(\mathbf{u}_{i_1}\otimes \mathbf{u}_{i_2} .....\otimes \mathbf{u}_{i_k}) = |\mathbf{u}_{i_1}> \otimes |\mathbf{u}_{i_2}>.....\otimes |\mathbf{u}_{i_k}> = |\mathbf{u}_{i_1}, \mathbf{u}_{i_2} ....., \mathbf{u}_{i_k}>_k \tag{2.11.e.3}$$

$$= |\mathbf{u}_{i_1}, \mathbf{u}_{i_2} ....., \mathbf{u}_{i_k}> . \qquad u_I \equiv \mathbf{u}_{i_1}\otimes \mathbf{u}_{i_2} .....\otimes \mathbf{u}_{i_k}$$

On the right in red we show our equations expressed in the multi-index notation introduced in (2.10.20) and (2.10.21). The letter Z which appears below is used to represent the set of integers 1,2...k.

Then the rank-k tensor T in V^k is represented in Dirac notation as

$$|T> = \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} |\mathbf{u}_{i_1}, \mathbf{u}_{i_2} ....., \mathbf{u}_{i_k}> . \qquad |T> = \Sigma_I T^I |u_I> \tag{2.11.e.4}$$

The rank-k tensor *functional* <T| of V^{*k} is then

$$<T| = \Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} <\mathbf{u}^{i_1}, \mathbf{u}^{i_2} ....., \mathbf{u}^{i_k}| \qquad <T| = \Sigma_I T_I <u^I|$$

or

$$\mathcal{T} = \Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} \lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_k} \ . \qquad \mathcal{T} = \Sigma_I T_I \lambda^I \tag{2.11.e.5}$$

A general pure element of V^k is specified by

$$|\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k> = |\mathbf{v}_1> \otimes |\mathbf{v}_2> \otimes .... \otimes |\mathbf{v}_k> . \qquad |\mathbf{v}_Z> = |\mathbf{v}_1> \otimes |\mathbf{v}_2> \otimes .... \otimes |\mathbf{v}_k> \tag{2.11.e.6}$$

The corresponding rank-k tensor *function* is given by

$$\mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k) = \langle T \mid \mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k\rangle$$

$$= \Sigma_{i_1 i_2 ....i_k} T_{i_1 i_2 ....i_k} \langle \mathbf{u}^{i_1}, \mathbf{u}^{i_2}, ..., \mathbf{u}^{i_k} \mid \mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k\rangle$$

$$= \Sigma_{i_1 i_2 ....i_k} T_{i_1 i_2 ....i_k} \langle \mathbf{u}^{i_1}|\mathbf{v}_1\rangle\langle \mathbf{u}^{i_2}|\mathbf{v}_2\rangle .... \langle \mathbf{u}^{i_k}|\mathbf{v}_k\rangle \qquad (2.11.e.7)$$

$$= \Sigma_{i_1 i_2 ....i_k} T_{i_1 i_2 ....i_k} (\mathbf{v}_1)^{i_1} (\mathbf{v}_2)^{i_2} ....(\mathbf{v}_k)^{i_k} . \qquad \mathcal{T}(\mathbf{v}_Z) = \Sigma_I T_I (\mathbf{v}_Z)^I$$

This shows that the rank-k tensor function is a linear combination of the products of the argument components weighted by the components of the corresponding rank-k tensor. Since this is the contraction of a rank-k tensor with k rank-1 tensors, the result transforms as a scalar, so then

$$\mathcal{T}'(\mathbf{v'}_1, \mathbf{v'}_2, ...\mathbf{v'}_k) = \mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k) . \qquad \mathcal{T}'(\mathbf{v'}_Z) = \mathcal{T}(\mathbf{v}_Z) \qquad (2.11.e.8)$$

That is to say, the rank-k tensor function transforms as a scalar field, where the term "rank-k" is associated with the functional $\mathcal{T} = \langle T|$ which is an element of the dual space V^{*k} . Finally we see that

$$\mathcal{T}(\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, ... \mathbf{u}_{j_k}) = \langle T \mid \mathbf{u}_{j_1}, \mathbf{u}_{j_2}, ... \mathbf{u}_{j_k} \rangle$$

$$= \Sigma_{i_1 i_2 ....i_k} T_{i_1 i_2 ....i_k} (\mathbf{u}_{j_1})^{i_1} (\mathbf{u}_{j_2})^{i_2} ....(\mathbf{u}_{j_k})^{i_k}$$

$$= T_{j_1 j_2 ....j_k} . \qquad \mathcal{T}(\mathbf{u}_J) = T_J \qquad (2.11.e.9)$$

From (2.11.e.7) one sees that the tensor function $\mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k)$ is manifestly **k-multilinear**, which is the generalization of linear for k = 1 and bilinear for k = 2.

Once can construct a rank-k tensor functional purely from the dual basis vectors,

$$(\lambda^{i_1}\otimes\lambda^{i_2}\otimes ... \otimes\lambda^{i_k}) = \langle\mathbf{u}^{i_1}| \otimes \langle\mathbf{u}^{i_2}| ... \langle\mathbf{u}^{i_k}| \quad \text{rank-k tensor functional} \qquad \lambda^I = \langle u^I| \qquad (2.11.e.10)$$

$$(\lambda^{i_1}\otimes\lambda^{i_2}\otimes ... \otimes\lambda^{i_k})(\mathbf{v}_1, \mathbf{v}_2 ....\mathbf{v}_k) = \lambda^{i_1}(\mathbf{v}_1)\lambda^{i_2}(\mathbf{v}_2)....\lambda^{i_k}(\mathbf{v}_k)$$
$$= (\mathbf{v}_1)^{i_1}(\mathbf{v}_2)^{i_2} ... (\mathbf{v}_k)^{i_k} \qquad \text{rank-k tensor function} \qquad \lambda^I(\mathbf{v}_Z) = (\mathbf{v}_Z)^I \qquad (2.11.e.11)$$

$$(\lambda^{i_1}\otimes\lambda^{i_2}\otimes ... \otimes\lambda^{i_k})(\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, ... \mathbf{u}_{j_k}) = \lambda^{i_1}(\mathbf{u}_{j_1})\lambda^{i_2}(\mathbf{u}_{j_2})....\lambda^{i_k}(\mathbf{u}_{j_k})$$
$$= (\mathbf{u}_{j_1})^{i_1}(\mathbf{u}_{j_2})^{i_2} ... (\mathbf{u}_{j_k})^{i_k}$$
$$= \delta_{j_1}{}^{i_1}\delta_{j_2}{}^{i_2} ... \delta_{j_k}{}^{i_k} \qquad \text{evaluated at basis vectors} . \qquad \lambda^I(u_J) = \delta^I{}_J \qquad (2.11.e.12)$$

As an alternative to the most general rank-k tensor functional $\mathcal{T}$ and the all-basis-vector rank-k tensor functional $(\lambda^{i_1}\otimes\lambda^{i_2}\otimes ... \otimes\lambda^{i_k})$, one can consider a "pure" rank-k tensor functional constructed from k dual vectors which we shall call $\langle\boldsymbol{\alpha}_i|$. In this case we find,

$<\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2....\boldsymbol{\alpha}_k| = <\boldsymbol{\alpha}_1| \otimes <\boldsymbol{\alpha}_2| .... \otimes <\boldsymbol{\alpha}_k|$ pure rank-k tensor functional

$= \alpha_1 \otimes \alpha_2... \otimes \alpha_k \quad = (\alpha_1 \otimes \alpha_2 ... \otimes \alpha_k)$ (2.11.e.13)

$(\alpha_1 \otimes \alpha_2 ... \otimes \alpha_k)(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k) = \alpha_1(\mathbf{v}_1)\alpha_2(\mathbf{v}_2) ....\alpha_k(\mathbf{v}_k)$

$= (\boldsymbol{\alpha}_1 \bullet \mathbf{v}_1)(\boldsymbol{\alpha}_2 \bullet \mathbf{v}_2) ....(\boldsymbol{\alpha}_k \bullet \mathbf{v}_k)$ pure rank-k tensor function (2.11.e.14)

$(\alpha_1 \otimes \alpha_2 ... \otimes \alpha_k)(\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, ... \mathbf{u}_{j_k}) = \alpha_1(\mathbf{u}_{j_1})\alpha_2(\mathbf{u}_{j_2}) ....\alpha_k(\mathbf{u}_{j_k})$

$= (\boldsymbol{\alpha}_1 \bullet \mathbf{u}_{j_1}) (\boldsymbol{\alpha}_2 \bullet \mathbf{u}_{j_2}) ... (\boldsymbol{\alpha}_k \bullet \mathbf{u}_{j_k})$ evaluated at $\mathbf{u}_r$

$= (\alpha_1)_{j_1}(\alpha_2)_{j_2}...(\alpha_k)_{j_k} = (\alpha_1 \otimes \alpha_2 ... \otimes \alpha_k)_{j_1 j_2 ... j_k}$ outer product notation (2.11.e.15)

Hopefully after this long slog, the following paragraph makes some sense to the reader:

A rank-k tensor function is the bra-ket closure (inner product) of a rank-k dual tensor functional $<T|$ of V^{*k} with a pure rank-k non-dual tensor $|\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k>$ of V^k such that $\mathcal{T}(\mathbf{v}_1,\mathbf{v}_2,...\mathbf{v}_k) = <T|\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k>$. The tensor function is k-multilinear in its arguments, and transforms as a scalar field with k vector arguments. When the rank-k tensor function is evaluated at the basis vectors $\mathbf{u}_r$, it replicates the non-dual rank-k tensor with which is it associated, which is to say, $\mathcal{T}(\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, ... \mathbf{u}_{j_k}) = T_{j_1 j_2 ... j_k}$. Spivak on page 75 refers to a rank-k tensor function as a "k-tensor". (2.11.e.16)

As we shall see later, a motivation for using tensor functions is their crashingly simple description of the tensor product of an arbitrary rank-k tensor with an arbitrary rank-k' tensor to produce a rank-(k+k') tensor :

$_k<T | \mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k>_k \ _{k'}< S| \mathbf{v}_{k+1},\mathbf{v}_{k+2}...\mathbf{v}_{k+k'}>_{k'}$

$= [\ _k<T \otimes \ _{k'}<S|\ ] \ \ [|\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k>_k \otimes |\mathbf{v}_{k+1},\mathbf{v}_{k+2}...\mathbf{v}_{k+k'}>_{k'}]$

$= \ _{k+k'}<T \otimes S | \mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_{k+k'}>_{k+k'}$ (2.11.e.17)

or

$\mathcal{T}(\mathbf{v}_1,\mathbf{v}_2,...\mathbf{v}_k)\, \mathcal{S}(\mathbf{v}_{k+1},\mathbf{v}_{k+2},...\mathbf{v}_{k+k'}) = (\mathcal{T} \otimes \mathcal{S})(\mathbf{v}_1,\mathbf{v}_2 .... \mathbf{v}_{k+k'})$. (2.11.e.18)

This equation appears below as (6.6.13) and also appears in Spivak page 75.

As noted by Benn and Tucker page 2, the relationship between the vector space V^k and the dual vector space V^{*k} is a reciprocal one. One could, as they say, perversely regard V^{*k} as the starting vector space and then V^k would be the dual space of V^{*k}. This amounts to swapping bra ↔ ket in the Dirac notation outlined above. Instead of having a functional $\alpha(\mathbf{v}) = <\boldsymbol{\alpha}|\mathbf{v}>$, one would have a functional $v(\boldsymbol{\alpha}) = <\mathbf{v}|\boldsymbol{\alpha}>$.

We find that things are hard enough to understand without doing this "perverse" swapping of things right off the bat as they do. They refer to a rank-k tensor as a tensor of degree k, while other authors refer to rank as the order of a tensor. We us the term rank and promise not to confuse it with the different notion of the rank of a matrix which is the number of linearly independent rows or columns, or with various other meanings of the word "rank" in mathematics.

(f) The Covariant Transpose

Whereas the *matrix transpose* of a matrix $M_a{}^b$ would be $(M^{\mathrm{T}})_a{}^b = M_b{}^a$ (swap the rows and columns), it is the *covariant transpose* $(M^{\boldsymbol{T}})_a{}^b = M^b{}_a$ that is significant in covariant notation. We quote from *Tensor* where M is a general rank-2 tensor while R and S are the "differentials" of (2.1.2),

$$\begin{array}{lll} (M^{\boldsymbol{T}})^{ab} = M^{ba} & (R^{\boldsymbol{T}})^{ab} = R^{ba} & (S^{\boldsymbol{T}})^{ab} = S^{ba} \\ (M^{\boldsymbol{T}})^a{}_b = M_b{}^a & (R^{\boldsymbol{T}})^a{}_b = R_b{}^a & (S^{\boldsymbol{T}})^a{}_b = S_b{}^a \\ (M^{\boldsymbol{T}})_a{}^b = M^b{}_a & (R^{\boldsymbol{T}})_a{}^b = R^b{}_a & (S^{\boldsymbol{T}})_a{}^b = S^b{}_a \\ (M^{\boldsymbol{T}})_{ab} = M_{ba} & (R^{\boldsymbol{T}})_{ab} = R_{ba} & (S^{\boldsymbol{T}})_{ab} = S_{ba} \ . \end{array} \qquad (7.9.3)' \qquad (2.11.f.1)$$

Equations in any column can be obtained by lowering one or both indices in the top equation, so that the covariant transpose $M^{\boldsymbol{T}}$ is a rank-2 tensor if M is a rank-2 tensor.

For all-up or all-down indices, the two kinds of transposes are the same: $(M^{\mathrm{T}})^{ab} = (M^{\boldsymbol{T}})^{ab} = M^{ba}$.

The covariant transpose has the indices reflected in a vertical line between the indices.

The following facts involve the covariant transpose,

$$\det(M) = \det(M^{\mathrm{T}}) = \det(M^{\boldsymbol{T}}) \qquad (7.9.7)' \qquad (2.11.f.2)$$

$$\begin{array}{lll} RR^{\boldsymbol{T}} = R^{\boldsymbol{T}}R = 1 & SS^{\boldsymbol{T}} = S^{\boldsymbol{T}}S = 1 & RS = SR = 1 \\ R^{\boldsymbol{T}} = R^{-1} = S & S^{\boldsymbol{T}} = S^{-1} = R \ . & \end{array} \qquad (7.9.8)' \qquad (2.11.f.3)$$

The fact that $R^{\boldsymbol{T}} = S$ and $S^{\boldsymbol{T}} = R$ and is just a restatement of (2.1.4) and RS = SR = 1 is (2.1.3). The fact that $\det(M) = \det(M^{\mathrm{T}})$ is well known, where one swaps the rows and columns. The fact that $\det(M) = \det(M^{\boldsymbol{T}})$ is proven in (A.1.22).

(g) Linear Dirac Space Operators

Consider these three ways of writing *the same* real number, where M is a matrix sandwiched between vector **b** on the right and transpose vector **a** on the left,

$\mathbf{a}^{\mathrm{T}}(M\mathbf{b})$ M acts to the right $\qquad (* \ * \ *) \left[\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \begin{pmatrix} * \\ * \\ * \end{pmatrix} \right]$

$(\mathbf{a}^{\mathrm{T}}M)\mathbf{b}$ M acts to the left, and note that $(a^{\mathrm{T}}M) = (M^{\boldsymbol{T}}a)^{\mathrm{T}}$ † $\qquad \left[(* \ * \ *) \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \right] \begin{pmatrix} * \\ * \\ * \end{pmatrix}$

$\mathbf{a}^{\mathrm{T}}M\,\mathbf{b}$ can think of M acting either to the right or to the left. (2.11.g.1)

† $[(M^T a)^T]^i = (M^T a)^i = (M^T)^i{}_j a^j = M_j{}^i a^j = a^j M_j{}^i = (a^T M)^i$

In writing these equations, one normally thinks of M as being a matrix

$M^i{}_j = (M^{[u]})^i{}_j$ or $M = M^{[u]}$.

By default, the matrix elements are taken in the axis-aligned $\mathbf{u}_i$ basis on both left and right (and this applies to all indices as discussed at the end of Section 2.4 so that

$$(\mathbf{u}^i)^T M (\mathbf{u}_j) = \Sigma_{a,b} (\mathbf{u}^i)_a M^a{}_b (\mathbf{u}_j)^b = \Sigma_{a,b} \delta^i{}_a M^a{}_b \delta_j{}^b = M^i{}_j \quad // = (M^{[u]})^i{}_j. \tag{2.11.g.2}$$

and then $M^{[u]} = M$. One could, however, do this in some other basis, for example using the tangent base vectors $\mathbf{e}_i$,

$$(\mathbf{e}^i)^T M (\mathbf{e}_j) = \Sigma_{a,b} (\mathbf{e}^i)_a M^a{}_b (\mathbf{e}_j)^b = \Sigma_{a,b} R^i{}_a M^a{}_b R_j{}^b = (RMR^T)^i{}_j \quad // = (M^{[e]})^i{}_j \tag{2.11.g.3}$$

and the result is a completely different matrix. In this case the matrices are related by a covariant similarity transformation by R

$$M^{[e]} = R\, M^{[u]} R^T . \qquad // M' = RMR^T \tag{2.11.g.4}$$

It is useful to think of the object M as being a basis-independent abstract linear *operator* which, when sandwiched between certain basis vectors, has certain matrix elements. Different types of basis vectors yield different matrices. One could also have mixed basis elements,

$$(\mathbf{u}^i)^T M (\mathbf{e}_j) = \Sigma_{a,b} (\mathbf{u}^i)_a M^a{}_b (\mathbf{e}_j)^b = \Sigma_{a,b} \delta^i{}_a M^a{}_b R_j{}^b = (MR^T)^i{}_j \quad // = (M^{[u,e]})^i{}_j \tag{2.11.g.5}$$

so in this case we get

$$M^{[u,e]} = M^{[u]} R^T . \tag{2.11.g.6}$$

The abstract operator M only becomes a matrix when it is properly sandwiched between basis vectors.

This notion of thinking of the object M as a basis-independent linear operator becomes more pronounced in the Dirac notation. We restate the above equations as follows, all of which evaluate to the same real number,

$\langle \mathbf{a}| (M |\mathbf{b}\rangle)$ M acts to the right $= \langle \mathbf{a} | M\mathbf{b} \rangle$

$(\langle \mathbf{a}|M) |\mathbf{b}\rangle$ M acts to the left, and note that $\langle \mathbf{a}|M = \langle M^T\mathbf{a} |$ $= \langle M^T\mathbf{a} | \mathbf{b} \rangle$

$\langle \mathbf{a}| M |\mathbf{b}\rangle$ can think of M acting either to the right or to the left. (2.11.g.7)

The space between the vertical bars is inhabited by abstract linear operators like M. The matrix elements shown above are then

$$<\mathbf{u}^i \,|\, M \,|\, \mathbf{u}_j> \;=\; (M^{[u]})^i{}_j \;= M^i{}_j$$

$$<\mathbf{e}^i \,|\, M \,|\, \mathbf{e}_j> \;=\; (M^{[e]})^i{}_j \;=\; (RMR^T)^i{}_j$$

$$<\mathbf{u}^i \,|\, M \,|\, \mathbf{e}_j> \;=\; (M^{[u,e]})^i{}_j \;=\; (MR^T)^i{}_j\,. \qquad // \; M = M^{[u]} \qquad (2.11.g.8)$$

To emphasize this notion of abstract operator, we shall write the operator in a different font, so M is a matrix and $\mathcal{M}$ is a Dirac-space operator, and then

$$<\mathbf{a}|\, \mathcal{M}\, |\mathbf{b}> \;=\; <\mathbf{a}|\,(\; \mathcal{M}\, |\mathbf{b}>) \;=\; <\mathbf{a}\, |M\, \mathbf{b}> \qquad = \text{a scalar product of two vectors}$$
$$<\mathbf{a}|\, \mathcal{M}\, |\mathbf{b}> \;= (<\mathbf{a}|\mathcal{M}\,)\;\; |\mathbf{b}> \;\; = <M^T\mathbf{a}\,|\, \mathbf{b}\,> \qquad = \text{a scalar product of two vectors}$$

$$<\mathbf{u}^i \,|\, \mathcal{M} \,|\, \mathbf{u}_j> \;=\; (M^{[u]})^i{}_j \;= M^i{}_j \qquad \text{etc .} \qquad (2.11.g.9)$$

Here then is a review of the matrix and Dirac notations,

$$\mathbf{a}' \;=\; (M\mathbf{a}) = (M)\mathbf{a} \Rightarrow \qquad (\mathbf{b}')^T = (M\mathbf{b})^T \;= b^T\, M^T \qquad \text{matrix notation}$$
$$|\mathbf{a}'> = \;\; |M\mathbf{a}> = \mathcal{M}|\mathbf{a}> \qquad <\mathbf{b}'| \;=\; <M\mathbf{b}| \;=\; <\mathbf{b}|\mathcal{M}^T\;. \qquad \text{Dirac notation} \qquad (2.11.g.10)$$

Then consider the following claim

Fact: $<\mathbf{a}\,|\, \mathcal{M}\,|\, \mathbf{b}> = <\mathbf{b}|\, \mathcal{M}^T\,|\, \mathbf{a}>$ (2.11.g.11)

where both $\mathcal{M}$ and $\mathcal{M}^T$ are the names of abstract linear operators.

Proof:

$$<\mathbf{a}\,|\, \mathcal{M}\,|\, \mathbf{b}> \equiv\; <\mathbf{a}\,|\, M\, \mathbf{b}> \;=\; a^i(M\mathbf{b})_i \;=\; a^i[\, M_i{}^j b_j] \;=\; a^i\, M_i{}^j\;\, b_j$$

$$=\;\; b_j\, M_i{}^j\;\, a^i \;=\; (b^T)_j\, (M^T)^j{}_i\;\, a^i \;=\;\; (b^T)_j\, [M^T\mathbf{a}]^j \;\;=\;\; <\mathbf{b}|\, M^T\mathbf{a}> \;=\; <\mathbf{b}|\, \mathcal{M}^T\;\, |\mathbf{a}>\,.$$

Operator $\mathcal{M}$ is defined by its action on an arbitrary ket vector $\mathcal{M}\,|\,\mathbf{b}> = \;|\,M\,\mathbf{b}>$
Operator $\mathcal{M}^T$ is defined by its action on an arbitrary ket vector $\mathcal{M}^T\,|\,\mathbf{b}> = \;|\,M^T\,\mathbf{b}>$
Notice in the proof that the covariant transpose M^T is the correct transpose to use since $M_i{}^j = (M^T)^j{}_i$.

Exercise: Show that **w•v** is a scalar under any transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$:

$$\mathbf{w}'\bullet\mathbf{v}' = <\mathbf{w}'\,|\,\mathbf{v}'> \;= <R\mathbf{w}|R\mathbf{v}> = <\mathbf{w}|\mathcal{R}^T\mathcal{R}|\mathbf{v}> = <\mathbf{w}|\; 1\; |\mathbf{v}> \;= <\mathbf{w}|\mathbf{v}> \;= \mathbf{w}\bullet\mathbf{v}\;\;. \qquad (2.11.g.12)$$

In this example R is a matrix, whereas $\mathcal{R}$ is the corresponding Dirac space operator. The statement

$$\mathcal{R}^T\mathcal{R} = 1 \qquad (2.11.g.13)$$

is the operator version of our (2.11.f.3) matrix statement

$$R^T R = 1 \qquad (2.11.g.14)$$

which we verify as follows,

$$(R^T R)^a{}_c = (R^T)^a{}_b R^b{}_c = R_b{}^a R^b{}_c = \delta^a{}_c \qquad // (2.1.9) \#1 \qquad (2.11.g.15)$$

and which is valid for any transformation differential matrix $R^i{}_j$.

<u>Ways of representing $\mathcal{M}$</u>

One may represent a Dirac operator $\mathcal{M}$ in various ways, for example,

$$\mathcal{M} = \Sigma_{ij} \,|\, \mathbf{u}_i> M^i{}_j <\mathbf{u}^j|$$

$$= \Sigma_{ij} \,|\, \mathbf{e}_i> [M^{[e]}]^i{}_j <\mathbf{e}^j|$$

$$= \Sigma_{ij} \,|\, \mathbf{u}_i> [M^{[u,e]}]^i{}_j <\mathbf{e}^j| \qquad (2.11.g.16)$$

as can be verified by closing with the appropriate basis vectors. For example, for the last line above,

$$<\mathbf{u}^a \,|\, \mathcal{M} \,|\, \mathbf{e}_b> = <\mathbf{u}^a \,|\, \{ \Sigma_{ij} \,|\, \mathbf{u}_i> [M^{[u,e]}]^i{}_j <\mathbf{e}^j| \} \,|\, \mathbf{e}_b>$$

$$= \Sigma_{ij} <\mathbf{u}^a \,|\, \mathbf{u}_i> [M^{[u,e]}]^i{}_j <\mathbf{e}^j| \,\mathbf{e}_b>$$

$$= \Sigma_{ij} \,\delta^a{}_i \,[M^{[u,e]}]^i{}_j \,\delta^j{}_b$$

$$= [M^{[u,e]}]^a{}_b \,. \qquad (2.11.g.17)$$

When $\mathcal{M} = 1$ we find

$$1 = \Sigma_{ij} \,|\, \mathbf{u}_i> \delta^i{}_j <\mathbf{u}^j| = \Sigma_i \,|\, \mathbf{u}_i><\mathbf{u}^i| \qquad (2.11.g.18)$$

which is just a statement that the $|\, \mathbf{u}_i>$ basis is complete, as discussed more below in section (h).

We can compare the abstract Dirac operator $\mathcal{M}$ with the abstract rank-2 "vector" M,

$$\mathcal{M} = \Sigma_{ij} \,|\, \mathbf{u}_i> M^i{}_j <\mathbf{u}^j| \qquad // \text{ Dirac operator}$$

$$M = \Sigma_{ij} M^i{}_j \,\mathbf{u}_i \otimes \mathbf{u}^j \qquad // (2.8.10), \text{"vector" in vector space } V^2$$

or

$$|M> = \Sigma_{ij} M^i{}_j \,|\, \mathbf{u}_i> \otimes \,|\, \mathbf{u}^j> \,. \qquad (2.11.g.19)$$

The first object $\mathcal{M}$ is an *operator* in the Dirac Hilbert Space V.

The second object M or |M> is a *vector* in the tensor product space V ⊗ V.

M and $\mathcal{M}$ are completely different objects, though they both involve the same matrix elements $M^i{}_j$. In each case, we can project out those matrix elements in an appropriate fashion:

$$<\mathbf{u}^a \,|\, \mathcal{M} \,|\, \mathbf{u}_b> \;=\; <\mathbf{u}^a \,|\, \{\ \Sigma_{ij} \,|\, \mathbf{u}_i> M^i{}_j <\mathbf{u}^j|\}\ \,|\, \mathbf{u}_b> \;=\; M^a{}_b$$

$$[< \mathbf{u}_a| \otimes < \mathbf{u}^b| \,]\,|\, M > = [< \mathbf{u}_a| \otimes < \mathbf{u}^b| \,]\, \Sigma_{ij} M^i{}_j \,\,|\, \mathbf{u}_i> \otimes |\, \mathbf{u}^j> \;=\; M^a{}_b \;. \qquad (2.11.g.20)$$

Non-square matrices

The above discussion is presented implicitly for a square matrix M, but only small adjustments are needed for it to apply to a non-square matrix. In this case, in $\mathbf{a}^T M\, \mathbf{b}$ one thinks of vectors **a** and **b** as having different dimensions. Perhaps **b** lies in x-space which is R^n while **a** lies in x'-space which is R^m with m > n, and then $M^i{}_j$ is an m x n matrix. The x'-space V' has n basis vectors $|\mathbf{u'}_i>$ while the x-space V has m basis vectors $|\mathbf{u}_i>$. Then one would have, for example,

$$<\mathbf{u}^i \,|\, \mathcal{M} \,|\, \mathbf{u'}_j> \;=\; M^i{}_j$$

$$\mathcal{M} = \Sigma_{i=1}{}^m\, \Sigma_{j=1}{}^n \,|\, \mathbf{u}_i> M^i{}_j <\mathbf{u'}^j|$$

$$|M> \;=\; \Sigma_{i=1}{}^m\, \Sigma_{j=1}{}^n\, M^i{}_j \,\,|\, \mathbf{u}_i> \otimes |\, \mathbf{u'}^j>$$

$$1' \;=\; \Sigma_i \,|\, \mathbf{u'}_i><\mathbf{u'}^i| \qquad \text{completeness in V'}$$

$$1 \;=\; \Sigma_i \,|\, \mathbf{u}_i><\mathbf{u}^i| \qquad \text{completeness in V} \qquad (2.11.g.21)$$

This is exactly the situation we shall encounter in Chapter 10 where the matrix R is an m x n matrix.

Linearity of $\mathcal{M}$

We emphasize that any Dirac operator like $\mathcal{M}$ is a *linear* operator. This is so because the action of $\mathcal{M}$ is defined in terms of the matrix M which is of course a linear operator. Specifically,

$$\mathcal{M} \,|\, s_1\mathbf{a} + s_2\mathbf{b}> = \,|\, M(s_1\mathbf{a} + s_2\mathbf{b})> \qquad \text{// definition of } \mathcal{M}$$

$$= \,|\, s_1 M\mathbf{a} + s_2\, M\mathbf{b} > \qquad \text{// matrix algebra}$$

$$= \,|\, s_1 M\mathbf{a}> + |\, s_2\, M\mathbf{b} > \qquad \text{// the ket vector space V is a linear space}$$

$$= \, s_1|\, M\mathbf{a}> + s_2|\, M\mathbf{b} > \qquad \text{// the ket vector space V is a linear space}$$

$$= s_1\mathcal{M}|\, \mathbf{a} > + s_2\mathcal{M}|\, \mathbf{b} > \;. \qquad (2.11.g.22)$$

Following the same steps, one finds that M is also linear when it acts to the left on vectors in the dual space V*,

$$< s_1\mathbf{a} + s_2\mathbf{b} \,|\, \mathcal{M} = < M^T(s_1\mathbf{a} + s_2\mathbf{b}) \,| = s_1 <M^T\mathbf{a}\,| + s_2 <M^T\mathbf{b}\,| = s_1 <\mathbf{a}\,|\mathcal{M} + s_2 <\mathbf{b}\,|\mathcal{M} \qquad (2.11.g.23)$$

<u>$\mathcal{M}$ acting on tensor product spaces</u>

Tensor product spaces and wedge product spaces (regular and dual) are described in later Chapters of this document, so our presentation is a little out of order here. We just want to have all material for Dirac operators collected in one place.

It is possible to extend the definition of $\mathcal{M}$ to describe its action on a tensor product space. Suppose |T> and |S> are elements of $V^2 = V\otimes V$ (to be discussed in Section 4.1). Calling this extended operator $\mathcal{M}^{(2)}$, we first define it to be a *linear* operator,

$$\mathcal{M}^{(2)} [\, s_1 |T> + s_2 |S>] \equiv s_1\mathcal{M}^{(2)} |T> + s_2\mathcal{M}^{(2)} |S> . \qquad // \text{ definition}$$

Then we state instructions for how $\mathcal{M}^{(2)}$ acts on a vector in V^2, using a general expansion for |T>,

$$\mathcal{M}^{(2)} |T> = \mathcal{M}^{(2)} [\, \Sigma_{ij}T_{ij} |\mathbf{u}_i> \otimes |\mathbf{u}_j>] = \Sigma_{ij}T_{ij}\, \mathcal{M}^{(2)}[\, |\mathbf{u}_i> \otimes |\mathbf{u}_j>]$$

$$\equiv \Sigma_{ij}T_{ij}\; \mathcal{M}|\mathbf{u}_i> \otimes \mathcal{M}|\mathbf{u}_j> \qquad // \text{ definition}$$

$$= \Sigma_{ij}T_{ij}\; |M\mathbf{u}_i> \otimes |M\mathbf{u}_j> .$$

In general, if $|\,\mathbf{a}>$ and $|\,\mathbf{b}>$ are vectors in V^1, then

$$\mathcal{M}^{(2)} [\, |\,\mathbf{a}> \otimes |\,\mathbf{b}>] \equiv \mathcal{M}|\,\mathbf{a}> \otimes \mathcal{M}|\,\mathbf{b}> = |\,M\mathbf{a}> \otimes |\,M\mathbf{b}>.$$

Normally we write $\mathcal{M}^{(2)}$ just as $\mathcal{M}$ so the nature of $\mathcal{M}$ is implied by the space on which it acts. Then

$$\mathcal{M} [\, s_1 |T> + s_2 |S>] \equiv s_1\mathcal{M} |T> + s_1\mathcal{M} |S>$$

$$\mathcal{M}[\, |\,\mathbf{a}> \otimes |\,\mathbf{b}>] \equiv \mathcal{M}|\,\mathbf{a}> \otimes \mathcal{M}|\,\mathbf{b}> = |\,M\mathbf{a}> \otimes |\,M\mathbf{b}> . \qquad // \mathcal{M} \text{ acting on } V^2 \qquad (2.11.g.24)$$

In generalizing the above equations to the tensor product space $V^n = V\otimes V\otimes ...V$ (Chapter 5), the first equation above stays the same, where then $\mathcal{M}$ on the left side means $\mathcal{M}^{(n)}$, while the second equation changes, so

$$\mathcal{M} [\, s_1 |T> + s_2 |S>] \equiv s_1\mathcal{M} |T> + s_1\mathcal{M} |S>$$

$$\mathcal{M}[\, |\,\mathbf{v}_1> \otimes |\,\mathbf{v}_2> \otimes .... \otimes |\mathbf{v}_n>] \qquad // \mathcal{M} \text{ acting on } V^n \quad [\text{ see } (5.6.17)\,]$$

$$\equiv \mathcal{M}|\,\mathbf{v}_1> \otimes \mathcal{M}|\,\mathbf{v}_2> \otimes .... \otimes \mathcal{M}|\mathbf{v}_n> = |\,M\mathbf{v}_1> \otimes |\,M\mathbf{v}_2> \otimes ....\otimes |M\mathbf{v}_n> . \qquad (2.11.g.25)$$

Parallel statements apply for the dual space V^{*n} (Chapter 6)

$$[\ s_1 \langle T| + s_2 \langle S|\ ]\mathcal{M} \equiv s_1 \langle T|\mathcal{M} + s_1 \langle S|\mathcal{M}$$

$$[\ \langle \mathbf{v}_1| \otimes \langle \mathbf{v}_2| \otimes ....\otimes \langle \mathbf{v}_n|\ ]\ \mathcal{M} \qquad // \ \mathcal{M} \text{ acting on } V^{*n} \text{ [see (6.6.18)]}$$

$$\equiv \langle \mathbf{v}_1|\mathcal{M} \otimes \langle \mathbf{v}_2|\mathcal{M} \otimes .... \otimes \langle \mathbf{v}_n|\mathcal{M} = \langle M^T\mathbf{v}_1| \otimes \langle M^T\mathbf{v}_2| \otimes .... \otimes \langle M^T\mathbf{v}_n| \ . \qquad (2.11.g.26)$$

$\mathcal{M}$ acting on wedge product spaces

As will be shown in Chapter 7, the last two equation sets above have the same form for action on wedge product spaces, but ⊗ is replaced by ^, so

$$\mathcal{M}\ [\ s_1 |T\rangle + s_2 |S\rangle\ ] \equiv s_1\mathcal{M} |T\rangle + s_1\mathcal{M} |S\rangle$$

$$\mathcal{M}[\ |\mathbf{v}_1\rangle \wedge |\mathbf{v}_2\rangle \wedge .... \wedge |\mathbf{v}_n\rangle\ ] \qquad // \ \mathcal{M} \text{ acting on } L^n \text{ [see (7.9.d.15)]}$$

$$\equiv \mathcal{M}|\mathbf{v}_1\rangle \wedge \mathcal{M}|\mathbf{v}_2\rangle \wedge .... \wedge \mathcal{M}|\mathbf{v}_n\rangle = |M\mathbf{v}_1\rangle \wedge |M\mathbf{v}_2\rangle \wedge .... \wedge |M\mathbf{v}_n\rangle \ . \qquad (2.11.g.27)$$

For the dual space Λ^n,

$$[\ s_1 \langle T| + s_2 \langle S|\ ]\mathcal{M} \equiv s_1 \langle T|\mathcal{M} + s_1 \langle S|\mathcal{M}$$

$$[\ \langle \mathbf{v}_1| \wedge \langle \mathbf{v}_2| \wedge ....\wedge \langle \mathbf{v}_n|\ ]\ \mathcal{M} \qquad // \ \mathcal{M} \text{ acting on } \Lambda^n \text{ [see (8.9.d.15)]}$$

$$\equiv \langle \mathbf{v}_1|\mathcal{M} \wedge \langle \mathbf{v}_2|\mathcal{M} \wedge .... \wedge \langle \mathbf{v}_n|\mathcal{M} = \langle M^T\mathbf{v}_1| \wedge \langle M^T\mathbf{v}_2| \wedge .... \wedge \langle M^T\mathbf{v}_n| \ . \qquad (2.11.g.28)$$

We give the $\mathcal{M}$ definitions above for tensor products and wedge products of *vectors*, but the equation numbers "[see (...)]" show the results generalized further to the products of an arbitrary set of *tensors*.

Special Case $\mathcal{R}$

As a special case of the general matrix M and its Dirac linear operator $\mathcal{M}$, we can consider the transformation differential matrix R and its corresponding Dirac operator $\mathcal{R}$, where then $\mathcal{R}|\mathbf{a}\rangle = |R\mathbf{a}\rangle$. The matrix R can be non-square as was noted above for M, and this situation will arise in Chapter 10. The main point is this:

Fact: The operator $\mathcal{R}$ is a *linear* operator with respect to any of the Dirac spaces it acts upon. (2.11.g.29)

These spaces could be V^n, V^{*n}, L^n, Λ^n or any tensor/wedge products of these spaces such as $\Lambda^n \wedge \Lambda^m$. For activities in x'-space, the vector space names are V'^n, V'^{*n}, L'^n, Λ'^n .

(h) Completeness

Let $\mathbf{b_i}$ be a set of basis vectors for vector space V. Then $\mathbf{b^i}$ is the dual basis and we have

$$\delta^i{}_j = \mathbf{b^i} \bullet \mathbf{b_j} = (\mathbf{b^i})^T \mathbf{b_j} = (\,*\ *\,)\begin{pmatrix} * \\ * \end{pmatrix} \qquad (2.11.h.1)$$

where we show the dot product and vector forms of the scalar product.

By the definition of a basis, any set of basis vectors for vector space V is "complete", which means that those vectors are sufficient to expand *any* vector **v** in V,

$$\mathbf{v} = \Sigma_{i=1}{}^n\, v^i \mathbf{b_i} \qquad \text{where} \qquad v^i = \mathbf{b^i} \bullet \mathbf{v} = (\mathbf{b^i})^T\mathbf{v} \ . \qquad (2.11.h.2)$$

One can then write the above equation as

$$\begin{aligned}
\mathbf{v} &= \Sigma_{i=1}{}^n\ \mathbf{b_i}\, [\, v^i\,] & \begin{pmatrix} * \\ * \end{pmatrix} &= \Sigma_{i=1}{}^n \begin{pmatrix} * \\ * \end{pmatrix} [\, v^i\,] \\
&= \Sigma_{i=1}{}^n\, \mathbf{b_i}\ [(\mathbf{b^i})^T\mathbf{v}] & &= \Sigma_{i=1}{}^n \begin{pmatrix} * \\ * \end{pmatrix} [\, (\,*\ *\,)\begin{pmatrix} * \\ * \end{pmatrix}] \\
&= \Sigma_{i=1}{}^n\ [\mathbf{b_i}(\mathbf{b^i})^T]\, \mathbf{v} & &= \Sigma_{i=1}{}^n\ [\, \begin{pmatrix} * \\ * \end{pmatrix}(\,*\ *\,)\,] \begin{pmatrix} * \\ * \end{pmatrix} \\
&= \{\, \Sigma_{i=1}{}^n\ [\mathbf{b_i}(\mathbf{b^i})^T]\}\ \mathbf{v} & &= \{\, \Sigma_{i=1}{}^n\ [\, \begin{pmatrix} * & * \\ * & * \end{pmatrix}\,]\, \} \begin{pmatrix} * \\ * \end{pmatrix} \\
&\equiv M\ \mathbf{v} & &= \begin{pmatrix} * & * \\ * & * \end{pmatrix} \begin{pmatrix} * \\ * \end{pmatrix} \ .
\end{aligned} \qquad (2.11.h.3)$$

On the right we show the vector/matrix structure of each expression in the simple case of R^2. We end up then with **v** = M **v**. Since this must be true for *any* **v** in V, it must be that M = 1, the identity matrix, so

$$\Sigma_{i=1}{}^n\ \mathbf{b_i}(\mathbf{b^i})^T = 1 \qquad (2.11.h.4)$$

which is the official statement that the basis $\mathbf{b_i}$ is "complete". To make this statement in terms of the components of the basis vectors, we can apply $(\mathbf{u_j})^T$ on the left and $(\mathbf{u_k})$ on the right to get

$$(\mathbf{u^j})^T\, [\, \Sigma_{i=1}{}^n\ \mathbf{b_i}(\mathbf{b^i})^T\,]\, \mathbf{u_k} = (\mathbf{u^j})^T\ 1\ \mathbf{u_k} = (\mathbf{u^j})^T\mathbf{u_k} = \mathbf{u^j} \bullet \mathbf{u_k} = \delta^j{}_k$$

or

$$\Sigma_{i=1}{}^n\ (\mathbf{u^j})^T\, [\mathbf{b_i}(\mathbf{b^i})^T]\mathbf{u_k} = \delta^j{}_k$$

or

$$\Sigma_{i=1}{}^n\ [(\mathbf{u^j})^T\, \mathbf{b_i}]\, [(\mathbf{b^i})^T]\mathbf{u_k}] = \delta^j{}_k$$

or

$$\Sigma_{i=1}{}^n\ [\mathbf{u^j} \bullet \mathbf{b_i}]\, [\mathbf{b^i} \bullet \mathbf{u_k}] = \delta^j{}_k$$

or

$$\Sigma_{i=1}{}^n\, (\mathbf{b_i})^j(\mathbf{b^i})_k = \delta^j{}_k \ . \qquad \text{// like (2.3.5)} \qquad (2.11.h.5)$$

We now repeat the above development in the Dirac notation shown on the right,

$$|\mathbf{v}\rangle = \Sigma_{i=1}{}^{n}\, v^i\, |\mathbf{b}_i\rangle \qquad \text{where} \qquad v^i = \langle \mathbf{b}^i|\, \mathbf{v}\rangle \tag{2.11.h.6}$$

$$\begin{aligned}
\mathbf{v} &= \Sigma_{i=1}{}^{n}\ \mathbf{b}_i\, [\, v^i\,] & |\mathbf{v}\rangle &= \Sigma_{i=1}{}^{n}\, |\mathbf{b}_i\rangle\, [\, v^i\,] \\
&= \Sigma_{i=1}{}^{n}\, \mathbf{b}_i\ [(\mathbf{b}^i)^T\mathbf{v}] & &= \Sigma_{i=1}{}^{n}\ |\mathbf{b}_i\rangle\, [\ \langle\mathbf{b}^i\,|\ \mathbf{v}\rangle\,] \\
&= \Sigma_{i=1}{}^{n}\ [\mathbf{b}_i(\mathbf{b}^i)^T]\, \mathbf{v} & &= \Sigma_{i=1}{}^{n}\ [\ |\mathbf{b}_i\rangle\langle\mathbf{b}^i|\ ]\, |\,\mathbf{v}\rangle \\
&= \{\, \Sigma_{i=1}{}^{n}\ [\mathbf{b}_i(\mathbf{b}^i)^T]\}\ \mathbf{v} & &= \{\, \Sigma_{i=1}{}^{n}\, |\mathbf{b}_i\rangle\langle\mathbf{b}^i|\, \}\, |\,\mathbf{v}\rangle\, .
\end{aligned} \tag{2.11.h.7}$$

Completeness expressed in Dirac notation is then

$$\Sigma_{i=1}{}^{n}\, |\mathbf{b}_i\rangle\langle\mathbf{b}^i| = \mathit{1}\ . \tag{2.11.h.8}$$

where *1* is the Dirac operator form of the matrix identity matrix 1.

Applying $\langle\mathbf{u}^j|$ on the left and $|\mathbf{u}_k\rangle$ on the right this becomes

$$\Sigma_{i=1}{}^{n}\, \langle\mathbf{u}^j|\ \mathbf{b}_i\rangle\langle\mathbf{b}^i|\mathbf{u}_k\rangle = \langle\mathbf{u}^j|\mathbf{u}_k\rangle = \delta^i{}_k$$

or

$$\Sigma_{i=1}{}^{n}\, (\mathbf{b}_i)^j(\mathbf{b}^i)_k = \delta^j{}_k \tag{2.11.h.9}$$

which replicates (2.11.h.5)

3. Outer Products and Kronecker Products

3.1 Outer Products Reviewed: Compatibility of Chapter 1 and Chapter 2

Vectors were unbolded in Chapter 1, but were bolded for clarity in Chapter 2. Here we express all vectors in unbolded notation. Also, we quietly switch from contravariant (upper) to covariant (lower) tensor indices.

Chapter 1 developed the idea of the tensor product space V⊗W with elements v⊗w which satisfy a set of bilinear rules (1.1.5),

$$(v_1+v_2) \otimes w = (v_1\otimes w) + (v_2\otimes w) \qquad \text{for all } v_1,v_2 \in V \text{ and all } w \in W \qquad \textit{(1.1.5)}$$
$$v \otimes (w_1+w_2) = (v\otimes w_1) + (v\otimes w_2) \qquad \text{for all } v \in V \text{ and all } w_1, w_2 \in W$$
$$s(v\otimes w) = (sv)\otimes w = v\otimes(sw) \qquad \text{for all } v \in V \text{ and all } w \in W \text{ and all } s \in K \qquad (3.1.1)$$

The essence of the tensor product is this bilinearity, and there is no requirement to describe the objects v⊗w in more detail. In the formal sense we are done and fini. However, for "engineering purposes", it is useful to add more structure to the tensor product by defining "tensor components", and that was the subject of Chapter 2. We conjured up a way to add components to the theory by defining tensor product components in terms of the outer product of two vectors,

$$(v\otimes w)_{ij} \equiv v_i w_j \qquad v \in V \quad w \in W . \qquad \textit{(2.8.10)} \qquad (3.1.2)$$

The components v_i and w_j can be elements of any field K and the juxtaposition of $v_i w_j$ implies multiplication in that field (we have in mind that K = R, the real numbers).

The key point: because the function $v_i w_j$ is manifestly bilinear, this extra specification does not conflict with any of the earlier tensor product "rules". For example we can evaluate,

$$[(v_1+v_2) \otimes w]_{ij} = (v_1+v_2)_i w_j$$

$$(v_1\otimes w)_{ij} + (v_2\otimes w)_{ij} = (v_1)_i\, w_j + (v_2)_i\, w_j . \qquad (3.1.3)$$

The first ⊗ rule of (3.1.1) says the left sides of these two equations must be equal, but we can see that the right sides are also equal, so our "tensor componentization" does not conflict with the no-components theory of Chapter 1. Thus it is that we simply glom this component structure onto the tensor product concepts of Chapter 1.

We extended the tensor product idea to include the tensor product of k spaces V⊗W⊗...⊗Z with this associated set of k-multilinear rules,

$$(s_1v_1+s_2v_2)\otimes w\otimes \;.....\; \otimes z = s_1\,(v_1\otimes w\otimes \;....\; \otimes z) + s_2\,(v_2\otimes w\otimes \;....\; \otimes z)$$
$$v \otimes (s_1w_1+s_2w_2)\otimes \;...\otimes z = s_1\,(v\otimes w_1\otimes \;....\; \otimes z) + s_2\,(v\otimes w_2\otimes \;....\; \otimes z).$$
etc. *(1.1.16)* (3.1.4)

Onto this skeleton we hang a component structure again using the outer product of vectors,

$$(v\otimes w\otimes z\otimes ....)_{ijk....} = v_i w_j z_k.... \qquad \textit{(2.8.18)} \qquad (3.1.5)$$

The function $v_i w_j z_k$.... is manifestly k-multilinear, so this structural enhancement is compatible with the general theory of Chapter 1.

We have tried to keep things general up this point by using V⊗W⊗...⊗Z where all the vector spaces can be different, but now we assume they are all the same,

$$V^k \equiv V\otimes V\otimes ....\otimes V \qquad \text{// k copies, fancy notation } {}^{\otimes}\Pi_{i=1}{}^{k}\, V \qquad \textit{(2.11.e.1)} \qquad (3.1.6)$$

and this is our main interest, since the elements are then true "tensors" in the sense of Chapter 2. There is then only one set of basis functions $\{e_i\}$ to worry about, the basis for V. In this case, if a and b transform as vectors, then a⊗b transforms as a rank-2 tensor and thus provides a name for the outer product tensor whose components are $a_i b_j$.

It has already been shown in Section 2.8 (in the components world) how the ⊗ product can combine tensors into tensors of higher rank using the outer product idea. We had for example for the combination of a vector with two rank-2 tensors, the following rank-5 tensor

$$(K\otimes K\otimes \mathbf{v})^{abcde} = K^{ab}K^{cd}\, v^e \ . \qquad \textit{(2.8.7)} \qquad (3.1.7)$$

Another example would be this,

$$\begin{aligned} A &= a\otimes b \qquad & A_{ij} &= (a\otimes b)_{ij} = a_i b_j \\ B &= c\otimes d \qquad & B_{ij} &= (c\otimes d)_{ij} = c_i d_j \end{aligned} \qquad (3.1.8)$$

One can then define the tensor product of A and B in a fairly obvious manner,

$$A\otimes B \equiv (a\otimes b)\otimes(c\otimes d) = a\otimes b\otimes c\otimes d \in V^4 \ . \qquad \text{// associative (2.8.22)} \qquad (3.1.9)$$

This equation has no indices and so is acceptable in the component-free world of Chapter 1. The components are then taken in the following obvious manner,

$$[A\otimes B]_{ijk\ell} = [\, a\otimes b\otimes c\otimes d]_{ijk\ell} = a_i b_j c_k d_\ell = A_{ij} B_{k\ell}\, . \qquad (3.1.10)$$

The above lines shows that A⊗B is in fact a rank-4 tensor constructed by taking the outer product of two rank-2 tensors (or the outer product of four rank-1 tensors). In Section 3.2 we shall have use for an object defined in this strange manner

$$[A\otimes B]_{ik,j\ell} \equiv [A\otimes B]_{ijk\ell} = A_{ij} B_{k\ell} \qquad (3.1.11)$$

and we just mention it here in passing. Note that the indices are shuffled relative to the LHS of (3.1.10).

Using the same method as above, one can construct a rank-6 tensor from the tensor product of three rank-2 tensors,

$$[A\otimes B\otimes C]_{abcdef} = A_{ab}B_{cd}C_{ef} \qquad (3.1.12)$$

or from the tensor product of two rank-3 tensors

$$[A\otimes B]_{abcdef} = A_{abc}B_{def} \,. \qquad (3.1.13)$$

In general, one can take the tensor product of any set of tensors to create a new tensor whose rank is the sum of the ranks of the tensors that were combined by the ⊗ symbol. If A,B,C... are arbitrary tensors, having multiindices I,J,K (for example I = $\{i_1,i_2,i_3\}$ if A is rank-3), one could write a general formula for the components of the tensor product of any number of pure tensor objects in this manner,

$$(A \otimes B \otimes C \otimes ....)_{IJK...} = A_I\, B_J\, C_K...... \qquad \text{// outer product} \qquad (3.1.14)$$

The tensor here is A ⊗ B ⊗ C ⊗ ...., and it is the tensor product and the outer product of the individual tensors A,B,C.... The equation specifies its components.

3.2 Kronecker Products

The subject here is the tensor product of two linear *operators* and is included here because it seems therefore to fit into the topic of "tensor products". This is a stand-alone section and nothing in it is referenced in later sections of our document. For that reason, a reader uninterested in Kronecker Products would do well to skip this section and continue into Chapter 4 on the wedge product development. The energetic reader can regard this section as an exercise in using the covariant tensor product machinery of Chapter 2.

Let V and X be vector spaces of dimension n and m. Basis(V) = u_i Basis(X) = u_i
Let W and Y be vector spaces of dimension n' and m' Basis(W) = u'_i Basis(Y) = u'_i . (3.2.1)

We imagine that vector spaces V,X,W,Y have metric tensors g, *g*, g', *g*' which can be used to raise and lower subscripts in the standard manner shown in (2.2.1). Often one assumes that all these spaces have a Cartesian metric tensor, so up and down indices are the same, but we shall carry out the development below in full covariant notation as part of our "exercise".

Rather than use Einstein implied sums, we shall display all sums explicitly in this section.

Consider linear operators S and T such that,

$$x = Sv = \text{a vector in X} \qquad S: V\rightarrow X \qquad x^i = \Sigma_{a=1}^{n}\, S^i{}_a v^a \qquad i = 1,2..m$$
$$y = Tw = \text{a vector in Y} \qquad T:W\rightarrow Y \qquad y^j = \Sigma_{b=1}^{n'}\, T^j{}_b w^b \qquad j = 1,2..m' \,. \qquad (3.2.2)$$

Notice that on $S^i{}_a$ the first index is an X-space index which can be raised and lowered by metric tensor g, whereas the second index on $S^i{}_a$ is a V-space index which can be raised and lowered by g. So we can regard $S^i{}_a$ as the components of a "cross tensor" involving the spaces X and V. In any equation below, we are free to change the "tilt" of any contracted index pair in the manner of (2.9.1) because such tilted index pairs will always be associated with the same metric tensor. Similar comments apply to $T^j{}_b$.

The linear operator S is represented by matrix $S^i{}_a$ which has m rows and n columns (m x n).
The linear operator T is represented by matrix $T^j{}_b$ which has m' rows and n' columns (m' x n').

We want to create a meaning for S⊗T which is the tensor product of these two operators S and T.
A candidate definition for this meaning is the following,

$$(S\otimes T)(v\otimes w) = (Sv)\otimes(Tw) \;. \qquad // = (x\otimes y) \qquad S\otimes T : V\otimes W \rightarrow X\otimes Y \;. \qquad (3.2.3)$$

$$(S\otimes T)|\, v\otimes w\rangle = (S\otimes T)\, |v\rangle \otimes w\rangle = S|v\rangle \otimes T|w\rangle = |x\rangle \otimes |y\rangle = |x\otimes y\rangle \qquad // \text{ Dirac notation}$$

Consider the following processing steps,

$$(S\otimes T)([\alpha v_1 + \beta v_2]\otimes w) = (S[\alpha v_1 + \beta v_2])\otimes(Tw) \qquad // \; (3.2.3)$$

$$= (\alpha\, Sv_1 + \beta Sv_2) \otimes (Tw) \qquad // \; S:V\rightarrow X \text{ is linear}$$

$$= \alpha\, (Sv_1)\otimes(Tw) + \beta(Sv_2)\otimes(Tw) \qquad // \text{ using the first } \otimes \text{ rule in } (3.1.1)$$

$$= \alpha\, (S\otimes T)(v_1\otimes w) + \beta\, (S\otimes T)(v_2\otimes w) \;. \qquad // \; (3.2.3) \text{ used twice} \qquad (3.2.4)$$

This shows that (S⊗T)(v⊗w) is linear in v. A similar argument shows it is also linear in w. Thus, the operator (S⊗T) as defined above is a bilinear operator on V⊗W, and we confirm the essential characteristic of the tensor product, which is its bilinearity. We accept the candidate definition (3.2.3).

<u>Exercise:</u> Compute the action of (S⊗T) on a general element of V⊗W .

Apply (S⊗T) to a general element of V⊗W using tensor expansion like (2.10.3b) and then (3.2.3),

$$(S\otimes T)[\, \Sigma_{ij}F^{ij}\, u_i\otimes u'_j] = \Sigma_{ij}F^{ij}\, (S\otimes T)(u_i\otimes u'_j) = \Sigma_{ij}F^{ij}\, (Su_i)\otimes(Tu'_j) \;. \qquad (3.2.5)$$

The action of S on a vector v (and T on w) can be written

$$x = (Sv) = \Sigma_a[Sv]^a u_a = \Sigma_a(\Sigma_b S^a{}_b v^b) u_a = \Sigma_{ab}(S^a{}_b v^b) u_a$$

$$y = (Tw) = \Sigma_c[Tw]^c u'_c = \Sigma_c(\Sigma_d T^c{}_d w^d) u'_c = \Sigma_{cd}(T^c{}_d w^d) u'_c \;. \qquad (3.2.6)$$

Select $v = u^i$ and $w = u'^j$ in these last two equations to get,

$$(Su_i) = \Sigma_{ab} S^a{}_b (u_i)^b u_a$$
$$(Tu'_j) = \Sigma_{cd} T^c{}_d (u'_j)^d u'_c \quad . \qquad (3.2.7)$$

Then the tensor product appearing in (3.2.5) can be written

$$(Su_i)\otimes(Tu'_j) = [\, \Sigma_{ab} S^a{}_b (u_i)^b u_a] \otimes [\, \Sigma_{cd} T^c{}_d (u'_j)^d u'_c]$$

$$= \Sigma_{abcd}\, S^a{}_b (u_i)^b T^c{}_d (u'_j)^d (u_a \otimes u'_c) \qquad (3.2.8)$$

and so the action of the tensor product operator (S⊗T) is given by,

$$(S\otimes T)[\, \Sigma_{ij} F^{ij}\, u_i \otimes u'_j] = \Sigma_{ij} F^{ij}\, (Su_i)\otimes(Tu'_j) \qquad // \ (3.2.3)$$

$$= \Sigma_{ijabcd}\, F^{ij} S^a{}_b (u_i)^b T^c{}_d (u'_j)^d (u_a \otimes u'_c) \qquad // \ (3.2.8)$$

$$= \Sigma_{ac} \{\, \Sigma_{ijbd}\, F^{ij} S^a{}_b (u_i)^b T^c{}_d (u'_j)^d \,\} (u_a \otimes u'_c) \qquad // \text{ regroup}$$

$$= \Sigma_{ac}\, G^{ac} (u_a \otimes u'_c) \qquad \text{where} \qquad G^{ac} = \Sigma_{ijbd}\, F^{ij} S^a{}_b (u_i)^b T^c{}_d (u'_j)^d \quad . \qquad (3.2.9)$$

We have then shown the action of operator S⊗T on a general element of V⊗W :

$$(S\otimes T)\ \{\, \Sigma_{ij} F^{ij}\, u_i \otimes u'_j \,\} = \Sigma_{ac}\, G^{ac} (u_a \otimes u'_c) \qquad (S\otimes T) : V\otimes W \ \rightarrow\ X\otimes Y$$

$$\text{where} \quad G^{ac} = \Sigma_{ijbd}\, F^{ij}\ S^a{}_b\ (u_i)^b\ T^c{}_d\ (u'_j)^d \quad . \qquad (3.2.10)$$

It is useful now to consider the component analysis of the action of S⊗T on a pure element of V⊗W in the sense of outer products. Then

$$(x\otimes y) = (S\otimes T)(v\otimes w) = (Sv)\otimes(Tw) \qquad \textit{(3.2.3)}$$

so

$$(x\otimes y)^{ii'} = [(S\otimes T)(v\otimes w)]^{ii'} = [(Sv)\otimes(Tw)]^{ii'} \quad . \qquad (3.2.11)$$

The right side of this last equation can be expanded using (3.1.2) and (3.2.2) to get

$$[(Sv)\otimes(Tw)]^{ii'} = (Sv)^i (Tw)^{i'} = (\Sigma_j\, S^i{}_j v^j)(\Sigma_{j'} T^{i'}{}_{j'} w^{j'})$$

$$= \Sigma_{jj'}\, S^i{}_j T^{i'}{}_{j'}\, v^j w^{j'} = \Sigma_{jj'}\, S^i{}_j T^{i'}{}_{j'}\, (v\otimes w)^{jj'} \quad . \qquad // = (x\otimes y)^{ii'} \qquad (3.2.12)$$

so then (3.2.11) may be written

$$[(S\otimes T)(v\otimes w)]^{ii'} = \Sigma_{jj'}\, [\, S^i{}_j T^{i'}{}_{j'}]\ (v\otimes w)^{jj'} \quad . \qquad // = (x\otimes y)_{ii'} \qquad (3.2.13)$$

We now define

$$(S\otimes T)^{ii'}{}_{,jj'} \equiv S^i{}_j T^{i'}{}_{j'} \qquad (3.2.14)$$

The comma is used to distinguish the left side from the rank-4 tensor $(S\otimes T)^{ii'}{}_{jj'} = S^{ii'}T_{jj'}$ which is a different animal.

Since S and T are (cross) tensors, we can raise and lower indices on the right side of (3.2.14) using the appropriate metric tensors as discussed below (3.2.1), and then the left side indices follow since this is a definition. For example.

$$(S\otimes T)_{ii',jj'} \equiv S_{ij}T_{i'j'} \ . \qquad (3.2.15)$$

This definition was mentioned in (3.1.11) where it was compared to the usual notation used for a rank-4 outer product tensor $(S\otimes T)_{iji'j'} = S_{ij}T_{i'j'}$. In (3.2.15) the two first indices of S and T are listed before the comma while the two second indices appear after the comma.

Installing (3.2.14) into (3.2.13), one gets

$$[(S\otimes T)(v\otimes w)]^{ii'} = \Sigma_{jj'}(S\otimes T)^{ii'}{}_{,jj'}\,(v\otimes w)^{jj'} \ . \qquad // = (x\otimes y)^{ii'} \qquad (3.2.16)$$

The structure of this equation suggests that we are multiplying a vector (v⊗w) by a matrix (S⊗T), but the usual summation index is replaced by two summation indices j and j'. In a multiindex notation one might write the above as

$$x^I = [(S\otimes T)(v\otimes w)]^I = \Sigma_J\ (S\otimes T)^I{}_J\,(v\otimes w)^J \ . \qquad I = \{i,i'\} \qquad J = \{j,j'\} \qquad (3.2.17)$$

Is there some way to write S⊗T as a standard matrix with two indices instead of four?

Start with (3.2.16) written as

$$(x\otimes y)^{ii'} = \Sigma_{jj'}\,(S\otimes T)^{ii'}{}_{,jj'}\,(v\otimes w)^{jj'}$$

or

$$(x^i y^{i'}) = \Sigma_{jj'}\ (S\otimes T)^{ii'}{}_{,jj'}\ (v^j w^{j'}). \qquad (S\otimes T)^{ii'}{}_{,jj'} = (S^i{}_j T^{i'}{}_{j'}) \ . \qquad (3.2.18)$$

We want to write this somehow in a form

$$q'_r = \Sigma_s\ M_{rs}\, q_s \ . \qquad (3.2.19)$$

For illustration purposes, assume n = 2 and n' = 3. Then write the components $(v^j w^{j'})$ as a single column vector in this obvious manner, where the w component index moves fastest,

$$\begin{pmatrix} v^1w^1 \\ v^1w^2 \\ v^1w^3 \\ v^2w^1 \\ v^2w^2 \\ v^2w^3 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \end{pmatrix} = q \text{ with components } q_s \text{ where } s = 1,2....n*n' \quad . \tag{3.2.20}$$

If $v^j w^{j'} \rightarrow q_s$, one can compute s from j,j' as follows: (here 3 = n' = dim(W) for this special case)

$$s = (j-1)3 + j' \quad \Rightarrow \quad (s-1) = (j-1)3 + (j'-1) \quad \Rightarrow \quad \frac{s-1}{3} = (j-1) + \frac{j'-1}{3}$$

$$\Rightarrow \ \text{int}(\frac{s-1}{3}) = j-1 \quad \text{and} \quad \text{rem}\,(\frac{s-1}{3}) = j'-1 \ . \tag{3.2.21}$$

Thus for general n' we can compute j and j' from s in this way (integer part and remainder)

$$j = 1+\text{int}(\frac{s-1}{n'}) \qquad j' = 1+\text{rem}(\frac{s-1}{n'}) \qquad s = 1,2....n*n' \quad . \tag{3.2.22}$$

One can similarly consider $x^i y^{i'} \rightarrow q'_r$ where the column vector q' has m*m' components. The rules here are analogous to those above,

$$i = 1+\text{int}(\frac{r-1}{m'}) \qquad i' = 1+\text{rem}(\frac{r-1}{m'}) \qquad r = 1,2...m*m' \quad . \tag{3.2.23}$$

Therefore, comparing (3.2.19) and (3.2.18), the desired M_{rs} is given by

$$M_{rs} = (S\otimes T)^{ii'}{}_{,jj'} = S^i{}_j T^{i'}{}_{j'} \quad \text{where}$$

$$i = 1+\text{int}(\frac{r-1}{m'}) \qquad j = 1+\text{int}(\frac{s-1}{n'}) \qquad s = 1,2....n*n'$$

$$i' = 1+\text{rem}(\frac{r-1}{m'}) \qquad j' = 1+\text{rem}(\frac{s-1}{n'}) \qquad r = 1,2...m*m' \ . \tag{3.2.24}$$

Thus we have reconfigured our multi-index equation $x^I = \Sigma_J\ (S\otimes T)^I{}_{,J}\ (v\otimes w)^J$ into an ordinary matrix equation $q'_r = \Sigma_s\ M_{rs}\, q_s$ where M_{rs} is given as stated above.

This matrix $M_{rs} = (S\otimes T)^{ii'}{}_{,jj'} = S^i{}_j T^{i'}{}_{j'}$ is known as the **Kronecker product** of the matrices S and T. The subscripts i,i',j'j' are all functions of r and s as shown in (3.2.24).

Symbolically we write this Kronecker product as M = S⊗T. Normally in writing M = S⊗T one would imply $M^a{}_b{}^c{}_d = S^a{}_b T^c{}_d$ which is unrelated to the Kronecker product.

It is a bit tedious to compute and display one of these M matrices by hand, so we let Maple do it for us. For this example we use the following dimensions m, n, m', n' for the spaces X, V, Y, W :

$$S = m \times n \ = 2 \times 3 \qquad \text{rows} = m*m' = 6$$

$$T = m' \times n' \ = 3 \times 4 \qquad \text{cols} = n*n' = 12 \tag{3.2.25}$$

The code simply does what (3.2.24) says to do:

```
restart;
m := 2:    n := 3:
mp := 3:   np := 4:
smax := n*np:    rmax := m*mp:
i  := 1 + iquo(r-1,mp):   # 1 + int((r-1)/mp))
ip := 1 + irem(r-1,mp):   # 1 + rem((r-1)/mp))
j  := 1 + iquo(s-1,np):   # 1 + int((s-1)/np))
jp := 1 + irem(s-1,np):   # 1 + rem((s-1)/np))
M := matrix(rmax,smax):
for r from 1 to rmax do
   for s from 1 to smax do
      M[r,s] := S[i,j]*T[ip,jp];
   od;
od:
print(M);
```

(3.2.26)

$$\begin{bmatrix} S_{1,1}T_{1,1} & S_{1,1}T_{1,2} & S_{1,1}T_{1,3} & S_{1,1}T_{1,4} & S_{1,2}T_{1,1} & S_{1,2}T_{1,2} & S_{1,2}T_{1,3} & S_{1,2}T_{1,4} & S_{1,3}T_{1,1} & S_{1,3}T_{1,2} & S_{1,3}T_{1,3} & S_{1,3}T_{1,4} \\ S_{1,1}T_{2,1} & S_{1,1}T_{2,2} & S_{1,1}T_{2,3} & S_{1,1}T_{2,4} & S_{1,2}T_{2,1} & S_{1,2}T_{2,2} & S_{1,2}T_{2,3} & S_{1,2}T_{2,4} & S_{1,3}T_{2,1} & S_{1,3}T_{2,2} & S_{1,3}T_{2,3} & S_{1,3}T_{2,4} \\ S_{1,1}T_{3,1} & S_{1,1}T_{3,2} & S_{1,1}T_{3,3} & S_{1,1}T_{3,4} & S_{1,2}T_{3,1} & S_{1,2}T_{3,2} & S_{1,2}T_{3,3} & S_{1,2}T_{3,4} & S_{1,3}T_{3,1} & S_{1,3}T_{3,2} & S_{1,3}T_{3,3} & S_{1,3}T_{3,4} \\ S_{2,1}T_{1,1} & S_{2,1}T_{1,2} & S_{2,1}T_{1,3} & S_{2,1}T_{1,4} & S_{2,2}T_{1,1} & S_{2,2}T_{1,2} & S_{2,2}T_{1,3} & S_{2,2}T_{1,4} & S_{2,3}T_{1,1} & S_{2,3}T_{1,2} & S_{2,3}T_{1,3} & S_{2,3}T_{1,4} \\ S_{2,1}T_{2,1} & S_{2,1}T_{2,2} & S_{2,1}T_{2,3} & S_{2,1}T_{2,4} & S_{2,2}T_{2,1} & S_{2,2}T_{2,2} & S_{2,2}T_{2,3} & S_{2,2}T_{2,4} & S_{2,3}T_{2,1} & S_{2,3}T_{2,2} & S_{2,3}T_{2,3} & S_{2,3}T_{2,4} \\ S_{2,1}T_{3,1} & S_{2,1}T_{3,2} & S_{2,1}T_{3,3} & S_{2,1}T_{3,4} & S_{2,2}T_{3,1} & S_{2,2}T_{3,2} & S_{2,2}T_{3,3} & S_{2,2}T_{3,4} & S_{2,3}T_{3,1} & S_{2,3}T_{3,2} & S_{2,3}T_{3,3} & S_{2,3}T_{3,4} \end{bmatrix}$$

(3.2.27)

One should interpret each matrix element of the form $S_{ab}T_{cd}$ as $S^a{}_bT^c{}_d$ -- we don't know how to make Maple display things this way. If all metric tensors are Cartesian, then (3.2.27) is correct as is.

Staring at the above matrix, one can see that the T submatrix is repeated six times, and one can write this matrix in a shorthand notation as

$$M = \begin{pmatrix} S^1{}_1T & S^1{}_2T & S^1{}_3T \\ S^2{}_1T & S^2{}_2T & S^2{}_3T \end{pmatrix} \qquad \text{where } T = \begin{pmatrix} T^1{}_1 & T^1{}_2 & T^1{}_3 & T^1{}_4 \\ T^2{}_1 & T^2{}_2 & T^2{}_3 & T^2{}_4 \\ T^3{}_1 & T^3{}_2 & T^3{}_3 & T^3{}_4 \end{pmatrix} . \qquad (3.2.28)$$

This provides an easy way to manually construct such matrices. This construction can be understood if we look back at the M matrix definition,

$$M_{rs} = (S\otimes T)^{ii'}{}_{,jj'} = S^i{}_jT^{i'}{}_{j'} \quad \text{where}$$

$$i = 1+\text{int}(\frac{r-1}{m'}) \qquad j = 1+\text{int}(\frac{s-1}{n'}) \qquad s = 1,2....n*n'$$

$$i' = 1+\text{rem}(\frac{r-1}{m'}) \qquad j' = 1+\text{rem}(\frac{s-1}{n'}) \qquad r = 1,2...m*m' . \qquad \textit{(3.2.24)}$$

The indices i,j on S select a rectangular subregion of the M matrix due to their integer part definitions. Then within each subregion the i'j' indices run through their full ranges so a copy of matrix T appears in that subregion, multiplied by the $S^i{}_j$ for that subregion.

One is commonly interested in the case where

$$\begin{array}{ll} S: V\rightarrow V & S = n \times n \text{ matrix} \\ T: W\rightarrow W & T = n' \times n' \text{ matrix} \end{array} \qquad (3.2.29)$$

With n = m = 2 and n' = m' = 2 the above code generates this matrix M,

$$\begin{bmatrix} S_{1,1}T_{1,1} & S_{1,1}T_{1,2} & S_{1,2}T_{1,1} & S_{1,2}T_{1,2} \\ S_{1,1}T_{2,1} & S_{1,1}T_{2,2} & S_{1,2}T_{2,1} & S_{1,2}T_{2,2} \\ S_{2,1}T_{1,1} & S_{2,1}T_{1,2} & S_{2,2}T_{1,1} & S_{2,2}T_{1,2} \\ S_{2,1}T_{2,1} & S_{2,1}T_{2,2} & S_{2,2}T_{2,1} & S_{2,2}T_{2,2} \end{bmatrix} \qquad (3.2.30)$$

which can be compared with a result quoted on the (current) wiki tensor product page.

<u>Some other properties</u>

Suppose S_1 and S_2 are two matrices of the same dimension as S, and T_1 and T_2 are two matrices of the same dimension as T. Recall from above,

$$(S\otimes T)^{ii'}{}_{,jj'} \equiv S^i{}_j T^{i'}{}_{j'} \quad . \qquad \textit{(3.2.14)}$$

It then follows that

$$\begin{aligned} (S_1 + S_2) \otimes T &= S_1\otimes T + S_2\otimes T \\ S \otimes (T_1 + T_2) &= S\otimes T_1 + S\otimes T_2 \\ (S_1 + S_2) \otimes (T_1 + T_2) &= S_1\otimes T_1 + S_2\otimes T_1 + S_2\otimes T_1 + S_2\otimes T_2 \end{aligned} \qquad (3.2.31)$$

which is just the statement that S ⊗ T is a bilinear operator. To prove the first line use (3.2.14),

$$\begin{aligned} [(S_1 + S_2) \otimes T]^{ii'}{}_{,jj'} &= (S_1 + S_2)^i{}_j T^{i'}{}_{j'} = (S_1)^i{}_j T^{i'}{}_{j'} + (S_2)^i{}_j T^{i'}{}_{j'} \\ &= (S_1\otimes T)^{ii'}{}_{,jj'} + (S_2\otimes T)^{ii'}{}_{,jj'} \quad . \end{aligned} \qquad (3.2.32)$$

Suppose S_1 and S_2 are both n x n and T_1 and T_2 are both n' x n' . Then one can write,

$$(S_1S_2)\otimes(T_1T_2) = (S_1\otimes T_1)(S_2\otimes T_2) \qquad (3.2.33)$$

Proof: Again use (3.2.14),

$$[(S_1S_2)\otimes(T_1T_2)]^{ii'}{}_{,jj'} = (S_1S_2)^i{}_j(T_1T_2)^{i'}{}_{j'}$$

$$= (S_1)^i{}_k (S_2)^k{}_j (T_1)^{i'}{}_{k'} (T_2)^{k'}{}_{j'} = (S_1)^i{}_k (T_1)^{i'}{}_{k'} (S_2)^k{}_j(T_2)^{k'}{}_{j'}$$

$$= (S_1\otimes T_1)^{ii'}{}_{kk'} (S_2\otimes T_2)^{kk'}{}_{jj'} \qquad (3.2.34)$$

so that in multiindex notation,

$$[(S_1S_2)\otimes(T_1T_2)]^I{}_J = (S_1\otimes T_1)^I{}_K(S_2\otimes T_2)^K{}_J \ . \qquad (3.2.35)$$

4. The Wedge Product of 2 vectors built on the Tensor Product

We now back up and reconsider the space V⊗W and its elements v⊗w. The goal of the next two sections is to establish the parallelism between the vector space V⊗W and the "dual" vector space V*⊗W*. Some repetition is used to review and reinforce earlier stated facts. Then Sections 4.3 and 4.4 introduce the wedge product developed in a similar parallel fashion.

At the end of each of the four sections below a selection of equations is re-expressed in Dirac notation.

4.1 The tensor product of 2 vectors in V^2

<u>Note</u>: The u'_i used below are unrelated to the u'_i of Chapter 2.

<u>Basics</u>. Consider two vector spaces V and W (defined over field K) of dimension n and n'. Let

$\{u_i\}$ = basis of V   dim(V) = n $v = \Sigma_{i=1}^{n} v^i u_i$ = general vector in V $v^i \in K$
$\{u'_i\}$ = basis of W   dim(W) = n' $w = \Sigma_{j=1}^{n'} w^j u'_j$ = general vector in W $w^j \in K$

$u_i = |u_i>$ $u'_i = |u'_i>$ // Dirac notation

$\{u_i \otimes u'_j\}$ = basis for the tensor product space V⊗W   dim(V⊗W) = n*n'

v⊗w = a pure "vector" in the tensor product space V⊗W   v⊗w ≠ w⊗v if v ≠ w

$\otimes : V\text{x}W \rightarrow V\otimes W \qquad \otimes : (v,w) \mapsto v\otimes w$ (4.1.1)

The last line shows ⊗ as a mapping → between two sets, while ↦ shows how set elements map.

Note that v⊗w ≠ w⊗v. For V≠W, w⊗v does not even make sense since that requires w ∈ V and v ∈ W. For V = W the objects v⊗w and w⊗v are still different unless v = w.

<u>Outer Product Revisited.</u> The notion of an outer product was discussed in Sections 2.8 and 3.1. We had for example (where a_i and b_j are the covariant components of vectors a and b both ∈ V),

$(a \otimes b)_{ij} = a_i b_j$ // outer product of two vectors   *(3.1.8)*

$(A \otimes B)_{abcd} = A_{ab}B_{cd}$.   // outer product of two rank-2 tensors   *(3.1.10)*

The "outer product" of two vectors a and b may be written in vector/matrix notation as follows,

$$(a \otimes b)_{**} = ab^T = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix} (\, b_1 .\, b_2 ....b_n) = \begin{pmatrix} a_1b_1 & a_1b_2 & ... \\ a_2b_1 & a_2b_2 & ... \\ ... & ... & ... \end{pmatrix} \qquad (a \otimes b)_{ij} = (ab^T)_{ij} \qquad (4.1.2)$$

The same vector/matrix notation used above can also be used to express the "inner product" (dot product) appearing in (2.2.5), with the caveat noted below,

$$a \bullet b = a^T b = (a^1 . a^2 .... a^n) \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix} = \Sigma_{k=1}^n a^k b_k = \langle a \mid b \rangle \qquad (4.1.3)$$

If the a components are contravariant, the b components must be covariant, and vice versa.

Chapter 1 Tensor Product Revisited. By convention one represents an element of a tensor product space using the ⊗ symbol. It is a certain kind of "product" between a vector in one vector space and a vector in another vector space. On can treat ⊗ as an operator $\otimes : VxW \rightarrow (V\otimes W)$ in the sense that

$\otimes(v,w) = (v) \otimes (w) = (v\otimes w)$ = element of tensor product space $(V\otimes W)$.

Certain ⊗ rules were declared in (1.1.5) which make the tensor product space be a vector space, and which in an intuitive sense just seem "reasonable",

$(sv) \otimes w = v \otimes (sw) = s\,(v\otimes w)$ // s = scalar ($\in$ K)
$v \otimes (w_1+ w_2) = v\otimes w_1 + v\otimes w_2$ // left distributive property
$(v_1 + v_2) \otimes w = v_1\otimes w + v_2\otimes w$. // right distributive property *(1.1.5)* (4.1.4)

In the last two equations, the + on the left represents addition in either W or V, whereas the + on the right side represents addition in V⊗W. These lines say that multiplication ⊗ "distributes" over addition +. The scalar rule can be combined with the distributive rules to obtain this equivalent rules restatement:

$v \otimes (s_1w_1+ s_2w_2) = s_1(v\otimes w_1) + s_2(v\otimes w_2)$ // s_1,s_2 = scalar ($\in$ K)
$(s_1v_1 + s_2v_2) \otimes w = s_1(v_1\otimes w)+ s_2(v_2\otimes w)$ .// s_1,s_2 = scalar ($\in$ K) *(1.1.7)* (4.1.5)

The above rules in effect say that ⊗ defines a "bilinear" operation -- it is linear separately in each of its operands.

Notice that the following two rules are incorrect:

$v \otimes w = w \otimes v$ // wrong! (unless V = W and v = w)
$(sv) \otimes (sw) = s(v\otimes w)$ // wrong! (unless s = 1)

As noted in Appendix B the second rule applies to a direct sum ⊕.

Using the correct "rules" above, one may write

$$v \otimes w = (\Sigma_i v^i u_i)\otimes(\Sigma_j w^j u_j') = \Sigma_{ij} v^i w^j \,(u_i\otimes u'_j) \qquad (4.1.6)$$

showing how this pure tensor product vector can be expressed in terms of the basis functions.

General tensors in V⊗W and V^2. A general "vector" (rank-2 cross tensor) in W⊗V can be written as a linear combination of the basis vectors, as was shown in (2.10.4),

$$T \equiv \Sigma_{ij}\ T^{ij}\ u_i \otimes u'_j \qquad T \in V\otimes W\ . \qquad \Sigma_{ij} \equiv \Sigma_{i=1}{}^{n}\Sigma_{j=1}{}^{n'} \qquad (4.1.7)$$

If W = V, we refer to the space V⊗W = V⊗V as V^2, and then

$$T \equiv \Sigma_{ij}\ T^{ij}\ u_i \otimes u_j \qquad T \in V\otimes V = V^2. \qquad \Sigma_{ij} \equiv \Sigma_{i=1}{}^{n}\Sigma_{j=1}{}^{n} \qquad (4.1.8)$$

Although we have said T is a "vector" in the abstract sense that a vector space (even a tensor product vector space) has "vectors" as elements, the usual terminology is to say that T is a "rank-2 tensor" in the space V^2.

Meanings of tensor. The word "tensor" has a weak and a strong meaning. In the weak meaning, a rank-2 tensor is something that has components with two indices like T^{ij}. In the strong meaning, a rank-2 tensor is a set of components T^{ij} which transform in a certain manner with respect to some underlying transformation,

$$T'^{ab} = R^a{}_{a'}\ R^b{}_{b'}\ T^{a'b'} \qquad \text{Picture A} \qquad T \in V\otimes V \qquad \textit{(2.1.7)}$$

Covariant expansion forms. The rank-2 tensor T can be expanded in various ways as shown in Section 2.10, and each such expansion has its own characteristic coefficients. Here are all four versions of (4.1.7) obtained using the tilt reversal rule (2.9.1) :

$$T \equiv \Sigma_{ij}\ T^{ij}\ u_i \otimes u'_j \qquad T \in V\otimes W$$
$$T \equiv \Sigma_{ij}\ T_i{}^j\ u^i \otimes u'_j$$
$$T \equiv \Sigma_{ij}\ T^i{}_j\ u_i \otimes u'^j$$
$$T \equiv \Sigma_{ij}\ T_{ij}\ u^i \otimes u'^j\ . \qquad (4.1.9)$$

Notation Comments:

• In (4.1.9) one could replace T^{ij} by $[T^{(u,u')}]^{ij}$ to be more precise about the meaning of the coefficients, namely, that they are those which arise when one expands on the basis $u_i \otimes u'_j$.

• Then in (4.1.8) one could write T^{ij} as $[T^{(u,u)}]^{ij}$ or $[T^{(u)}]^{ij}$, but in this case T^{ij} is the "default" notation for expanding on the axis-aligned $\mathbf{u}_i$ basis vectors as shown back in (2.10.3b).

Dot Products. One can define a covariant dot product between two elements of V^2 in this manner

$$A \bullet B \equiv \Sigma_{ij} A^{ij} B_{ij} \;=\; \Sigma_{ij} A_{ij} B^{ij} \;=\; \Sigma_{ij} B^{ij} A_{ij} \;=\; B \bullet A \;. \qquad (4.1.10)$$

If A or B is a pure rank-2 tensor, one can write as well

$$(a\otimes b)\bullet B = \Sigma_{ij} a^i b^j B_{ij}$$

$$A\bullet(c\otimes d) = \Sigma_{ij} A^{ij} c_i d_j$$

$$(a\otimes b)\bullet(c\otimes d) = \Sigma_{ij} a^i b^j c_i d_j \;=\; (a\bullet c)(b\bullet d)\,. \qquad (4.1.11)$$

The last line appears as (2.9.13).

Dirac Notation for Section 4.1 . It seemed best not to clutter that above text with these alternate forms. The Dirac version of an equation gets a D subscript on its equation number.

$\mathbf{u}_i = |\mathbf{u}_i\rangle \qquad \mathbf{u}'_i = |\mathbf{u}'_i\rangle$ bases for V and W

$\mathbf{u}_i\otimes\mathbf{u}'_j = |\mathbf{u}_i\rangle \otimes |\mathbf{u}'\,\rangle$ basis for the tensor product space V⊗W

$\mathbf{v}\otimes\mathbf{w} = |\mathbf{v}\rangle \otimes |\mathbf{w}\rangle$ a pure "vector" in the tensor product space V⊗W $\qquad (4.1.1)_D$

$T = |T\rangle = \Sigma_{ij}\, T^{ij}\, |\mathbf{u}_i\rangle \otimes |\mathbf{u}'_j\rangle$ rank-2 tensor in V⊗W

$(\mathbf{a} \otimes \mathbf{b})_{ij} = \langle\mathbf{u}_i| \otimes \langle\mathbf{u}_j| \;\; |\mathbf{a}\rangle \otimes |\mathbf{b}\rangle = \langle\mathbf{u}_i|\mathbf{a}\rangle \langle\mathbf{u}_j|\mathbf{b}\rangle = a_i b_j$ outer product

$\mathbf{a} \bullet \mathbf{b} = \langle\mathbf{a} \,|\, \mathbf{b}\rangle$ dot product

$|\mathbf{v}\rangle \otimes |\mathbf{w}\rangle = \Sigma_{ij} v^i w^j\, |\mathbf{u}_i\rangle \otimes |\mathbf{u}'_j\rangle$ expansion of pure vector on basis vectors $\qquad (4.1.6)_D$

$T \equiv |T\rangle = \Sigma_{ij}\, T^{ij}\, |\mathbf{u}_i\rangle \otimes |\mathbf{u}'_j\rangle$ expansion of a rank-2 tensor $\qquad (4.1.7)_D$

$$\langle\mathbf{u}^a| \otimes \langle\mathbf{u}'^b| \;\; |T\rangle = \langle\mathbf{u}^a| \otimes \langle\mathbf{u}'^b|\, \Sigma_{ij}\, T^{ij}\, |\mathbf{u}_i\rangle \otimes |\mathbf{u}'_j\rangle =$$

$$\Sigma_{ij}\, T^{ij} \langle\mathbf{u}^a|\mathbf{u}_i\rangle\langle\mathbf{u}'^b|\mathbf{u}'_j\rangle = \Sigma_{ij}\, T^{ij}\, \delta^a{}_i \delta^b{}_j \;=\; T^{ab}$$

$\langle\mathbf{a}| \otimes \langle\mathbf{b}| \;\; |\mathbf{c}\rangle \otimes |\mathbf{d}\rangle = \langle\mathbf{a}|\mathbf{c}\rangle\langle\mathbf{b}|\mathbf{d}\rangle = \Sigma_i a^i c_i\, \Sigma_j b^j d_j = \Sigma_{ij} a^i b^j c_i d_j$ dot product in $V^2 \qquad (4.1.11)_D$

4.2 The tensor product of 2 dual vectors in V^{*2}

The dual space V* of V was discussed in Section 2.11. Space W* is dual to W. We continue our convention of using Greek or script letters for dual space objects. The current section is basically a generalization of Section 2.11 to the case where V and W are different vector spaces. We show corresponding Section 2.11 equations in italics.

<u>Note</u>: The λ'^i used below are unrelated to the λ'^i of Chapter 2.

<u>Basics</u>. Consider the two dual vector spaces V* and W* (defined over field K) of dimension n and n'. Let

$\{\lambda^i\}$ = basis of V* $\quad$ dim(V*) = n $\quad$ $\alpha = \Sigma_{i=1}^n \alpha_i \lambda^i$ = general linear functional in V*
$\{\lambda'^i\}$ = basis of W* $\quad$ dim(W*) = n' $\quad$ $\beta = \Sigma_{j=1}^{n'} \beta_j \lambda'^j$ = general linear functional in W*

$\lambda^i = (u^i)^T = \langle u^i|$ $\qquad$ $\lambda'^i = (u'^i)^T = \langle u'^i|$ $\qquad$ // matrix and Dirac notation

$\{\lambda^i\otimes\lambda'^j\}$ = basis for the tensor product space V*⊗W* $\qquad$ dim(V*⊗W*) = n*n'

$\alpha\otimes\beta$ = a pure "vector" in the tensor product space V*⊗W* $\qquad$ $\alpha\otimes\beta \neq \beta\otimes\alpha$ if $\alpha \neq \beta$

$$\otimes : V^*xW^* \rightarrow V^*\otimes W^* \qquad \otimes : (\alpha,\beta) \mapsto \alpha\otimes\beta \qquad (4.2.1)$$

Note that $\alpha\otimes\beta \neq \beta\otimes\alpha$. For $V^*\neq W^*$, $\beta\otimes\alpha$ does not even make sense since that requires $\beta \in V^*$ and $\alpha \in W^*$. For V* = W* the objects $\alpha\otimes\beta$ and $\beta\otimes\alpha$ are still different unless $\alpha = \beta$.

<u>Vector expansions in V* and W*</u>. Linear functionals in V* and W* can be written as linear combinations of the basis functionals,

$$\alpha = \Sigma_i\alpha_i\lambda^i \qquad \alpha(v) = \Sigma_i\alpha_i\lambda^i(v) \qquad \alpha: V \rightarrow K \qquad \textit{(2.11.c.8)}$$
$$\beta = \Sigma_j\beta_j\lambda'^j \qquad \beta(v) = \Sigma_j\beta_j\lambda^{j}{}'(v) \qquad \beta: W \rightarrow K\ . \qquad (4.2.2)$$

The middle column shows the corresponding *functions* $\alpha(v)$ and $\beta(v)$, and we now have

$$\alpha(u_i) = \alpha_i \qquad \textit{(2.11.d.15)}$$
$$\beta(u'_j) = \beta_j\ . \qquad (4.2.3)$$

<u>Basis tensors in V*⊗W*</u>. The basis functionals for V*⊗W* are the $\lambda^i\otimes\lambda'^j$ where,

$$(\lambda^i\otimes\lambda'^j)(v,w) = \lambda^i(v)\lambda'^j(w) = \text{scalar * scalar = scalar} \in K \qquad \textit{(2.11.d.13)} \qquad (4.2.4)$$

where (recall these are called the "i^{th} coordinate functions")

$$\lambda^i(v) = v^i$$
$$\lambda'^i(w) = w^i \qquad \textit{(2.11.c.5)} \qquad (4.2.5)$$

so that

$$(\lambda^i \otimes \lambda'^j)(v,w) = v^i w^j \ . \qquad \textit{(2.11.d.13)} \qquad (4.2.6)$$

This function is manifestly bilinear in its two vector arguments.

The ⊗ Rules for V*⊗W* . The "rules" (1.1.5) for the ⊗ operator in the space V*⊗W* are the same as those for ⊗ in the space V⊗W, since V*⊗W* is, after all, a tensor product of two spaces,

$$(s\alpha) \otimes \beta = \alpha \otimes (s\beta) = s\,(\alpha \otimes \beta) \qquad s \in K,\ \alpha \in V^*\ \ \beta \in W^*$$
$$\alpha \otimes (\beta_1 + \beta_2) = \alpha \otimes \beta_1 + \alpha \otimes \beta_2 \qquad \text{// distributive property}$$
$$(\alpha_1 + \alpha_2) \otimes \beta = \alpha_1 \otimes \beta + \alpha_2 \otimes \beta \ . \qquad \text{// same idea as above} \qquad \textit{(4.1.4)} \qquad (4.2.7)$$

Rank-2 cross-tensor expansion in V*⊗W*. A general functional of the dual tensor product space V*⊗W* can be written

$$\mathcal{T} \equiv \Sigma_{ij}\ T_{ij}\ \lambda^i \otimes \lambda'^j \qquad \mathcal{T} \in V^* \otimes W^* \qquad \Sigma_{ij} \equiv \Sigma_{i=1}{}^{n} \Sigma_{j=1}{}^{n'} \qquad \textit{(2.11.d.11)} \qquad (4.2.8)$$

The T_{ij} here are exactly the same T_{ij} which appear in the V⊗W expansion (4.1.7), $T = \Sigma_{ij} T^{ij}\ u_i \otimes u'_j$. Evaluating at a point (v,w) in VxW one gets,

$$\mathcal{T}(v,w) = \Sigma_{ij}\ T_{ij}\ (\lambda^i \otimes \lambda'^j)(v,w) = \Sigma_{ij}\ T_{ij}\ \lambda_i(v)\, \lambda_j'(w) = \Sigma_{ij}\ T_{ij}\ v^i w^j \qquad \textit{(2.11.d.8)} \qquad (4.2.9)$$

so one may regard $\mathcal{T}$: VxW → K, and $\mathcal{T}(v,w)$ as manifestly bilinear in its arguments.

Setting $v = u_i$ and $w = u'_j$, one finds that

$$\mathcal{T}(u_i,u'_j) = T_{ij} \quad \in K \qquad \textit{(2.11.d.10)} \qquad (4.2.10)$$

This may be compared with $\alpha(u_i) = \alpha_i$ in (4.2.3).

An arbitrary rank-2 tensor functional $\mathcal{T}$ can be represented either by its expansion $\mathcal{T} = \Sigma_{ij} T_{ij} \lambda^i \otimes \lambda'^j$ or by the corresponding tensor function $\mathcal{T}(v,w)$.

As a special case, consider $\alpha \in V^*$ and $\beta \in W^*$ as shown above. Then,

$$\alpha \otimes \beta = (\Sigma_a \alpha_a \lambda^a) \otimes (\Sigma_b \beta_b \lambda'^b) = \Sigma_{ab}\ \alpha_a \beta_b\ \lambda^a \otimes \lambda'^b \quad \in V^* \otimes W^* \qquad (4.2.11)$$

$$(\alpha \otimes \beta)(v,w) = \Sigma_{ab} \alpha_a \beta_b (\lambda^a \otimes \lambda'^b)(v,w) = \Sigma_{ab} \alpha_a \beta_b \lambda^a(v) \lambda^{b\prime}(w) = \Sigma_{ab} \alpha_a \beta_b v^a w^b \qquad (4.2.12)$$

which then is just a particular example of (4.2.9). Continuing the above,

$$(\alpha \otimes \beta)(v,w) = \Sigma_{ab} \alpha_a \beta_b v^a w^b = [\Sigma_a \alpha_a v^a][\Sigma_b \alpha_b w^b]$$
$$= \alpha(v)\, \beta(w) \ . \qquad \textit{(2.11.d.15)} \qquad (4.2.13)$$

In particular,

$$(\alpha \otimes \beta)(u_i,u'_j) = \alpha(u_i)\, \beta(u'_j) = \alpha_i\beta_j = (\alpha \otimes \beta)_{ij} \,. \qquad \textit{(2.11.d.15)} \qquad (4.2.14)$$

If it happens that W = V, then W* = V* and we write $V^*\otimes W^* = V^*\otimes V^* = V^{*2}$. The equations above then revert to those given in Section 2.11 (referenced in italics above).

<u>The vector spaces V^{2*} and V^{2*}_f</u>

We can regard both $\mathcal{T} = <T|$ and $\mathcal{T}(v,w) = <T|v,w>$ as representations of the same rank-2 tensor functional $<T|$ in $V^*\otimes W^*$. The object $\mathcal{T}$ is a bilinear rank-2 tensor functional, whereas $\mathcal{T}(v,w)$ is a bilinear rank-2 tensor function (a Spivak 2-tensor). There is a 1-to-1 correspondence between $\mathcal{T}$ and $\mathcal{T}(v,w)$. We shall say $\mathcal{T} \in V^*\otimes W^*$ while $\mathcal{T}(v,w) \in (V^*\otimes W^*)_f$ (f = function), and the two spaces are isomorphic. If W = V, then $\mathcal{T} \in V^{2*}$ and $\mathcal{T}(v,w) \in V^{2*}{}_f$ and V^{2*} and $V^{2*}{}_f$ are isomorphic.

Fact: The vector space V^{*2} is equivalent to the vector space $V^{*2}{}_f$ of bilinear *functions* on V^2. (4.2.15)

<u>Dirac Notation for Section 4.2</u>

$\lambda^i = (\mathbf{u}^i)^T = <\mathbf{u}^i|$ $\qquad \lambda'^i = (\mathbf{u}'^i)^T = <\mathbf{u}'^i|$ $\qquad$ bases for V* and W*

$\lambda^i\otimes\lambda'^j = <\mathbf{u}^i| \otimes <\mathbf{u}'^j|$ $\qquad$ basis for the dual tensor product space V*⊗W*

$\alpha\otimes\beta = <\boldsymbol{\alpha}| \otimes <\boldsymbol{\beta}|$ $\qquad$ pure element of V* ⊗W* $\qquad (4.2.1)_D$

$\alpha = <\boldsymbol{\alpha}| = \Sigma_i\alpha_i<\mathbf{u}^i|$ $\qquad \alpha(v) = <\boldsymbol{\alpha}|\mathbf{v}> = \Sigma_i\alpha_i<\mathbf{u}^i|\mathbf{v}>$ $\qquad$ vector functional expansions
$\beta = <\boldsymbol{\beta}| = \Sigma_i\beta_i<\mathbf{u}'^i|$ $\qquad \beta(v) = <\boldsymbol{\beta}|\mathbf{v}> = \Sigma_i\beta_i<\mathbf{u}'^i|\mathbf{v}>$ $\qquad (4.2.2)_D$

$\alpha(\mathbf{u}_i) = <\boldsymbol{\alpha}|\mathbf{u}_i> = \alpha_i$ $\qquad$ vector function $\alpha(\mathbf{v})$ evaluated at $\mathbf{v} = \mathbf{u}_i$
$\beta(\mathbf{u}_i) = <\boldsymbol{\alpha}|\mathbf{u}_i> = \beta_i$ $\qquad$ vector function $\beta(\mathbf{v})$ evaluated at $\mathbf{v} = \mathbf{u}_i$ $\qquad (4.2.3)_D$

$$(\lambda^i\otimes\lambda'^j)(\mathbf{v},\mathbf{w}) = <\mathbf{u}^i| \otimes <\mathbf{u}'^j|\ |\mathbf{v}>\otimes |\mathbf{w}> = <\mathbf{u}^i|\,\mathbf{v}><\mathbf{u}'^j|\mathbf{w}> = v^iw^j \qquad (4.2.4)_D \qquad (4.2.5)_D$$

$$\mathcal{T} = <T| = \Sigma_{ij}T_{ij}<\mathbf{u}^i| \otimes <\mathbf{u}'^j| = \Sigma_{ij}\ T_{ij}\ \lambda^i \otimes \lambda'^i \quad \text{rank-2 tensor functional in } V^*\otimes W^* \qquad (4.2.8)_D$$

$$\mathcal{T}(\mathbf{v},\mathbf{w}) = <T|\ \ |\mathbf{v}> \otimes |\mathbf{w}> = <T|\mathbf{v},\mathbf{w}> \qquad \text{rank-2 tensor function for } V^*\otimes W^*$$

$$= \Sigma_{ij}T_{ij}<\mathbf{u}^i| \otimes <\mathbf{u}'^j|\ |\mathbf{v}> \otimes |\mathbf{w}> = \Sigma_{ij}T_{ij}<\mathbf{u}^i|\mathbf{v}> <\mathbf{u}'^j|\mathbf{w}> = \Sigma_{ij}T_{ij}v^iw^j \qquad (4.2.9)_D$$

$$(\alpha \otimes \beta)(\mathbf{v},\mathbf{w}) = <\boldsymbol{\alpha}| \otimes <\boldsymbol{\beta}|\ |\mathbf{v}> \otimes |\mathbf{w}> = <\boldsymbol{\alpha}|\,\mathbf{v}><\boldsymbol{\beta}|\,\mathbf{w}> \qquad (4.2.13)_D$$

4.3 The wedge product of 2 vectors in L^2

(a) Definition of the wedge product of 2 vectors and the space L^2

Momentarily jumping ahead, consider this equation,

$$v \wedge w = (v\otimes w - w\otimes v)/2 \ . \qquad v \in V \text{ and } w \in W$$

If V and W are different vector spaces, this makes no sense since the second term w⊗v implies that w lies in the left space V and v lies in the right space W. So in our discussion of wedge products, we require that W = V. This being the case, instead of using letters v and w as representative vectors, we shall use a and b. Then u_i are the basis vectors for both component spaces in the tensor product space V⊗V.

So, we start off by defining the following "wedge product" ("exterior product") of two vectors a,b ∈ V,

$$a \wedge b \equiv (a\otimes b - b\otimes a)/2 \ . \qquad \dim(V) = n \qquad (4.3.1)$$

Notice therefore that a ^ b is an element of $V\otimes V = V^2$, since it is a linear combination of elements of V⊗V. It is "antisymmetrized" under a ↔ b. Since not all elements of V⊗V can be written this way, the set of elements a ^ b exist in a subset of V⊗V which we shall call L^2, so $L^2 \subset V^2$. Some authors write L^2 as V^V.

The above definition trivially implies that

$$a \wedge b = - b \wedge a \qquad a, b \in V \qquad (4.3.2)$$

and

$$a \wedge a = 0 \qquad a \in V \ . \qquad (4.3.3)$$

In (1.1.5) we stated certain scalar and distributive properties of the ⊗ operator. These properties are passed through to the wedge ^ operator by the above definition. For example,

$$(sa) \wedge b = [\, (sa)\otimes b - b\otimes(sa)]/2 = s\,[\, a\otimes b - b\otimes a\,]/2 = s\,(a \wedge b) \qquad s = \text{scalar}$$

$$(a+c) \wedge b = [(a+c)\otimes b - b\otimes(a+c)]/2 = [a\otimes b + c\otimes b - b\otimes a - b\otimes c]/2$$

$$= [\, a\otimes b - b\otimes a\,]/2 + [\, c\otimes b - b\otimes c\,]/2 = (a \wedge b) + (c \wedge b) \qquad \text{distributive}$$

and similarly for a ^ (sb) and a ^ (b + c). To summarize, we have a set of ^ rules as follows:

$$(sa) \wedge b = s\,(a \wedge b) \qquad (a+c) \wedge b = (a \wedge b) + (c \wedge b) \qquad s \in K$$
$$a \wedge (sb) = s\,(a \wedge b) \qquad a \wedge (b + c) = (a \wedge b) + (a \wedge c) \qquad a,b,c \in V \ . \qquad (4.3.4)$$

The operator ^ is then seen to be "bilinear" over elements of V: it is separately linear in each operand.

To more precisely define the space L^2, we claim that the most general element $T_\wedge$ of the space L^2 can be written this way,

$$T_\wedge = \Sigma_{ij} T^{ij} u_i \wedge u_j . \qquad \Sigma_{ij} \equiv \Sigma_{i=1}^n \Sigma_{j=1}^n \tag{4.3.5}$$

where T^{ij} are the expansion coefficients. For example, if $T^{ij} = a^i b^j$ this would be,

$$T_\wedge = \Sigma_{ij} a^i b^j u_i \wedge u_j = (\Sigma_i a^i u_i) \wedge (\Sigma_j b^j u_j) = a \wedge b \tag{4.3.6}$$

and then $a \wedge b$ is included in L^2 for any vectors a and b in V.

We can take the ab component of (4.3.5) as follows

$$\begin{aligned} T_\wedge^{ab} &= \Sigma_{ij} T^{ij} (u_i \wedge u_j)^{ab} \\ &= \Sigma_{ij} T^{ij}(u_i\otimes u_j - u_j\otimes u_i)^{ab} / 2 = \Sigma_{ij} T^{ij}[(u_i\otimes u_j)^{ab} - (u_j\otimes u_i)^{ab}] / 2 \\ &= \Sigma_{ij} T^{ij} [u_i^a u_j^b - u_j^a u_i^b] / 2 = \Sigma_{ij} T^{ij} [\delta_i^a \delta_j^b - \delta_j^a \delta_i^b] / 2 \\ &= (1/2)[T^{ab} - T^{ba}] \end{aligned}$$

$$\Rightarrow \quad T_\wedge^{ab} = - T_\wedge^{ba} . \tag{4.3.7}$$

This shows that the expansion (4.3.5) can only represent an antisymmetric rank-2 tensor $T_\wedge$.

One could rearrange the n^2 basis vectors of $V\otimes V$ into these two groups,

$$\begin{aligned} &(u_i \wedge u_j) = [u_i\otimes u_j - u_j\otimes u_i]/2 && n(n-1)/2 \text{ independent elements in this set} \\ &(u_i * u_j) \equiv [u_i\otimes u_j + u_j\otimes u_i]/2 && n(n)/2 \text{ independent elements in this set} \end{aligned} \tag{4.3.8}$$

for a total of $n(n-1)/2+ n(n)/2 = n^2$ basis vectors. One would say then that L^2 is spanned by just the first set of basis vectors.

It was noted above that L^2 is a subset of V^2. A stronger statement is that L^2 is a subspace of V^2. First of all, L^2 is obviously closed under addition of vectors since

$$\Sigma_{ij} T_{ij} u_i \wedge u_j + \Sigma_{ij} T'_{ij} u_i \wedge u_j = \Sigma_{ij} (T_{ij}+T'_{ij}) u_i \wedge u_j . \tag{4.3.9}$$

And if $(a \wedge b)$ is an element of L^2 then so is $s(a \wedge b) = (s\alpha) \wedge b \in L^2$. Finally, since $a \wedge a = 0$, L^2 includes the 0 element. So L^2 then is a vector space which is a subspace of V^2.

(b) How big is the space L^2 compared to the space V^2?

Consider this most general element of L^2:

$$T_\wedge = \Sigma_{ij} T^{ij} (u_i \wedge u_j) = \Sigma_{i\neq j} T^{ij} (u_i \wedge u_j) \qquad // (u_i \wedge u_i) = 0$$

$$= \Sigma_{i<j} T^{ij} (u_i \wedge u_j) + \Sigma_{i>j} T^{ij} (u_i \wedge u_j)$$

$$= \Sigma_{i<j} T^{ij} (u_i \wedge u_j) + \Sigma_{j>i} T^{ji} (u_j \wedge u_i) \qquad // i\leftrightarrow j \text{ in second sum}$$

$$= \Sigma_{i<j} T^{ij} (u_i \wedge u_j) - \Sigma_{i<j} T^{ji} (u_i \wedge u_j) \qquad // (u_j \wedge u_i) = -(u_i \wedge u_j)$$

$$= \Sigma_{i<j} (T^{ij} - T^{ji}) (u_i \wedge u_j)$$

$$= \Sigma_{i<j} A^{ij} (u_i \wedge u_j) \qquad A^{ij} \equiv (T^{ij} - T^{ji}) \qquad A^{ij} = - A^{ji} . \qquad (4.3.10)$$

Thus, the number of elements in L^2 is equal to the number of antisymmetric n x n matrices A one can construct which contain elements of field K. An n x n antisymmetric matrix has only n(n-1)/2 places to insert independent values since the diagonal is all zeros and one triangle is the negative of the other. If the scalar space K contains N elements ($N = \infty$ for the reals), one could then construct exactly $N^{n(n-1)/2}$ antisymmetric matrices A.

Meanwhile, the most general element of V^2 can be written

$$T = \Sigma_{ij} T^{ij} (u_i \otimes u_j) . \qquad \textit{(4.1.9)}$$

Now each matrix T^{ij} defines an element of V^2. Using the same counting method as above, the total number of elements of V^2 is N^{n^2}. We conclude that

$$\frac{\#\text{ elements in } L^2}{\#\text{ elements in } V^2} = \frac{N^{n(n-1)/2}}{N^{n^2}} = (1/2)\frac{n^2 - n}{n^2} = (1/2)\left(1 - \frac{1}{n}\right) . \qquad (4.3.11)$$

The conclusion is that L^2 contains less than half the number of elements in V^2. This ratio is of course the same as the (4.3.8) count ratio of L^2 basis vectors to V^2 basis vectors: $[n(n-1)/2] / [n^2] = (n-1)/2n$.

Below we use this terminology,

$$T_\wedge = \Sigma_{ij} T^{ij} (u_i \wedge u_j) \qquad = \text{the "symmetric expansion" of } T_\wedge$$
$$T_\wedge = \Sigma_{i<j} A^{ij} (u_i \wedge u_j) \qquad = \text{the "ordered expansion" of } T_\wedge .$$

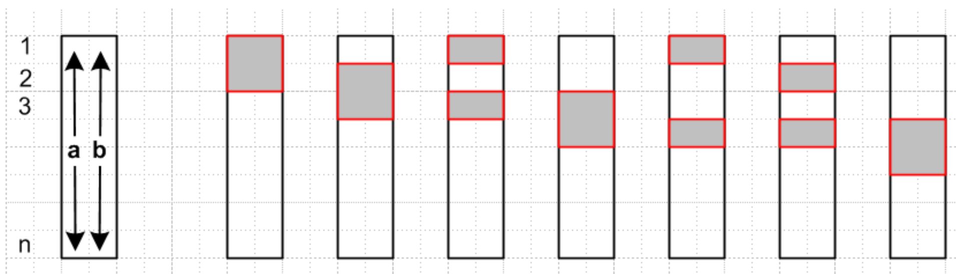
(c) Wedge products and determinants: the geometry connection

From (4.3.6) and (4.3.10) with $T^{ij} = a^i b^j$ we get,

$$a \wedge b = \Sigma_{ij} a^i b^j (u_i \wedge u_j) = \Sigma_{i<j} (a^i b^j - a^j b^i)\ (u_i \wedge u_j)$$

$$= \Sigma_{i<j} \det\begin{pmatrix} a^i & b^i \\ a^j & b^j \end{pmatrix} (u_i \wedge u_j) \qquad A^{ij} = (a^i b^j - a^j b^i) = \det\begin{pmatrix} a^i & b^i \\ a^j & b^j \end{pmatrix} . \qquad (4.3.12)$$

The determinants which appear here are 2x2 minors of a matrix having n rows and 2 columns. The two columns are the vectors **a** and **b,** each of which has n components. Below that matrix is shown on the left, and some of the 2x2 minors (row i < row j) are shown in gray on the right:



(4.3.13)

If $V = R^2$ (so n=2) there is only one term in the sum (4.3.12), the one with i=1 and j=2, so

$$\mathbf{a} \wedge \mathbf{b} = \det\begin{pmatrix} a^1 & b^1 \\ a^2 & b^2 \end{pmatrix} \mathbf{u_1} \wedge \mathbf{u_2} = \det(\mathbf{a},\mathbf{b})\, \mathbf{u_1} \wedge \mathbf{u_2} = [\, a^1 b^2 - a^2 b^1]\ \mathbf{u_1} \wedge \mathbf{u_2} . \qquad (4.3.14)$$

If one draws a parallelogram (2-piped) in the x-y plane with edges **a** and **b**, one knows that the area of that 2-piped is |**a** x **b**| which is then $|a^1 b^2 - a^2 b^1| = |\det(\mathbf{a},\mathbf{b})|$. There is then some connection between the wedge product of two vectors in R^2 and the geometry of R^2. Later in (7.5.6) we will show that for $V = R^3$ the *triple* wedge product of three vectors is given by,

$$\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} = \det(\mathbf{a},\mathbf{b},\mathbf{c})\ (\mathbf{u_1} \wedge \mathbf{u_2} \wedge \mathbf{u_3}) \qquad (4.3.15)$$

and here det(**a**,**b**,**c**) is the volume of the 3-piped spanned by the vectors **a**,**b**,**c,** so again there is a geometry connection. However, for R^3 the wedge product of *two* vectors is more complicated. Using the above expression, we find

$$\mathbf{a} \wedge \mathbf{b} = \det\begin{pmatrix} a^1 & b^1 \\ a^2 & b^2 \end{pmatrix} \mathbf{u_1} \wedge \mathbf{u_2} + \begin{pmatrix} a^1 & b^1 \\ a^3 & b^3 \end{pmatrix} \mathbf{u_1} \wedge \mathbf{u_3} + \begin{pmatrix} a^2 & b^2 \\ a^3 & b^3 \end{pmatrix} \mathbf{u_2} \wedge \mathbf{u_3}$$

$$= [a^1 b^2 - a^2 b^1]\, \mathbf{u_1} \wedge \mathbf{u_2} + [a^3 b^1 - a^1 b^3]\, \mathbf{u_3} \wedge \mathbf{u_1} + [a^2 b^3 - a^3 b^2]\, \mathbf{u_2} \wedge \mathbf{u_3} . \qquad (4.3.16)$$

The coefficients are those which appear in the normal "cross product" of two contravariant vectors,

$$\mathbf{a} \times \mathbf{b} = [a^1 b^2 - a^2 b^1]\ \mathbf{u_3} + [a^3 b^1 - a^1 b^3]\, \mathbf{u_2} + [a^2 b^3 - a^3 b^2]\, \mathbf{u_1} .$$

We do not wish, however, to identify for example $\mathbf{u_1}$^$\mathbf{u_2}$ with $\mathbf{u_3}$. After all, $\mathbf{u_3}$ is a basis vector in V, whereas $\mathbf{u_1}$^$\mathbf{u_2}$ is a vector in the tensor product space V⊗V. One can, on the other hand, define a *correspondence* of sorts where one says (each line in cyclic order, and ↔ means "corresponds to")

$$\mathbf{u_1}\wedge\mathbf{u_2} \leftrightarrow \mathbf{u_3}$$
$$\mathbf{u_2}\wedge\mathbf{u_3} \leftrightarrow \mathbf{u_1}$$
$$\mathbf{u_3}\wedge\mathbf{u_1} \leftrightarrow \mathbf{u_2} \qquad // = -\,\mathbf{u_1}\wedge\mathbf{u_3} \qquad (4.3.17)$$

in which case one can say

$$\mathbf{a} \wedge \mathbf{b} = [a^1b^2 - a^2b^1]\,\mathbf{u_1}\wedge\mathbf{u_2} + [a^3b^1 - a^1b^3]\,\mathbf{u_3}\wedge\mathbf{u_1} + [a^2b^3 - a^3b^2]\,\mathbf{u_2}\wedge\mathbf{u_3}$$
$$\leftrightarrow$$
$$\mathbf{a} \times \mathbf{b} = [a^1b^2 - a^2b^1]\,\mathbf{u_3} + [a^3b^1 - a^1b^3]\,\mathbf{u_2} + [a^2b^3 - a^3b^2]\,\mathbf{u_1} \qquad (4.3.18a)$$

so there is then a correspondence between the wedge product and the cross product in R^3. This correspondence was described by Scottish mathematician William Hodge (1903-1975) around 1941 and the relationship ↔ is formalized by the Hodge dual star operator *, see Appendix H. For example $*(\mathbf{u_1}\wedge\mathbf{u_2}) = \mathbf{u_3}$ and $*\mathbf{u_3} = \mathbf{u_1}\wedge\mathbf{u_2}$ in R^3. We can make the correspondence between ^ and x more explicit by writing $\mathbf{u_i} \wedge \mathbf{u_j} = \varepsilon_{ijk} A_k$ where for example $\mathbf{u_1} \wedge \mathbf{u_2} = \varepsilon_{123} A_3 = A_3 = *\mathbf{u_3}$ (defines A_3 which suggests area). Then the Hodge correspondence takes this form,

$$\mathbf{a} \wedge \mathbf{b} = a_ib_j\,\mathbf{u_i}\wedge\mathbf{u_j} = a_ib_j\varepsilon_{ijk} A_k = A_k [\varepsilon_{kij}a_ib_j] = \mathbf{A} \bullet (\mathbf{a} \times \mathbf{b}) \ . \qquad (4.3.18b)$$

For R^n with n > 3 there is no cross product of two vectors, but there *is* a wedge product. With $V = R^4$ for example, using the result (4.3.12) stated above,

$$\mathbf{a} \wedge \mathbf{b} = \det\begin{pmatrix} a^1 & b^1 \\ a^2 & b^2 \end{pmatrix} \mathbf{u_1}\wedge\mathbf{u_2} + \det\begin{pmatrix} a^1 & b^1 \\ a^3 & b^3 \end{pmatrix} \mathbf{u_1}\wedge\mathbf{u_3} + \det\begin{pmatrix} a^1 & b^1 \\ a^4 & b^4 \end{pmatrix} \mathbf{u_1}\wedge\mathbf{u_4}$$
$$+ \det\begin{pmatrix} a^2 & b^2 \\ a^3 & b^3 \end{pmatrix} \mathbf{u_2}\wedge\mathbf{u_3} + \det\begin{pmatrix} a^2 & b^2 \\ a^4 & b^4 \end{pmatrix} \mathbf{u_2}\wedge\mathbf{u_4} + \det\begin{pmatrix} a^3 & b^3 \\ a^4 & b^4 \end{pmatrix} \mathbf{u_3}\wedge\mathbf{u_4} \ . \qquad (4.3.19)$$

There are enthusiastic workers (e.g. Denker) who recommend deep-sixing the cross product altogether and replacing it with the wedge product for the study of topics like angular momentum. The wedge product plays a role in the so-called Clifford algebras, and a very famous such algebra is that involved in Dirac's relativistic theory of the electron, which theory predicts antiparticles. Elements of this Clifford algebra are the 4 x 4 "gamma matrices" γ^μ. This is the same Dirac whose bra-ket notation we are using, so somehow we have come full circle.

(d) Components

For the tensor product of two basis vectors we have these outer product forms,

$$(u_i\otimes u'_j)^{rs} = (u_i)^r(u'_j)^s = \delta_i{}^r\delta_j{}^s \qquad // \ V\otimes W$$
$$(u_i\otimes u_j)^{rs} = (u_i)^r(u_j)^s = \delta_i{}^r\delta_j{}^s \qquad // \ V\otimes V = V^2 \quad . \qquad (4.3.20)$$

For the wedge product $(u_i\wedge u_j)$ we have instead,

$$(u_i\wedge u_j)^{rs} = (1/2)[u_i\otimes u_j - u_j\otimes u_i]^{rs} = (1/2)[(u_i\otimes u_j)^{rs} - (u_j\otimes u_i)^{rs}] = (1/2)[(u_i)^r(u_j)^s - (u_j)^r(u_i)^s]$$
$$= (1/2)\,(\delta_i{}^r\,\delta_j{}^s - \delta_i{}^s\,\delta_j{}^r)$$

$$(u_i\wedge u_j)^{rs} = -(u_i\wedge u_j)^{sr} = -(u_j\wedge u_i)^{rs} \ . \qquad // \text{ two forms of antisymmetry} \qquad (4.3.21)$$

We now examine the pure wedge product $a \wedge b$ using both the symmetric expansion (4.3.5) and the ordered expansion (4.3.10).

Using the *symmetric* double sum expansion form (4.3.5) with $T^{ij} = a^ib^j$ one has from (4.3.6) and (4.3.21),

$$(a\wedge b)^{rs} = \Sigma_{ij}\, a^ib^j\,(u_i\wedge u_j)^{rs} = \Sigma_{ij}\, a_ib_j\,[\delta_i{}^r\delta_j{}^s - \delta_j{}^r\delta_i{}^s]/2$$
$$= (a^rb^s - a^sb^r)/2 \ . \qquad (4.3.22)$$

Using the *ordered* double sum expansion (4.3.10) with $T^{ij} = a^ib^j$, we find instead

$$(a\wedge b)^{rs} = \Sigma_{i<j}\,(a^ib^j - a^jb^i)\,(u_i\wedge u_j)^{rs} = \Sigma_{i<j}\det\begin{pmatrix} a^i & b^i \\ a^j & b^j\end{pmatrix}(u_i\wedge u_j)^{rs}$$
$$= (1/2)\,\Sigma_{1\le i<j\le n}\det\begin{pmatrix} a^i & b^i \\ a^j & b^j\end{pmatrix}[\delta_i{}^r\delta_j{}^s - \delta_j{}^r\delta_i{}^s] \ . \qquad (4.3.23)$$

For $r = s$, one clearly has $(a\wedge b)^{rs} = 0$. If $r < s$, then only the $\delta_i{}^r\delta_j{}^s$ term can contribute to the ordered sum, since this will make $i < j$, otherwise only the second term contributes. Then using θ(Boolean) = 1 if true else 0, we can evaluate as follows,

$$2(a\wedge b)^{rs} = \det\begin{pmatrix} a^r & b^r \\ a^s & b^s\end{pmatrix}\theta(r<s) - \det\begin{pmatrix} a^s & b^s \\ a^r & b^r\end{pmatrix}\theta(s<r)$$

$$= \det\begin{pmatrix} a^r & b^r \\ a^s & b^s\end{pmatrix}\theta(r<s) + \det\begin{pmatrix} a^r & b^r \\ a^s & b^s\end{pmatrix}\theta(s<r) \qquad // \text{ swap rows 2nd term}$$

$$= \det\begin{pmatrix} a^r & b^r \\ a^s & b^s\end{pmatrix}[\theta(r<s) + \theta(s<r)] = \det\begin{pmatrix} a^r & b^r \\ a^s & b^s\end{pmatrix} = a^rb^s - a^sb^r \ . \qquad (4.3.24)$$

Combining the results for r=s and r≠s we get

$$(a \wedge b)^{rs} = (a^r b^s - a^s b^r)/2 \qquad // \ T_\wedge^{rs} = (T^{rs} - T^{sr})/2 = A^{rs}/2 \qquad (4.3.25)$$

in agreement with (4.3.22) which used the symmetric sum.

<u>We now repeat this comparison for *general* elements of L^2.</u>

Using the *symmetric* double sum (4.3.5),

$$T_\wedge^{rs} = \Sigma_{ij} T^{ij} (u_i \wedge u_j)^{rs} = \Sigma_{ij} T^{ij} (\delta_i^r \delta_j^s - \delta_i^s \delta_j^r)/2$$

$$= (T^{rs} - T^{sr})/2 = A^{rs}/2 . \qquad (4.3.26)$$

Using the *ordered* double sum (4.3.10),

$$T_\wedge^{rs} = \Sigma_{i<j} A^{ij} (u_i \wedge u_j)^{rs} = \Sigma_{i<j} A^{ij} [\delta_i^r \delta_j^s - \delta_j^r \delta_i^s]/2 . \qquad (4.3.27)$$

For r = s one has [..] = 0 so $T_\wedge^{rs} = 0$. Otherwise,

$$2\,T_\wedge^{rs} = A^{rs}\,\theta(r<s) - A^{sr}\theta(s<r) = A^{rs}\,\theta(r<s) + A^{rs}\theta(s<r)$$

$$= A^{rs}[\,\theta(r<s) + \theta(s<r)] = A^{rs} \qquad (4.3.28)$$

with the conclusion that

$$T_\wedge^{rs} = A^{rs}/2 \qquad \text{for all } r,s \in (1,n) \qquad (4.3.29)$$

which agrees with (4.3.26) using the symmetric expansion.

(e) Dot Products

Since a^b is an element of V^2 as well as of L^2, we may use the V^2 dot product to write

$$(a\wedge b) \bullet (c\otimes d) = \{(a\otimes b - b\otimes a)/2\} \bullet (c\otimes d) = (1/2)\,[\,(a\otimes b)\bullet(c\otimes d) - (b\otimes a)\bullet(c\otimes d)]$$

$$= [(a\bullet c)(b\bullet d) - (b\bullet c)(a\bullet d)]/2 \qquad // \ (2.9.13) \qquad (4.3.30)$$

with this special case

$$(u_i\wedge u_j) \bullet (c\otimes d) = [(u_i\bullet c)(u_j\bullet d) - (u_j\bullet c)(u_i\bullet d)]/2 = [c_i d_j - c_j d_i]/2 . \qquad (4.3.31)$$

The dot product of two-vector wedge products is the same as (4.3.30),

$(a\wedge b) \bullet (c\wedge d) = \{(a\otimes b-b\otimes a)/2\} \bullet \{(c\otimes d-d\otimes c)/2\} =$

$= (1/4) [(a\otimes b)\bullet(c\otimes d) - (b\otimes a)\bullet(c\otimes d) - (a\otimes b)\bullet(d\otimes c) + (b\otimes a)\bullet(d\otimes c)]$

$= (1/4) [(a\bullet c)(b\bullet d) - (b\bullet c)(a\bullet d) - (a\bullet d)(b\bullet c) + (b\bullet d)(a\bullet c)]$

$= [(a\bullet c)(b\bullet d) - (b\bullet c)(a\bullet d)]/2$

$= (a\wedge b) \bullet (c\otimes d)$

$= (a\otimes b) \bullet (c\wedge d)$ (4.3.32)

so then the special case is the same as (4.3.31),

$(u_i\wedge u_j) \bullet (c\wedge d) = [c_i d_j - c_j d_i]/2$. (4.3.33)

<u>Dirac Notation for Section 4.3</u> (a selection)

Section 4.3 (a)

$|\mathbf{a}> \wedge |\mathbf{b}> = (|\mathbf{a}>\otimes |\mathbf{b}> - |\mathbf{b}>\otimes |\mathbf{a}>)/2$ $(4.3.1)_D$

$|\mathbf{a}> \wedge |\mathbf{b}> = - |\mathbf{b}> \wedge |\mathbf{a}>$ $(4.3.2)_D$

and $|\mathbf{a}> \wedge |\mathbf{a}> = 0$ $(4.3.3)_D$

$T_\wedge = |T_\wedge> = \Sigma_{ij} T^{ij} |\mathbf{u}_i> \wedge |\mathbf{u}_j>$ $(4.3.5)_D$

$T_\wedge{}^{ab} = < \mathbf{u}^a| \otimes <\mathbf{u}^b | \ |T_\wedge> = < \mathbf{u}^a| \otimes <\mathbf{u}^b | \ \Sigma_{ij} T^{ij} |\mathbf{u}_i> \wedge |\mathbf{u}_j>$

$= \Sigma_{ij} T^{ij} <\mathbf{u}^a| \otimes <\mathbf{u}^b | \ [|\mathbf{u}_i> \otimes |\mathbf{u}_j> - |\mathbf{u}_j> \otimes |\mathbf{u}_i>]/2$

$= \Sigma_{ij} T^{ij} [<\mathbf{u}^a |\mathbf{u}_i> <\mathbf{u}^b |\mathbf{u}_j> - <\mathbf{u}^a |\mathbf{u}_j> <\mathbf{u}^b |\mathbf{u}_i>]/2$

$= \Sigma_{ij} T^{ij} [\delta^a{}_i\delta^b{}_j - \delta^a{}_j\delta^b{}_i]/2 = (T^{ab} - T^{ba})/2 = - T_\wedge{}^{ba}$ $(4.3.7)_D$

Section 4.3 (b)

$T_\wedge = |T_\wedge> = \Sigma_{i<j} A^{ij} |\mathbf{u}_i> \wedge |\mathbf{u}_j>$ $(4.3.10)_D$

Section 4.3 (c)

$$|\mathbf{a}\rangle \wedge |\mathbf{b}\rangle = \Sigma_{ij}\, a^i b^j\ |\mathbf{u}_i\rangle \wedge |\mathbf{u}_j\rangle\ = \Sigma_{i<j} \det\begin{pmatrix} a^i & b^i \\ a^j & b^j \end{pmatrix}\ |\mathbf{u}_i\rangle \wedge |\mathbf{u}_j\rangle \qquad (4.3.12)_D$$

$$|\mathbf{a}\rangle \wedge |\mathbf{b}\rangle \wedge |\mathbf{c}\rangle\ = \det(\mathbf{a},\mathbf{b},\mathbf{c})\, |\mathbf{u}_1\rangle \wedge |\mathbf{u}_2\rangle \wedge |\mathbf{u}_3\rangle \qquad n = 3 \qquad (4.3.15)_D$$

Section 4.3 (d)

$$(\mathbf{a} \wedge \mathbf{b})^{rs}\ = \langle u^r| \otimes \langle u^s|\ \ |\mathbf{a}\rangle \wedge |\mathbf{b}\rangle = \ \langle u^r| \otimes \langle u^s|\ \ (\ |\mathbf{a}\rangle\otimes |\mathbf{b}\rangle\ -\ |\mathbf{b}\rangle\otimes |\mathbf{a}\rangle\ )/2$$

$$= [\ \langle u^r| \otimes \langle u^s|\ \ |\mathbf{a}\rangle\otimes |\mathbf{b}\rangle\ -\ \langle u^r| \otimes \langle u^s|\ \ |\mathbf{b}\rangle\otimes |\mathbf{a}\rangle\ ]/2$$

$$= [\ \langle u^r|\mathbf{a}\rangle\langle u^s|\mathbf{b}\rangle - \langle u^r|\mathbf{b}\rangle\langle u^s|\mathbf{a}\rangle\ ]/2\ = (a^r b^s - a^s b^r)/2 \qquad (4.3.25)_D$$

Section 4.3 (e)

$$(u_i \wedge u_j) \bullet (c\otimes d)\ =\ [(u_i\bullet c)(u_j\bullet d) - (u_j\bullet c)(u_i\bullet d)]/2\ = [c_i d_j - c_j d_i]/2\ .$$

$$\langle\mathbf{u}_i| \wedge \langle\mathbf{u}_j|\ \ |\mathbf{c}\rangle\otimes |\mathbf{d}\rangle\ = (1/2)[\langle\mathbf{u}_i| \otimes \langle\mathbf{u}_j|\ - \langle\mathbf{u}_j| \otimes \langle\mathbf{u}_i|\ ]\ \ |\mathbf{c}\rangle\otimes |\mathbf{d}\rangle$$

$$= (1/2)[\ \langle\mathbf{u}_i|\mathbf{c}\rangle\langle\mathbf{u}_j|\mathbf{d}\rangle -\ \langle\mathbf{u}_j|\mathbf{c}\rangle\langle\mathbf{u}_i|\mathbf{d}\rangle\ ] = (1/2)\ [\ c_i d_j - c_j d_i] \qquad (4.3.31)_D$$

4.4 The wedge product of 2 dual vectors in Λ^2

Section 4.3 considered the wedge product of two vectors in V^2. Here we consider the wedge product of two vectors in the dual space V^{*2}. We mimic the approach of Section 4.3, omitting some details, and we match equation numbers even though this leaves some "holes" in the sequence.

(a) Definition of the wedge product and the space Λ^2

We start off by defining the following wedge product of two vectors (linear functionals) α and β of V*,

$$\alpha \wedge \beta \equiv (\alpha \otimes \beta - \beta \otimes \alpha)/2 \ . \qquad (4.4.1)$$

Notice therefore that $\alpha \wedge \beta$ is an element of $V^*\otimes V^* = V^{*2}$, since it is a linear combination of elements of $V^*\otimes V^*$. It is "antisymmetrized" under $\alpha \leftrightarrow \beta$. Since not all elements of $V^*\otimes V^* = V^{*2}$ can be written this way, the set of elements $\alpha \wedge \beta$ exist in a subspace of V^{*2} which we shall call Λ^2, so $\Lambda^2 \subset V^{*2}$. Some authors write Λ^2 as $V^*\wedge V^*$. The proof that Λ^2 is a subspace and not just a subset of V^{*2} is the same as in the Section 4.3 (a).

The above definition trivially implies that

$$\alpha \wedge \beta = -\beta \wedge \alpha \qquad \alpha, \beta \in V^* \qquad (4.4.2)$$

and

$$\alpha \wedge \alpha = 0 \qquad \alpha \in V^* \ . \qquad (4.4.3)$$

The "rules" for the ^ operator in $\Lambda^2 \subset V^{*2}$ are found just as they were for $L^2 \subset V^2$, namely :

$$(s\alpha) \wedge \beta = s(\alpha \wedge \beta) \qquad (\alpha+\gamma) \wedge \beta = (\alpha \wedge \beta) + (\gamma \wedge \beta)$$
$$\alpha \wedge (s\beta) = s(\alpha \wedge \beta) \qquad \alpha \wedge (\beta + \gamma) = (\alpha \wedge \beta) + (\alpha \wedge \gamma) \qquad (4.4.4)$$

where α,β,γ are vectors in V* and s is a scalar in K.

To more precisely define the space Λ^2, we claim that the most general element of the space Λ^2 can be written this way (that is, Λ^2 is the space spanned by the $\lambda_i \wedge \lambda_j$ basis vectors)

$$\mathcal{T}_\wedge = \Sigma_{ij} T_{ij}\ \lambda^i \wedge \lambda^j \ . \qquad // \lambda^i \wedge \lambda^j = \langle u^i| \wedge \langle u^j| \text{ in Dirac notation} \qquad (4.4.5)$$

For example, if $T_{ij} = \alpha_i\beta_j$ this would be

$$\mathcal{T}_\wedge = \Sigma_{ij}\ \alpha_i\beta_j\ (\lambda^i \wedge \lambda^j) = (\Sigma_i\alpha_i\lambda^i) \wedge (\Sigma_j\beta_j\lambda^j) = \alpha \wedge \beta \qquad (4.4.6)$$

and then $\alpha \wedge \beta$ is included in Λ^2 for any vectors α and β in V*.

Λ^2 is spanned by the n(n-1)/2 independent basis vectors $(\lambda^i \wedge \lambda^j)$ for i < j. (4.4.8)
Λ^2 is a subspace of V^{*2}, just as L^2 is a subspace of V^2 as shown near (4.3.9). (4.4.9)

(b) How big is the space Λ^2 compared to the space V^{*2}?

Just as in Section 4.3 (b), we can show that

$$\mathcal{T}_\wedge = \Sigma_{ij}\, T_{ij}\, (\lambda^i \wedge \lambda^j)$$

$$= \Sigma_{i<j}\, A_{ij}\, (\lambda^i \wedge \lambda^j) \qquad A_{ij} \equiv (T_{ij} - T_{ji}) \qquad A_{ij} = - A_{ji} \qquad (4.4.10)$$

where A^{ij} is an antisymmetric n x n matrix. Using the same argument presented there, we find

$$\frac{\#\text{ elements in } \boldsymbol{\Lambda}^2}{\#\text{ elements in } V^{*2}} = \frac{Nn(n-1)/2}{Nn^2} = (1/2)\frac{n^2 - n}{n^2} = (1/2)\,(1 - \frac{1}{n}) \;. \qquad (4.4.11)$$

(c) Wedge products and determinants

From (4.4.6) and (4.4.10) with $T^{ij} = \alpha^i\beta^j$ we get,

$$\alpha \wedge \beta = \Sigma_{ij}\, \alpha_i\beta_j\, (\lambda^i \wedge \lambda^j) = \Sigma_{i<j}\, (\alpha_i\beta_j - \alpha_j\beta_i)\, (\lambda^i \wedge \lambda^j)$$

$$= \Sigma_{i<j} \det\begin{pmatrix} \alpha_i & \beta_i \\ \alpha_j & \beta_j \end{pmatrix} (\lambda^i \wedge \lambda^j) \qquad A_{ij} = (\alpha_i\beta_j - \alpha_j\beta_i) = \det\begin{pmatrix} \alpha_i & \beta_i \\ \alpha_j & \beta_j \end{pmatrix} \;. \qquad (4.4.12)$$

See Fig (4.3.13) for an interpretation of this sum.

If $V^* = R^2$ (so n=2) there is only one term in the sum (4.4.12), the one with i=1 and j=2, so

$$\alpha \wedge \beta = \det\begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \lambda^1 \wedge \lambda^2 = \det(\alpha,\beta)\, \boldsymbol{\lambda}^1 \wedge \boldsymbol{\lambda}^2 = [\, \alpha_1\beta_2 - \alpha_2\beta_1]\, \lambda^1 \wedge \lambda^2 \;. \qquad (4.4.14)$$

For $V = R^3$ the *triple* wedge product of three vectors is given by,

$$\alpha \wedge \beta \wedge \gamma = \det(\alpha,\beta,\gamma)\, (\lambda^1 \wedge \lambda^2 \wedge \lambda^3) \qquad (4.4.15)$$

It does not seem useful to discuss "geometry" in the space of functionals, but we could be wrong.

(d) Tensor Functions

In Section 4.3(d) we discussed components $(u_i \otimes u'_j)^{rs}$ and $(u_i \wedge u_j)^{rs}$. In the dual world, the corresponding objects are the tensor functions $(\lambda^i \otimes \lambda'^j)(v_r,w_s)$ and $(\lambda^i \wedge \lambda^j)(v_r,w_s)$.

For the tensor product of two dual basis vectors we have these tensor functions,

$(\lambda^i\otimes\lambda'^j)(v_r,w_s) = \lambda^i(v_r)\lambda'^j(w_s) = (v_r)^i(w_s)^j$ // V*⊗W*; r and s are vector labels
$(\lambda^i\otimes\lambda^j)(v_r,v_s) = \lambda^i(v_r)\lambda^j(v_s) = (v_r)^i(v_s)^j$ // V*⊗V*
and
$(\lambda^i\otimes\lambda'^j)(u_r,u'_s) = (u_r)^i(u'_s)^j = \delta_r{}^i\delta_s{}^j$ // V*⊗W*
$(\lambda^i\otimes\lambda^j)(u_r,u_s) = (u_r)^i(u_s)^j = \delta_r{}^i\delta_s{}^j$. // V*⊗V* (4.4.20)

For the wedge product $(\lambda^i\wedge\lambda^j)$ we have instead,

$$(\lambda^i\wedge\lambda^j)(v_r,v_s) = [(\lambda^i\otimes\lambda^j)(v_r,v_s) - (\lambda^j\otimes\lambda^i)(v_r,v_s)]/2 = [\lambda^i(v_r)\lambda^j(v_s) - \lambda^j(v_r)\lambda^i(v_s)]/2$$

$$= (1/2)\,[\,(v_r)^i(v_s)^j - (v_r)^j(v_s)^i]$$ // $\Lambda^2 = V^* \wedge V^*$

$$(\lambda^i\wedge\lambda^j)(u_r,u_s) = (1/2)\,[\,\delta_r{}^i\delta_s{}^j - \delta_r{}^j\delta_s{}^i]$$

$$(\lambda^i\wedge\lambda^j)(v_r,v_s) = -(\lambda^i\wedge\lambda^j)(v_s,v_r) = -(\lambda^j\wedge\lambda^i)(v_r,v_s)\ .$$ // two forms of antisymmetry (4.4.21)

We now examine the pure wedge product $\alpha \wedge \beta$ using both the symmetric expansion (4.4.5) and the ordered expansion (4.4.10).

Using the *symmetric* double sum expansion form (4.4.5) with $T_{ij} = \alpha_i\beta_j$ one has,

$$(\alpha \wedge \beta)(v_r,v_s) = \Sigma_{ij}\,\alpha_i\beta_j\,(\lambda^i \wedge \lambda^j)(v_r,v_s) = \Sigma_{ij}\,\alpha_i\beta_j\,[\,\lambda^i(v_r)\lambda^j(v_s) - \lambda^j(v_r)\lambda^i(v_s)]/2$$

$$= \Sigma_{ij}\,\alpha_i\beta_j[(v_r)^i(v_s)^j - (v_r)^j(v_s)^i]/2 = [\,\alpha(v_r)\beta(v_s) - \alpha(v_s)\beta(v_r)]/2\ .$$

$$(\alpha \wedge \beta)(u_r,u_s) = \Sigma_{ij}\,\alpha_i\beta_j\,[\delta^i{}_r\delta^j{}_s - \delta^j{}_r\delta^i{}_s]/2 = (\alpha_r\beta_s - \alpha_s\beta_r)/2\ . \qquad (4.4.22)$$

where in the last equation we set $v_r = u_r$ and $v_s = u_s$ and use (2.4.1) that $(u_r)^i = \delta_r{}^i$.

Using the *ordered* double sum expansion (4.4.10) with $T_{ij} = \alpha_i\beta_j$, we find instead

$$(\alpha \wedge \beta)(v_r,v_s) = \Sigma_{i<j}\,(\alpha_i\beta_j - \alpha_j\beta_i)\,(\lambda^i \wedge \lambda^j)(v_r,v_s) = \Sigma_{i<j}\,\det\begin{pmatrix}\alpha_i & \beta_i\\ \alpha_j & \beta_j\end{pmatrix}(\lambda^i \wedge \lambda^j)(v_r,v_s)$$

$$= \Sigma_{i<j}\det\begin{pmatrix}\alpha_i & \beta_i\\ \alpha_j & \beta_j\end{pmatrix}[\,\lambda^i(v_r)\lambda^j(v_s) - \lambda^j(v_r)\lambda^i(v_s)]/2$$

$$= \Sigma_{i<j}\det\begin{pmatrix}\alpha_i & \beta_i\\ \alpha_j & \beta_j\end{pmatrix}\ (1/2)\det\begin{pmatrix}\lambda^i(v_r) & \lambda^j(v_r)\\ \lambda^i(v_s) & \lambda^j(v_s)\end{pmatrix}.$$

$$(\alpha \wedge \beta)(u_r,u_s) = \Sigma_{i<j}\ \det\begin{pmatrix}\alpha_i & \beta_i\\ \alpha_j & \beta_j\end{pmatrix}\ (1/2)\,[\delta^i{}_r\delta^j{}_s - \delta^j{}_r\delta^i{}_s]\ . \qquad (4.4.23)$$

where again in the last equation $v_r = u_r$ and $v_s = u_s$.

Repeating the argument (4.3.24) this becomes

$$(\alpha \wedge \beta)(u_r,u_s) = (\alpha_r\beta_s - \alpha_s\beta_r)/2 \qquad \text{// later this will be called } [\mathrm{Alt}(\alpha\otimes\beta)]_{rs} \qquad (4.4.25)$$

in agreement with (4.4.22) which used the symmetric sum.

We now repeat this comparison for *general* elements of Λ^2.

Using the *symmetric* double sum (4.4.5),

$$\mathcal{T}_\wedge(v_r,v_s) = \Sigma_{ij}\, T_{ij}\, (\lambda^i \wedge \lambda^j)(v_r,v_s) = \Sigma_{ij}\, T_{ij}\, [\, \lambda^i(v_r)\lambda^j(v_s) - \lambda^j(v_r)\lambda^i(v_s)]/2$$

$$\mathcal{T}_\wedge(u_r,u_s) = \Sigma_{ij}\, T_{ij}\, [\delta^i{}_r\delta^j{}_s - \delta^j{}_r\delta^i{}_s] / 2 = (T_{rs} - T_{sr})/2 = A_{rs}/2\,. \qquad \text{// see (4.4.10)} \qquad (4.4.26)$$

Using the *ordered* double sum (4.4.10),

$$\mathcal{T}_\wedge(u_r,u_s) = \Sigma_{i<j}\, A_{ij}\, (\lambda^i \wedge \lambda^j)(u_r,u_s) = \Sigma_{i<j}\, A_{ij}\, [\delta^i{}_r\delta^j{}_s - \delta^j{}_r\delta^i{}_s]/2\,. \qquad (4.4.27)$$

For r = s one has [..] = 0 so $\mathcal{T}_\wedge(u_r,u_s) = 0$. Otherwise,

$$2\mathcal{T}_\wedge(u_r,u_s) = A_{rs}\, \theta(r<s) - A_{sr}\theta(s<r) = A_{rs}\, \theta(r<s) + A_{rs}\theta(s<r)$$

$$= A_{rs}[\, \theta(r<s) + \theta(s<r)] = A_{rs} \qquad (4.4.28)$$

with the conclusion that

$$\mathcal{T}_\wedge(u_r,u_s) = A_{rs}/2 \qquad \text{for all } r,s \in (1,n) \qquad (4.4.29)$$

in agreement with (4.4.26).

Section 4.3 (e) on dot products like (a^b) • (c⊗d) has no useful analog for functionals.

The vector spaces Λ^2 and $\Lambda^2{}_f$

Looking at (4.4.21), (4.4.22) and (4.4.26), one sees that $(\lambda^i \wedge \lambda^j)(v_r,v_s)$, $(\alpha \wedge \beta)(v_r,v_s)$ and $\mathcal{T}_\wedge(v_r,v_s)$ are all *antisymmetric* bilinear functions of the two arguments v_r, $v_s \in V$.

In Section 4.3 we declare that the rank-2 tensor $T_\wedge = |T_\wedge\rangle$ is antisymmetric (alternating) if $T_\wedge{}^{ab} = - T_\wedge{}^{ba}$. In similar fashion, we declare that the rank-2 tensor functional $\mathcal{T}_\wedge = \langle T_\wedge|$ is antisymmetric (alternating) if $\mathcal{T}_\wedge(v,v') = - \mathcal{T}_\wedge(v',v)$. That is to say, saying that the functional is alternating means that the corresponding tensor function is alternating.

We can regard both $\mathcal{T}_\wedge$ = <T^| and $\mathcal{T}_\wedge$(v,v') = <T^|v,v'> as representations of the same antisymmetric rank-2 tensor functional <T^| in V* ^ V* = Λ^2. The object $\mathcal{T}_\wedge$ is an antisymmetric bilinear rank-2 tensor *functional*, whereas $\mathcal{T}_\wedge$(v,v') is an antisymmetric bilinear rank-2 tensor *function* (a Spivak 2-tensor). There is a 1-to-1 correspondence between $\mathcal{T}_\wedge$ and $\mathcal{T}_\wedge$(v,v') . We shall say $\mathcal{T}_\wedge \in \Lambda^2$ while $\mathcal{T}_\wedge(v,v') \in \Lambda^2{}_f$ (f = function), and the two spaces are isomorphic. Therefore,

Fact: The vector space Λ^2 is equivalent to the vector space $\Lambda^2{}_f$ of *antisymmetric* bilinear functions on V^2. (4.4.34)

This may be compared to our earlier statement for the larger space $V^{*2} = V^*\otimes V^*$,

Fact: The vector space V^{*2} is equivalent to the vector space $V^{*2}{}_f$ of bilinear functions on V^2. *(4.2.15)*

Dirac Notation for Section 4.4 (a selection)

Section 4.4 (a)

$$<\boldsymbol{\alpha}| \wedge <\boldsymbol{\beta}| = (<\boldsymbol{\alpha}| \otimes <\boldsymbol{\beta}| - <\boldsymbol{\alpha}| \otimes <\boldsymbol{\beta}|)/2 \qquad (4.4.1)_D$$

$$<\boldsymbol{\alpha}| \wedge <\boldsymbol{\beta}| = - <\boldsymbol{\beta}| \wedge <\boldsymbol{\alpha}| \qquad (4.4.2)_D$$

$$<\boldsymbol{\alpha}| \wedge <\boldsymbol{\alpha}| = 0 \qquad (4.4.3)_D$$

$$\mathcal{T}_\wedge = <T_\wedge| = \Sigma_{ij} T_{ij} <\mathbf{u}^i| \wedge <\mathbf{u}^j| \qquad // \lambda^i = <\mathbf{u}^i| \qquad (4.4.5)_D$$

$$\mathcal{T}_\wedge = <T_\wedge| = \Sigma_{ij}\alpha_i\beta_j <\mathbf{u}^i| \wedge <\mathbf{u}^j| = (\Sigma_i\alpha_i<\mathbf{u}^i|)(\Sigma_j\beta_i<\mathbf{u}^j|) = <\boldsymbol{\alpha}| \wedge <\boldsymbol{\beta}| \qquad (4.4.6)_D$$

Section 4.4 (b)

$$\mathcal{T}_\wedge = <T_\wedge| = \Sigma_{i<j} A_{ij} <\mathbf{u}^i| \wedge <\mathbf{u}^j| \qquad (4.4.10)_D$$

Section 4.4 (c)

$$\alpha \wedge \beta = <\boldsymbol{\alpha}| \wedge <\boldsymbol{\beta}| = \Sigma_{i<j} \det\begin{pmatrix} \alpha_i & \beta_i \\ \alpha_j & \beta_j \end{pmatrix} <\mathbf{u}^i| \wedge <\mathbf{u}^j| \qquad (4.4.12)_D$$

Section 4.4 (d)

$$(\lambda^i\wedge \lambda^j)(\mathbf{v_r},\mathbf{v_s}) = <\mathbf{u}^i| \wedge <\mathbf{u}^j| \;\; | \mathbf{v_r}> \otimes | \mathbf{v_s}> = (1/2) (<\mathbf{u}^i| \otimes <\mathbf{u}^j| - <\mathbf{u}^j| \otimes <\mathbf{u}^i|) | \mathbf{v_r}> \otimes | \mathbf{v_s}>$$

$$= (1/2) (<\mathbf{u}^i|\mathbf{v_r}><\mathbf{u}^j| \mathbf{v_s}> - <\mathbf{u}^i|\mathbf{v_s}><\mathbf{u}^j| \mathbf{v_r}>) = (1/2) [(\mathbf{v_r})^i(\mathbf{v_s})^j - (\mathbf{v_r})^j(\mathbf{v_s})^i] \qquad (4.4.21)_D$$

5. The Tensor Product of k vectors : the vector spaces V^k and T(V)

Our task is now to generalize the tensor product from V^2 to V^k, where

$$V^k \equiv V \otimes V \otimes .... \otimes V \ . \qquad \text{// tensor product of k vector spaces, each one is V} \qquad (5.1)$$

We are setting up for a parallel treatment in Chapter 6 where ⊗ becomes ^, so certain rather obvious statements will be made here to allow for comparison later with the wedge product.

5.1 Pure elements, basis elements, and dimension of V^k

A generic pure ("decomposable") element of V^k is this tensor product of k vectors,

$$v_1 \otimes v_2 \otimes .....\otimes v_k \qquad \text{all } v_i \in V \qquad (5.1.1)$$

$$= |v_1\rangle \otimes |v_2\rangle ... \otimes |v_k\rangle = |v_1,v_2, ... v_k\rangle \ . \qquad \text{// Dirac notation}$$

Since ⊗ is associative by (2.8.21), one can install parentheses anywhere in (5.1.1) without altering the meaning of the object, for example, $v_1 \otimes (v_2 \otimes v_3) \otimes ....\otimes v_k = v_1 \otimes v_2 \otimes v_3 \otimes ....\otimes v_k$.

The basis elements of V^k are (these u_i are those of Section 2.4),

$$u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k} = |u_{i_1}\rangle \otimes |u_{i_2}\rangle ... \otimes |u_{i_k}\rangle = | u_{i_1}, u_{i_2} ...u_{i_k}\rangle \ . \qquad (5.1.2)$$

In (5.1.1) and (5.1.2) the subscripts are labels, not components. The components of these two tensor objects are given by the (2.8.17) outer product form,

$$(v_1 \otimes v_2 \otimes .....\otimes v_k)^{j_1j_2...j_k} = (v_1)^{j_1} (v_2)^{j_2} .... (v_k)^{j_k} \qquad (5.1.3)$$

$$(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k})^{j_1j_2...j_k} = (u_{i_1})^{j_1} (u_{i_2})^{j_2} .... (u_{i_k})^{j_k} = \delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k} \ . \qquad (5.1.4)$$

If n = dim(V), the total number of such basis elements is n^k, so

$$\dim(V^k) = n^k. \qquad (5.1.5)$$

In the full set of tensor-product basis elements shown in (5.1.2), two or more of the u_{i_r} might be the same. This will always be the case if $k > n$ where $n \equiv \dim(V)$. For example, for k = 3 and n = 2 one such element would be $u_1 \otimes u_1 \otimes u_2 \neq 0$.

In Dirac notation, we can write (5.1.3) and (5.1.4) as

$$\langle u^{j_1}, u^{j_2} ...,u^{j_k} | v_1,v_2, ...v_k\rangle = \langle u^{j_1}|v_1\rangle\langle u^{j_2}|v_2\rangle...\langle u^{j_k}|v_k\rangle = (v_1)^{j_1} (v_2)^{j_2} .... (v_k)^{j_k} \qquad (5.1.3)_D$$

$$\langle u^{j_1}, u^{j_2} .....,u^{j_k} | u_{i_1},u_{i_2}, ... u_{i_k}\rangle = \delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k} \ . \qquad \text{// } \langle u^J|u_I\rangle = \delta^J{}_I \ . \qquad (5.1.4)_D$$

5.2 Tensor Expansion for a tensor in V^k ; the ordinary multiindex

Note: This section is subset of Section 2.10 (b) with new equation numbers and with Dirac notation equations added at the end.

A rank-k tensor T in V^k has this general expansion on the u_r basis,

$$T = \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} (u_{i_1} \otimes u_{i_2} .....\otimes u_{i_k}) . \qquad (5.2.1)$$

As expected,

$$[T]^{j_1 j_2 ... j_k} = \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} (u_{i_1} \otimes u_{i_2} .....\otimes u_{i_k})^{j_1 j_2 ... j_k}$$

$$= \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} (\delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k}) = T^{j_1 j_2 ... j_k} . \qquad (5.2.2)$$

The coefficients $T^{i_1 i_2 .... i_k}$ can be projected out from T as in (2.10.19),

$$(\mathbf{u}^{i_1} \otimes \mathbf{u}^{i_2} \otimes ... \otimes \mathbf{u}^{i_k}) \bullet T = T^{i_1 i_2 ... i_k} \qquad (5.2.3)$$

with an appropriate generalization of the dot product • to the space $V^k = V\otimes V...\otimes V$,

$$(\mathbf{v}_1 \otimes \mathbf{v}_2 ... \otimes \mathbf{v}_k) \bullet (\mathbf{u}^1 \otimes \mathbf{u}^2 ... \otimes \mathbf{u}^k) \equiv \Sigma_{i_1 i_2 .... i_k} (\mathbf{v}_1 \otimes \mathbf{v}_2 ... \otimes \mathbf{v}_k)^{i_1 i_2 .... i_k} (\mathbf{u}^1 \otimes \mathbf{u}^2 ... \otimes \mathbf{u}^k)_{i_1 i_2 .... i_k}$$

$$= \Sigma_{i_1 i_2 .... i_k} (\mathbf{v}_1)^{i_1} (\mathbf{v}_2)^{i_2} ... (\mathbf{v}_k)^{i_k} (\mathbf{u}^1)_{i_1} (\mathbf{u}^2)_{i_2} ... (\mathbf{u}^k)_{i_k} \qquad \text{// outer products}$$

$$= (\mathbf{v}_1 \bullet \mathbf{u}^1)(\mathbf{v}_2 \bullet \mathbf{u}^2) .... (\mathbf{v}_k \bullet \mathbf{u}^k) . \qquad \text{// } = (\mathbf{v}_1)^1 (\mathbf{v}_2)^2 \cdots (\mathbf{v}_k)^k \qquad (5.2.4)$$

Using the notion of a multiindex I (an *ordinary* multiindex),

$$I \equiv i_1, i_2, ...i_k \qquad \text{// each } i_s \text{ ranges } 1,2....n \qquad n = \dim(V) \qquad (5.2.5)$$

and a shorthand notation for the basis vectors

$$u_I \equiv u_{i_1} \otimes u_{i_2} .....\otimes u_{i_k} \qquad u^I \equiv u^{i_1} \otimes u^{i_2} .....\otimes u^{i_k} \qquad (5.2.6)$$

the expansion (5.2.1) can be stated in the following compact form,

$$T = \Sigma_I T^I u_I . \qquad \textit{(5.2.1)} \qquad (5.2.7)$$

and the coefficients T^I can be projected out according to (5.2.3),

$$u^I \bullet T = T^I . \qquad \textit{(5.2.3)} \qquad (5.2.8)$$

The Dirac notation restatements of selected equations above are ,

$$|T\rangle = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} \, | u_{i_1}, u_{i_2} \ldots\ldots, u_{i_k}\rangle \, . \qquad (5.2.1)_D$$

$$[T]^{j_1 j_2 \ldots j_k} = \langle u^{j_1}, u^{j_2} \ldots\ldots, u^{j_k} \, | T\rangle$$

$$= \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} \langle u^{j_1}, u^{j_2} \ldots\ldots, u^{j_k} \, | u_{i_1}, u_{i_2} \ldots\ldots, u_{i_k}\rangle$$

$$= \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} \, (\delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} \ldots\ldots \delta_{i_k}{}^{j_k}) = T^{j_1 j_2 \ldots j_k} \qquad (5.2.2)_D$$

or

$$[T]^J = \langle u^J | T\rangle = \Sigma_I T^I \langle u^J | u_I\rangle = \Sigma_I T^I \, \delta^J{}_I = T^J \, .$$

$$(\mathbf{u}^{i_1} \otimes \mathbf{u}^{i_2} \otimes \ldots \otimes \mathbf{u}^{i_k}) \bullet T = \langle \mathbf{u}^{i_1}, \mathbf{u}^{i_2}, \ldots \mathbf{u}^{i_k} \, | T\rangle = T^{i_1 i_2 \ldots i_k} \quad // = \mathcal{T}(\mathbf{u}^{i_1}, \mathbf{u}^{i_2}, \ldots \mathbf{u}^{i_k}) \qquad (5.2.3)_D$$

$$\langle v_1, v_2 \ldots v_k | \, u^1, u^2 \ldots u^k\rangle = \langle v_1 | u^1\rangle \langle v_1 | u^1\rangle \ldots \langle v_k | u^k\rangle = (\mathbf{v}_1)^1 (\mathbf{v}_2)^3 \cdots (\mathbf{v}_k)^k \qquad (5.2.4)_D$$

$$|T\rangle = \Sigma_I \, T^I \, |u_I\rangle \qquad (5.2.7)_D$$

$$\langle u^I | \, T\rangle = T^I \, . \qquad (5.2.8)_D$$

5.3 Rules for product of k vectors

The tensor product of k vectors is "k-multilinear" meaning it is linear in each of its k factors. This was discussed in (1.1.16) and later in (3.1.4). For example,

$$v_1 \otimes (v_2 + v'_2) \otimes v_3 \otimes \ldots\ldots \otimes v_k = v_1 \otimes v_2 \otimes v_3 \otimes \ldots\ldots \otimes v_k + v_1 \otimes v'_2 \otimes v_3 \otimes \ldots\ldots \otimes v_k$$

$$v_1 \otimes (s v_2) \otimes v_3 \otimes \ldots\ldots \otimes v_k = s(v_1 \otimes v_2 \otimes v_3 \otimes \ldots\ldots \otimes v_k) \qquad s = \text{scalar} \qquad (5.3.1)$$

Here we show linearity in the 2^{nd} factor. All the other factors have similar equations. We impose this k-multilinearity by fiat with the result that:

Fact: The space V^k is a vector space. (5.3.2)

The proof of this fact follows that of the text near (1.1.9). For example, the "0" in V^k is represented by (5.1.1) with one or more vectors being 0, since for example,

$$v_1 \otimes 0 \otimes \ldots\ldots \otimes v_k = v_1 \otimes (v_2 - v_2) \otimes \ldots\ldots \otimes v_k = v_1 \otimes v_2 \otimes \ldots\ldots \otimes v_k - v_1 \otimes v_2 \otimes \ldots\ldots \otimes v_k = 0 \, . \qquad (5.3.3)$$

"Vector multiplication" is distributive over scalar addition (here the "vector" is $v_1 \otimes v_2 \otimes \ldots\ldots \otimes v_k$), as one finds applying the rules (5.3.1),

$$(s_1 + s_2)(v_1 \otimes v_2 \otimes \,.....\otimes v_k) = [(s_1+s_2)v_1]\otimes v_2 \otimes \,.....\otimes v_k = [s_1 v_1 + s_2 v_1]\otimes v_2 \otimes \,.....\otimes v_k \qquad (5.3.4)$$

$$= s_1(v_1 \otimes v_2 \otimes \,.....\otimes v_k) + s_2(v_1 \otimes v_2 \otimes \,.....\otimes v_k) \qquad s_1, s_2 \in K$$

and multiplication by a scalar is distributive over "vector addition",

$$s[(v_1 \otimes v_2 \otimes \,.....\otimes v_k) + (v'_1 \otimes v'_2 \otimes \,.....\otimes v'_k)] = s\,(v_1 \otimes v_2 \otimes \,.....\otimes v_k) + s\,(v'_1 \otimes v'_2 \otimes \,.....\otimes v'_k)\,. \qquad (5.3.5)$$

All the above equations are meaningful for any positive integer k, regardless of the value $n = \dim(V)$.

5.4 The Tensor Algebra T(V)

Direct Sums

A direct sum of two vector spaces $Z = V\oplus W$ is a new vector space and has elements $v\oplus w$. Similarly, a direct sum of three vector spaces $Z = V\oplus W\oplus X$ is a new vector space with elements $v\oplus w\oplus x$. The idea can be applied to any number of vector spaces. Below we use $Z = V^0\oplus V^1\oplus V^2\oplus$ .... The reader unfamiliar with direct sums will find a detailed description in Appendix B including a simple "tall vector" method of visualizing such spaces.

The Tensor Algebra

Normally one does not add apples and oranges, so one does not add items of the form $a\otimes b \in V^2$ to those of the form $a\otimes b\otimes c \in V^3$. However (as Denker notes) fruit salad is great, and so we could define a very large vector space of the form

$$T(V) \equiv V^0 \oplus V \oplus V^2 \oplus V^3 \oplus \,....... \;=\; \Sigma^{\oplus}{}_{k=0}{}^{\infty}\, V^k\,. \qquad (5.4.1)$$

Here V^0 = the space of scalars, $V^1 = V$ the space of vectors, $V^2 = V\otimes V$ = the space of rank-2 tensors, and so on. The most general element t of the space T(V) has the form

$$t = s \oplus \Sigma_i T^i\, u_i \oplus \Sigma_{ij}\, T^{ij}\, u_i\otimes u_j \oplus \Sigma_{ijk}\, T^{ijk}\, u_i\otimes u_j\otimes u_k + \,...... \qquad s \in K \qquad (5.4.2)$$

with all coefficients in a field K.

Fact: This large space T(V) is in fact itself a vector space. (5.4.3)

We know this is true since $T(V) = \Sigma^{\oplus}{}_{k=0}{}^{\infty}\, V^k$ and we showed in (5.3.2) that each V^k is a vector space. For example, the "0" element in T(V) is the direct sum of the "0" elements of all the V^k. See Appendix B for more detail.

To show that T(V) is an algebra, we must show that it is closed under both addition and multiplication. It should be clear to the reader that T(V) is closed under addition and has the right scalar rule. For example, if k_1 and s are scalars,

$$k_1 \oplus a \oplus b\otimes c \oplus f\otimes g\otimes h = \text{sum of 4 elements of } T(V) = \text{an element of } T(V)$$

$$s(k_1 \oplus a \oplus b\otimes c \oplus f\otimes g\otimes h) = (sk_1) \oplus (sa) \oplus (sb)\otimes c \oplus f\otimes(sg)\otimes h = \text{element of } T(V) . \qquad (5.4.4)$$

This additive closure is of course necessary for T(V) be a vector space.

The space is also closed under the multiplication operation ⊗. For example

$$(b\otimes c)\otimes(f\otimes g\otimes h) = b\otimes c\otimes f\otimes g\otimes h = \in V^5 = \in T(V) . \qquad // (b\otimes c) \in V^2 , \ (f\otimes g\otimes h) \in V^3 \qquad (5.4.5)$$

Here we have used the associative property (2.8.21) applied to vectors. This closure claim is stated more generally in (5.6.6).

For later comparison with the corresponding wedge picture, here we have:

Object	lin comb is	Rank(grade)	Space	
s	scalar $\in$ K	0	V^0	
a	vector	1	V^1	
a⊗b	rank-2 tensor	2	V^2	
a⊗b⊗c	rank-3 tensor	3	V^3	
a⊗b⊗c⊗d	rank-4 tensor	4	V^4	
.....				
a⊗b⊗c⊗d....	rank-k tensor	k	V^k	
.....				
arbitrary element of T(V)	multivector	mixed	T(V)	(5.4.6)

Since T(V) is closed under the operations ⊕ and ⊗, it is "an algebra" (the space V^k alone is not an algebra because it is not closed under ⊗). The T(V) algebra is different from that of the reals due to its definition as a direct sum of vector spaces. The elements of T(V) have different "grades" as shown in the right column above, and T(V) is known therefore as a "graded algebra". The grade here is just the tensor rank. Sometimes T(V) is called "*the* tensor algebra" over V, see for example Benn and Tucker page 3.

Any linear combination of a set of tensor products of k vectors is a **rank-k tensor**. More generally, a rank-k tensor has the form shown in (5.2.1). A **multivector** is any linear combination of rank-k tensors for any mixed values of k

The dimensionality of the space T(V) is as follows, where n = dim(V),

$$\dim[T(V)] = 1 + n + n^2 + n^3 + ... = \infty \qquad (5.4.7)$$

Here are a few Dirac notation restatements of equations above

$k_1 \oplus |a\rangle \oplus |b,c\rangle \oplus |f,g,h\rangle$ = sum of 4 elements of T(V) = an element of T(V) (5.4.4)$_D$

$s(k_1 \oplus |a\rangle \oplus |b,c\rangle \oplus |f,g,h\rangle) = (sk_1) \oplus |sa\rangle \oplus |sb,c\rangle \oplus |f,sg,h\rangle$ = element of T(V) .

$|b,c\rangle \otimes |f,g,h\rangle = |b,c,f,g,h\rangle = \in V^5 = \in T(V)$. // $|b,c\rangle \in V^2$, $|f,g,h\rangle \in V^3$ (5.4.5)$_D$

5.5 Comments about tensors

The following fact is doubtless obvious to the reader, but we feel it is worth stating explicitly. First, suppose T^{ij} are the components of a rank-2 tensor. Define $Q^{ij} \equiv T^{ji}$. Then Q is also a rank-2 tensor (although one different from T if T is not symmetric). Here is a formal proof of this claim:

transformation (2.1.7) i↔j and a↔b reorder

$$(T = \text{rank-2 tensor}) \Rightarrow T'^{ij} = R^i{}_a R^j{}_b T^{ab} \Rightarrow T'^{ji} = R^j{}_b R^i{}_a T^{ba} = R^i{}_a R^j{}_b T^{ba}$$

$$\Rightarrow Q'^{ij} = T'^{ji} = R^i{}_a R^j{}_b Q^{ab} \Rightarrow (Q = \text{rank-2 tensor}) \qquad (5.5.1)$$

In similar fashion the reader can verify the following :

Fact: If $T^{i_1 i_2 \ldots i_k}$ are the components of a rank-k tensor, then $T^{j_1 j_2 \ldots j_k}$ are the also components of a rank-k tensor, where the $\{j_n\}$ are any permutation of the $\{i_n\}$. The permuted tensor is in general a different rank-k tensor from the unpermuted one. (5.5.2)

Corollary: Any linear combination of permutations of $T^{i_1 i_2 \ldots i_k}$ is a rank-k tensor. (5.5.3)

<u>Example:</u> If T^{ij} is a rank-2 tensor, then so is $A^{ij} = (T^{ij} - T^{ji})$. Thus, the A^{ij} shown in (4.3.10) is a rank-2 tensor given that T^{ij} is a rank-2 tensor.

5.6 The Tensor Product of two or more tensors in T(V)

The tensor algebra T(V) shown in (5.4.1) is closed under both ⊕ and ⊗. It seems evident how one would add two tensors of T(V) of the form (5.4.2), but how would one multiply two tensors?

During this set of steps, we try to gracefully transition into multiindex notation by doing a "real-time translation" for each line.

Consider two tensors of rank k and k' expanded as in (5.2.1),

$$T = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1} \otimes u_{i_2} \ldots\ldots \otimes u_{i_k}) . \qquad \text{rank k}, \quad T \in V^k \qquad (5.6.1)$$

$$\Sigma_I T^I u_I$$

$$S = \Sigma_{j_1 j_2 \ldots j_{k'}} S^{j_1 j_2 \ldots j_{k'}} (u_{j_1} \otimes u_{j_2} \ldots\ldots \otimes u_{j_{k'}}) \qquad \text{rank k'}, \quad S \in V^{k'} \quad . \qquad (5.6.2)$$

$$\Sigma_J S^J u_J$$

Multiplying these together with ⊗ one gets, using the rules (5.3.1),

$$T\otimes S = [\Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k}(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k})]\otimes[\Sigma_{j_1j_2....j_{k'}} S^{j_1j_2....j_{k'}}(u_{j_1}\otimes u_{j_2} ....\otimes u_{j_{k'}})]$$
$$[\ \Sigma_I T^I u_I] \qquad\qquad \otimes \qquad\qquad [\Sigma_J S^J u_J]$$

(a) $= \Sigma_{i_1i_2....i_k} \Sigma_{j_1j_2....j_{k'}} T^{i_1i_2....i_k} S^{j_1j_2....j_{k'}}(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k}) \otimes (u_{j_1}\otimes u_{j_2} .....\otimes u_{j_{k'}})$
$$\Sigma_{I,J} T^I S^J (u_I) \otimes (u_J)$$

(b) $= \Sigma_{i_1i_2....i_kj_1j_2....j_{k'}} T^{i_1i_2....i_k} S^{j_1j_2....j_{k'}}(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k}\otimes u_{j_1}\otimes u_{j_2} .....\otimes u_{j_{k'}})$
$$\Sigma_{I,J} T^I S^J (u_I \otimes u_J)$$

(c) $= \Sigma_{i_1i_2....i_ki_{k+1}i_{k+2}....i_{k+k'}}\ T^{i_1i_2....i_k} S^{i_{k+1}i_{k+2}....i_{k+k'}}\ (u_{i_1}\otimes u_{i_2} ......\otimes u_{i_{k+k'}})$.
$$\Sigma_{I,I'} T^I S^{I'} (u_I \otimes u_{I'})$$

(d) $= \Sigma_{i_1i_2....i_ki_{k+1}i_{k+2}....i_{k+k'}}\ [T\otimes S]^{i_1i_2....i_k\ i_{k+1}i_{k+2}....i_{k+k'}}\ (u_{i_1}\otimes u_{i_2} ......\otimes u_{i_{k+k'}})$.
$$\Sigma_{I,I'} [T\otimes S]^{I,I'} (u_I \otimes u_{I'})$$

(u) $= \Sigma_{i_1i_2...i_{k+k'}}\ [T\otimes S]^{i_1i_2...i_{k+k'}}\ (u_{i_1}\otimes u_{i_2} ......\otimes u_{i_{k+k'}})$. (5.6.3)
$$\Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ u_{\mathcal{I}} \ . \qquad\qquad \text{// italic I's}$$

Comparing lines one sees that

$$I \equiv i_1, i_2...i_k \qquad I' \equiv i_{k+1}, i_{k+2}, ...i_{k+k'} \qquad \mathcal{I} \equiv I, I' = i_1,i_2...i_{k+k'}$$
$$u_I \equiv (u_{i_1}\otimes u_{i_2} ....\otimes u_{i_k}) \quad u_{I'} \equiv (u_{i_{k+1}}\otimes...\otimes u_{i_{k+k'}}) \quad u_{\mathcal{I}} \equiv (u_{i_1}\otimes u_{i_2} ....\otimes u_{i_{k+k'}}) \ . \qquad (5.6.4)$$

Notice that the (2.8.21) associativity of ⊗ is used going from (a) to (b). In step (c) we renamed the dummy j_r summation indices so that $j_1 = i_{k+1}$, $j_2 = i_{k+2}$ and so on. Step (d) uses the outer product form (3.1.14) to replace $T^I\ S^{I'} = (T\otimes S)^{I,I'} = (T\otimes S)^{\mathcal{I}}$.

The conclusion is that

$$T\otimes S = \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ u_{\mathcal{I}} \qquad \mathcal{I} \equiv I, I' = i_1,i_2...i_{k+k'},\ \ u_{\mathcal{I}} \equiv (u_{i_1}\otimes u_{i_2} ....\otimes u_{i_{k+k'}}) \ . \qquad (5.6.5)$$

Since the $u_{\mathcal{I}}$ are basis vectors in $V^{k+k'}$, we have shown that:

$$T \in V^k \text{ and } S \in V^{k'} \quad \Rightarrow \quad T\otimes S \in V^{k+k'} \subset T(V) \ . \qquad (5.6.6)$$

Thus we have strengthened the claim made in (5.4.5) that T(V) is closed under the operation ⊗.

We shall now undertake the tensor product of three tensors T,S,R of ranks k,k',k" by mimicking the above set of steps, but leaning more heavily now on multiindex notation (no training wheels here),

$T\otimes S\otimes R = [\Sigma_I T^I u_I]\otimes[\Sigma_J S^J u_J]\otimes[\Sigma_K S^K u_K]$

(a) $= \Sigma_{I,J,K}\, T^I S^J R^K\ (u_I)\otimes(u_J)\otimes(u_K)$

(b) $= \Sigma_{I,J,K}\, T^I S^J R^K\ (u_I\otimes u_J\otimes u_K)$ // associative of ⊗ used here

(d) $= \Sigma_{I,I',I''}\, T^I S^{I'} R^{I''}\ (u_I\otimes u_{I'}\otimes u_{I''})$ // rename multiindices J→I',K→I"

$I \equiv i_1, i_2...i_k$ $I' \equiv i_{k+1}, i_{k+2}, ...i_{k+k'}$ $I'' \equiv i_{k+k'+1}, i_{k+k'+2}, ...i_{k+k'+k''}$
$u_I \equiv (u_{i_1}\otimes....\otimes u_{i_k})$ $u_{I'} \equiv (u_{i_{k+1}}\otimes...\otimes u_{i_{k+k'}})$ $u_{I''} \equiv (u_{i_{k+k'+1}}\otimes...\otimes u_{i_{k+k'+k''}})$

(e) $= \Sigma_{\mathcal{I}}\,(T\otimes S\otimes R)^{\mathcal{I}}\ u_{\mathcal{I}}$ $u_{\mathcal{I}} \equiv (u_{i_1}\otimes....\otimes u_{i_{k+k'+k''}})$ $\mathcal{I} \equiv I, I',I'' = i_1,i_2...i_{k+k'+k''}$ (5.6.7)

Now the outer product form is $T^I S^{I'} R^{I''} = (T\otimes S\otimes R)^{I,I',I''} = (T\otimes S\otimes R)^{\mathcal{I}}$.

The conclusion is this:

$$T\otimes S\otimes R = \Sigma_{\mathcal{I}}\,(T\otimes S\otimes R)^{\mathcal{I}} u_{\mathcal{I}} \quad \mathcal{I} \equiv I, I',I'' = i_1,i_2...i_{k+k'+k''} \quad u_{\mathcal{I}} \equiv (u_{i_1}\otimes....\otimes u_{i_{k+k'+k''}}) . \qquad (5.6.8)$$

Since the $u_{\mathcal{I}}$ are basis vectors in $V^{k+k'+k''}$, we have shown that:

$$T \in V^k \text{ and } S \in V^{k'} \text{ and } R \in V^{k''} \quad \Rightarrow \quad T\otimes S\otimes R \in V^{k+k'+k''} \subset T(V) . \qquad (5.6.9)$$

To develop a more systematic approach, consider the first three tensors in a product of tensors,

T_1 = tensor of rank k_1 $I_1 = \{i_1, i_2.....i_{k_1}\}$
T_2 = tensor of rank k_2 $I_2 = \{i_{k_1+1}, i_{k_1+2}.....i_{k_1+k_2}\}$
T_3 = tensor of rank k_3 $I_3 = \{i_{k_1+k_2+1}, i_{k_1+k_2+2}.....i_{k_1+k_2+k_3}\}$. (5.6.10)

Define the following "cumulative ranks",

$\kappa_1 = k_1$
$\kappa_2 = k_1 + k_2$
$\kappa_3 = k_1 + k_2 + k_3$
...
$\kappa_N = k_1 + k_2 + ... + k_N = \Sigma_{i=1}^N k_i$. (5.6.11)

Then rewrite and extend (5.6.10),

$$\begin{array}{ll} T_1 = \text{tensor of rank } k_1 & I_1 = \{i_1, i_2.....i_{\kappa_1}\} \\ T_2 = \text{tensor of rank } k_2 & I_2 = \{i_{\kappa_1+1}, i_{\kappa_1+2}.....i_{\kappa_2}\} \\ T_3 = \text{tensor of rank } k_3 & I_3 = \{i_{\kappa_2+1}, i_{\kappa_2+2}.....i_{\kappa_3}\} \\ ... & \\ T_s = \text{tensor of rank } k_s & I_s = \{i_{\kappa_{s-1}+1}, i_{\kappa_{s-1}+2}.....i_{\kappa_s}\} \\ ... & \\ T_N = \text{tensor of rank } k_N & I_N = \{i_{\kappa_{N-1}+1}, i_{\kappa_{N-1}+2}.....i_{\kappa_N}\} \ . \end{array} \qquad (5.6.12)$$

In this notation, and generalizing the above development for the tensor product of three tensors, we find the following expansion for the tensor product of N tensors of T(V),

$$T_1\otimes T_2\otimes...\otimes T_N = \Sigma_{\mathit{I}} (T_1\otimes T_2....\otimes T_N)^{\mathit{I}} u_{\mathit{I}}$$

$$\text{where} \qquad u_{\mathit{I}} = u_{i_1}\otimes u_{i_2} .....\otimes u_{i_{k_1+k_2+...+k_N}} = u_{i_1}\otimes u_{i_2} .....\otimes u_{\kappa_N} \qquad (5.6.13)$$

$$\text{and} \qquad (T_1\otimes T_2....\otimes T_N)^{\mathit{I}} = T_1{}^{I_1} T_2{}^{I_2} .... T_N{}^{I_N} \ .$$

The rank of this product tensor is then $\kappa = \Sigma_{i=1}{}^{N} k_i$ and the tensor is an element of $V^\kappa \subset T(V)$. In Dirac notation, one rewrites (5.6.13) as

$$| T_1,T_2....T_N> = | T_1\otimes T_2....\otimes T_N> = |T_1>\otimes|T_2>...\otimes|T_N> = \Sigma_{\mathit{I}} (T_1\otimes T_2....\otimes T_N)^{\mathit{I}} |u_{\mathit{I}}> . \qquad (5.6.13)_D$$

Example 1: The tensor product of two rank-1 tensors.

$$T\otimes S = \Sigma_{i_1i_2}[T^{i_1} S^{i_2}] (u_{i_1}\otimes u_{i_2}) = \Sigma_{ij} T^i S^j (u_i\otimes u_j) = \Sigma_{ij} (T\otimes S)^{ij} (u_i\otimes u_j)$$

$$(T\otimes S)^{ab} = \Sigma_{ij} T^i S^j (u_i\otimes u_j)^{ab} = \Sigma_{ij} T^i S^j \delta_i{}^a \delta_j{}^b = T^a S^b \qquad (5.6.14)$$

Example 2: The tensor product of two rank-2 tensors.

$$T\otimes S = \Sigma_{i_1i_2i_3i_4} T^{i_1i_2} S^{i_3i_4} (u_{i_1}\otimes u_{i_2} \otimes u_{i_3}\otimes u_{i_4})$$

$$(T\otimes S)^{abcd} = T^{ab} S^{cd} \qquad (5.6.15)$$

In both examples the evaluation of components produces the expected outer product forms.

Special cases of the tensor product T⊗S.

Assume T and S have rank k and k'.

If $S = \kappa' \in K$ = a **scalar**, then rank(S) = k' = 0 and (5.6.3) (c) reads,

$$T\otimes S = \Sigma_{i_1i_2....i_kj_1j_2....j_{k'}} T^{i_1i_2....i_k} S^{j_1j_2....j_{k'}}(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k}\otimes u_{j_1}\otimes u_{j_2} .....\otimes u_{j_{k'}})$$

$$\rightarrow \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k} (\kappa')\,(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k}) \;\; = \kappa'T$$

and

$$S\otimes T = \Sigma_{j_1j_2....j_{k'}i_1i_2....i_k} S^{j_1j_2....j_{k'}} T^{i_1i_2....i_k}(u_{j_1}\otimes u_{j_2} .....\otimes u_{j_{k'}}\otimes u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k})$$

$$\rightarrow \Sigma_{i_1i_2....i_k} (\kappa')\; T^{i_1i_2....i_k} \;(u_{i_1}\otimes u_{i_2} .....\otimes u_{i_k}) = \kappa'T$$

so we find that $T\otimes S = S\otimes T = \kappa'T$.

If $T = \kappa$ and $S = \kappa'$, the result above would be $T\otimes S = \kappa\kappa'$ and $S\otimes T = \kappa'\kappa$ and so $T\otimes S = S\otimes T = \kappa\kappa'$. Thus,

$$\begin{aligned}
&T\otimes S \;= \kappa\otimes S \;= S\otimes T \;= S\otimes\kappa \;= \kappa S && \text{if } T = \kappa \in V^0 \\
&T\otimes S \;= T\otimes\kappa' \;= S\otimes T \;= \kappa'\otimes T = \kappa'T && \text{if } S = \kappa' \in V^0 \\
&T\otimes S \;= \kappa\otimes\kappa' \;= S\otimes T \;= \kappa'\otimes\kappa \;= \kappa\kappa' && \text{if } T,S = \kappa,\kappa' \in V^0 \;.
\end{aligned} \tag{5.6.16}$$

All equations above can be written in Dirac notation, for example,

$$|T>\otimes|S> = [\Sigma_I T^I|u_I>] \otimes [\Sigma_{I'} S^{I'}|u_{I'}>] \;= \Sigma_{I,I'} T^I S^{I'}|u_I>\otimes|u_{I'}> \;= \Sigma_{\boldsymbol{I}} (T\otimes S)^{\boldsymbol{I}}\; |u_{\boldsymbol{I}}> \qquad (5.6.5)_D$$

$$|T> \in V^k \text{ and } |S> \in V^{k'} \;\;\Rightarrow\;\; |T\otimes S> = |T>\otimes|S> \;\in V^{k+k'} \subset T(V) \;. \qquad (5.6.6)_D$$

<u>Operators on the tensor product space</u>

Recall from above the following tensor product space vector,

$$|T_1,T_2....T_N> \;= |T_1\otimes T_2....\otimes T_N> \;= \;|T_1> \otimes |T_2>...\otimes |T_N> \qquad \textit{(5.6.13)}_D$$

which is an element of the tensor product space $V^{k_1} \otimes V^{k_2} \otimes ...\otimes V^{k_N}$. The action of a linear operator $\mathscr{P}$ on such a tensor product vector is defined in terms of its action in the spaces from which the tensor product is composed,

$$\mathscr{P}\,[\;|T_1> \otimes |T_2>...\otimes |T_N>] \;=\; \mathscr{P}\,|T_1> \otimes \mathscr{P}\,|T_2>...\otimes \mathscr{P}\,|T_N> \;. \tag{5.6.17}$$

6. The Tensor Product of k dual vectors : the vector spaces V^{*k} and T(V*)

Every equation in Chapter 5 can be converted to an appropriate equation of Chapter 6 using this simple set of translation rules:

1. $|X> \rightarrow <X|$ and $<Y| \rightarrow |Y>$. That is, reverse all Dirac bras and kets.

2. Swap lower and upper indices, indices. eg. $\mathbf{u_i} \rightarrow \mathbf{u^i}$, $T^{ij} \rightarrow T_{ij}$ (really: reverse all tilts).

3. $|v_i> \rightarrow <\alpha_i|$ // use Greek/script names for functionals; $v_i \rightarrow \alpha_i$

4. $V^k \rightarrow V^{*k}$ // space goes to dual space

5. $< T | v_1,v_2.....v_k > = \mathcal{T}(v_1,v_2.....v_k)$ = a tensor function (a new item) (6.1)

In general, translation of a Chapter 5 equation to Chapter 6 is most easily done if the Chapter 5 equation is first stated in Dirac notation.

We could end Chapter 6 right here, allowing the reader to apply the above rules, but that seems unsportsmanlike, so we proceed with a partial mimicry of Chapter 5.

6.1 Pure elements, basis elements, and dimension of V^{*k}

A generic pure ("decomposable") element of V^{*k} is this tensor product of k functionals,

$$\alpha_1 \otimes \alpha_2 \otimes .....\otimes \alpha_k \quad . \qquad \text{all } \alpha_i \in V^* \tag{6.1.1}$$

$$= <\alpha_1| \otimes <\alpha_2| ... \otimes <\alpha_k| = <\alpha_1,\alpha_2, ... \alpha_k| \quad . \qquad \text{// Dirac notation}$$

Since ⊗ is associative by (2.8.21), one can install parentheses anywhere in (6.1.1) without altering the meaning of the object, for example, $\alpha_1 \otimes (\alpha_2 \otimes \alpha_3) \otimes ....\otimes \alpha_k = \alpha_1 \otimes \alpha_2 \otimes a_3 \otimes ....\otimes \alpha_k$.

The basis elements of V^{*k} are (these u^i are those of Section 2.4),

$$\lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_k} = <u^{i_1}| \otimes <u^{i_2}| ... \otimes <u^{i_k}| = < u^{i_1}, u^{i_2} ...u^{i_k}| \quad . \tag{6.1.2}$$

The subscripts in (6.1.1) and the superscripts in (6.1.2) are labels, not components.

Equations corresponding to (5.1.3) and (5.1.4) are these

$$(\lambda^{j_1}\otimes \lambda^{j_2} .....\otimes \lambda^{j_k})(v_1,v_2.....v_k) = \lambda^{j_1}(v_1)\, \lambda^{j_2}(v_2) .....\lambda^{j_k}(v_k) = (v_1)^{j_1} (v_2)^{j_2} .... (v_k)^{j_k} \tag{6.1.3}$$

$$(\lambda^{j_1}\otimes \lambda^{j_2} .....\otimes \lambda^{j_k})(u_{i_1},u_{i_2}...u_{i_k}) = \lambda^{j_1}(u_{i_1})\, \lambda^{j_2}(u_{i_2}) .....\lambda^{j_k}(u_{i_k}) = \delta^{j_1}{}_{i_1}\, \delta^{j_2}{}_{i_2} ....\, \delta^{j_k}{}_{i_k} \,. \tag{6.1.4}$$

If n = dim(V), the total number of such basis elements is n^k, so

$$\dim(V^{*k}) = n^k. \tag{6.1.5}$$

In the full set of dual tensor-product basis elements shown in (6.1.2), two or more of the λ^{i_r} might be the same. This will always be the case if k > n where n ≡ dim(V*). For example, for k = 3 and n = 2 one such element would be $\lambda_1 \otimes \lambda_1 \otimes \lambda_2 \neq 0$.

In Dirac notation, we can write (6.1.3) and (6.1.4) as

$$<\lambda^{j_1}, \lambda^{j_2}...\lambda^{j_k} \,|\, v_1, v_2 ...,v_k > = <\lambda^{j_1}|v_1> <\lambda^{j_2}|v_1>... <\lambda^{j_k}|v_1> = (v_1)^{j_1} (v_2)^{j_2}...(v_k)^{j_k} \qquad (6.1.3)_D$$

$$<\lambda^{j_1}, \lambda^{j_2}...\lambda^{j_k} \,|\, u_{i_1}, u_{i_2} ...,u_{i_k} > = <\lambda^{j_1}|u_{i_1}> <\lambda^{j_2}|u_{i_2}>...<\lambda^{j_k}|u_{i_k}> = \delta^{j_1}{}_{i_1}\, \delta^{j_2}{}_{i_2}...\delta^{j_k}{}_{i_k} . \qquad (6.1.4)_D$$

6.2 Tensor Expansion for a tensor in V^{*k} ; the ordinary multiindex

We apply our translation rules to get this dense translation of Section 5.2 :

$$<T| = \Sigma_{i_1 i_2 ....i_k} T_{i_1 i_2....i_k} < u^{i_1}, u^{i_2} .....,u^{i_k}| . \qquad \text{tensor functional} \tag{6.2.1}$$

$$<T| = \Sigma_I T_I <u^I| \qquad <u^I| = < u^{i_1}, u^{i_2} .....,u^{i_k} | \qquad <T| u_I> = T_I \tag{6.2.2}$$

$$\mathcal{T} = \Sigma_{i_1 i_2....i_k} T_{i_1 i_2....i_k} (\lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_k}) . \tag{6.2.3}$$

$$\mathcal{T} = \Sigma_I T_I \lambda^I \qquad \lambda^I \equiv \lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_k} \tag{6.2.4}$$

$$T \bullet (u_{i_1}\otimes u_{i_2}\otimes ... \otimes u_{i_k}) = <T| u_{i_1}, u_{i_2} .....,u_{i_k}> = T_{i_1 i_2...i_k} = \mathcal{T}(u_{i_1}, u_{i_2} .....,u_{i_k}) \tag{6.2.5}$$

$$T \bullet u_I = <T|u_I> = T_I = \mathcal{T}(u_I) \tag{6.2.6}$$

$$T \bullet (v_1\otimes v_2\otimes ... \otimes v_k) = <T| v_1, v_2 .....,v_k> = \mathcal{T}(v_1, v_2 .....,v_k) \quad \text{tensor function} \tag{6.2.7}$$

$$T \bullet v_Z = <T|v_Z> = \mathcal{T}(v_Z) \qquad Z = 1,2..k \tag{6.2.8}$$

6.3 Rules for product of k vectors

The tensor product of k vectors is "k-multilinear" meaning it is linear in each of its k factors. This was discussed in (1.1.16) and later in (3.1.4). For example,

$$\alpha_1\otimes(\alpha_2 + \alpha'_2)\otimes\alpha_3\otimes .....\otimes\alpha_k = \alpha_1\otimes\alpha_2\otimes\alpha_3\otimes .....\otimes\alpha_k + \alpha_1\otimes\alpha'_2\otimes\alpha_3\otimes .....\otimes\alpha_k$$

$$\alpha_1\otimes(s\alpha_2)\otimes\alpha_3\otimes .....\otimes \alpha_k = s(\alpha_1\otimes\alpha_2\otimes\alpha_3\otimes .....\otimes\alpha_k) \qquad s = \text{scalar} . \tag{6.3.1}$$

Here we show linearity in the 2^{nd} factor. All the other factors have similar equations. We impose this k-multilinearity by fiat with the result that:

Fact: The space V^{*k} is a vector space. (6.3.2)

Proof: Repeat the discussion of Section 5.3 with all $v_i \rightarrow \alpha_i$, meaning $|v_i\rangle \rightarrow \langle\alpha_i|$.

6.4 The Tensor Algebra T(V*)

$$T(V^*) \equiv V^{*0} \oplus V^* \oplus V^{*2} \oplus V^{*3} \oplus .... = \Sigma^{\oplus}{}_{k=1}^{\infty} V^{*k} . \qquad (6.4.1)$$

Here V^{*0} = the space of scalars, V^{*1} = V the space of dual vectors, $V^{*2} = V^*\otimes^*V$ = the space of rank-2 dual tensors, and so on (tensor = functional). The most general element t of the space T(V*) has the form

$$\tau = s \oplus \Sigma_i T_i \lambda^i \oplus \Sigma_{ij} T_{ij} \lambda^i\otimes\lambda^j \oplus \Sigma_{ijk} T_{ijk} \lambda^i\otimes\lambda^j\otimes\lambda^k + ... \qquad s \in K \qquad (6.4.2)$$

with all coefficients in a field K.

Fact: This large space T(V*) is in fact itself a vector space. (6.4.3)

The proof is the same as that shown in Section 5.4 with a,b,c,d,e,f replaced by Greek letters. For example

$$k_1 \oplus \alpha \oplus \beta\otimes\kappa \oplus \rho\otimes\sigma\otimes\eta = \text{sum of 4 elements of } T(V^*) = \text{an element of } T(V^*)$$

$$s(k_1 \oplus \langle\alpha| \oplus \langle\beta,\kappa| \oplus \langle\rho,\sigma,\eta|) = (sk_1) \oplus (s\langle\alpha|) \oplus (s\langle\beta,\kappa|) \oplus (s\langle\rho,\sigma,\eta|) = \text{an element of } T(V^*) \qquad (6.4.4)$$

For later comparison with the corresponding dual wedge picture, here we have:

Object	lin comb is	Rank(grade)	Space	
s	scalar $\in$ K	0	V^{*0}	
α	dual vector	1	V^{*1}	
$\alpha\otimes\beta$	dual rank-2 tensor	2	V^{*2}	
$\alpha\otimes\beta\otimes\gamma$	dual rank-3 tensor	3	V^{*3}	
$\alpha\otimes\beta\otimes\gamma\otimes\delta$	dual rank-4 tensor	4	V^{*4}	
.....				
$\alpha\otimes\beta\otimes\gamma\otimes\delta$....	dual rank-k tensor	k	V^{*k}	
.....				
arbitrary element of T(V*)	dual multivector	mixed	T(V*)	(6.4.6)

All objects listed in the left column are tensor functionals, but we just call them tensors above and below.

Any linear combination of a set of tensor products of k dual vectors is a **dual rank-k tensor**. More generally, a dual rank-k tensor has the form shown in (6.2.3). A **dual multivector** is any linear combination of dual rank-k tensors for any mixed values of k .

The dimensionality of the space T(V*) is as follows, where $n = \dim(V^*)$,

$$\dim[T(V^*)] = 1 + n + n^2 + n^3 + \ldots = \infty \qquad (6.4.7)$$

6.5 Comments about Tensor Functions

For every rank-k tensor functional $\langle T| = \mathcal{T}$ in V^{*k} there exists a corresponding tensor function:

$$\mathcal{T}(v_1, v_2 \ldots, v_k) = \langle T|\, v_1,v_2 \ldots v_k\rangle \qquad // \ \mathcal{T}(v_Z) = \langle T|v_Z\rangle$$

$$\mathcal{T}(u_{i_1}, u_{i_2} \ldots, u_{i_k}) = \langle T|\, u_{i_1}, u_{i_2} \ldots, u_{i_k}\rangle = T_{i_1 i_2 \ldots i_k} \qquad // \ (6.2.5) \qquad (6.5.1)$$

There is a simple one-to-one relationship between the rank-k tensors $|T\rangle$ of V^k and the rank-k tensor functionals $\langle T|$ of V^{*k} and the rank-k tensor functions $\mathcal{T}(v_Z)$ of V^{*k}. These functions are manifestly k-multilinear since $|\, v_1,v_2 \ldots v_k\rangle = |\, v_1\rangle \otimes |\, v_2\rangle \ldots \otimes |\, v_k\rangle$ is k-multilinear. That is to say, each V space in the tensor product $V^k = V\otimes V \ldots \otimes V$ is a linear (vector) space.

Fact: The vector space V^{*k} is equivalent to the vector space $V^{*k}{}_f$ of k-multilinear *functions* on V^k. (6.5.2)

This is the generalization of Fact (4.2.15) from k = 2 to k = k.

The point made in Section 5.5 about tensors remaining tensors if their indices are shuffled around is reflected in the space of tensor functions: if $\mathcal{T}(v_1,v_2 \ldots, v_k)$ is a rank-k tensor function, then so is the function $\mathcal{T}(v_{i_1}, v_{i_2} \ldots, v_{i_k})$ where the arguments are any permutation of $v_1,v_2 \ldots, v_k$.

6.6 The Tensor Product of two or more tensors in T(V*)

Were we to write out the full detailed development of Section 5.6, it would begin as follows :

Consider two tensor functionals of rank k and k' expanded as in (6.2.3),

$$\mathcal{T} = \Sigma_{i_1 i_2 \ldots i_k}\ T_{i_1 i_2 \ldots i_k}\ \lambda^{i_1} \otimes \lambda^{i_2} \ldots \otimes \lambda^{i_k} \qquad \text{rank k}, \quad \mathcal{T} \in V^{*k}$$
$$\mathcal{S} = \Sigma_{j_1 j_2 \ldots j_{k'}}\ S_{j_1 j_2 \ldots j_{k'}}\ \lambda^{j_1} \otimes \lambda^{j_2} \ldots \otimes \lambda^{j_{k'}} \qquad \text{rank k'}, \quad \mathcal{S} \in V^{*k'} \quad . \qquad (6.6.1)$$

In multiindex and then Dirac notation these equations say

$$\mathcal{T} = \Sigma_I\, T_I\, \lambda^I \qquad \text{or} \qquad \langle T| = \Sigma_I\, T_I \langle u^I|$$
$$\mathcal{S} = \Sigma_I\, S_I\, \lambda^I \qquad \text{or} \qquad \langle S| = \Sigma_J\, S_J \langle u^J| \qquad (6.6.2)$$

and the tensor product of interest is

$$\mathcal{T}\otimes S = \langle T| \otimes \langle S| \ . \qquad (6.6.3)$$

The entire development proceeds as shown in Section 5.6 but with the translation rules outlined at the start of Chapter 6, in particular, that all bra-kets are reversed. One then finds for the tensor product of a rank-k tensor functional with a rank-k' one,

$$\mathcal{T}\otimes\mathcal{S} = \Sigma_{I,J} T_I S_J \lambda^I\otimes\lambda^J = \Sigma_{\boldsymbol{I}} (T\otimes S)_{\boldsymbol{I}} \lambda^{\boldsymbol{I}} \qquad (T\otimes S)_{\boldsymbol{I}} = T_I S_{I'} \qquad (6.6.4)$$
$$\boldsymbol{I} = \{i_1, i_2, .. i_{k+k'}\} \qquad \lambda^{\boldsymbol{I}} \equiv \lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_{k+k'}}$$
$$I = \{i_1, i_2, .. i_k\} \qquad I' = \{i_{k+1}, i_{k+2}, .. i_{k+k'}\}$$

which compare to the non-dual (5.6.5) (recall that $u_{\boldsymbol{I}} = <u_{\boldsymbol{I}}|$ and $\lambda^{\boldsymbol{I}} = <u^{\boldsymbol{I}}|$)

$$T\otimes S = \Sigma_{\boldsymbol{I}} (T\otimes S)^{\boldsymbol{I}} u_{\boldsymbol{I}} . \qquad (5.6.5)$$

A triple tensor product is then

$$\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R} = \Sigma_{I,J,K} T_I S_J R_K \lambda^I\otimes\lambda^J\otimes\lambda^K = \Sigma_{\boldsymbol{I}} (T\otimes S\otimes R)_{\boldsymbol{I}} \lambda^{\boldsymbol{I}} \qquad (T\otimes S\otimes R)_{\boldsymbol{I}} = T_I S_{I'} R_{I''} \qquad (6.6.5)$$
$$\boldsymbol{I} = \{i_1, i_2, .. i_{k+k'+k''}\} \qquad \lambda^{\boldsymbol{I}} \equiv \lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_{k+k'+k''}}$$
$$I = \{i_1, i_2, .. i_k\} \qquad I' = \{i_{k+1}, i_{k+2}, .. i_{k+k'}\} \qquad I'' = \{i_{k+k'+1}, i_{k+k'+2}, .. i_{k+k'+k''}\} .$$

For an arbitrary set of tensor functionals $\mathcal{T}_i$ of rank k_i the tensor product is

$$\mathcal{T}_1\otimes\mathcal{T}_2\otimes...\otimes\mathcal{T}_N = \Sigma_{\boldsymbol{I}} [(T_1)_{I_1}(T_2)_{I_2} .... (T_N)_{I_N}] \lambda^{\boldsymbol{I}} = \Sigma_{\boldsymbol{I}} (T_1\otimes T_2....\otimes T_N)_{\boldsymbol{I}} \lambda^{\boldsymbol{I}} \qquad (6.6.6)$$

$\mathcal{T}_i$ = tensor functional of rank k_i , I_i = multiindex range of i_r values for tensor $\mathcal{T}_i$

$I_1 = \{i_1, i_2.....i_{k_1}\}$, $I_2 = \{i_{k_1+1}, i_{k_1+2}.....i_{k_1+k_2}\}$, etc.

$\boldsymbol{I} = I_1 \cup I_2 .... \cup I_N = \{i_1, i_2......i_{k_1+k_2+\cdots k_N}\}$

$$\lambda^{\boldsymbol{I}} \equiv \lambda^{i_1}\otimes \lambda^{i_2} .....\otimes \lambda^{i_{k_1+k_2+...+k_N}} \qquad (6.6.7)$$

where the resulting tensor has a rank equal to the sum of the ranks of the combined tensors.

In Dirac notation, equation (6.6.6) is written,

$$< T_1,T_2....T_N| = < T_1\otimes T_2....\otimes T_N | = <T_1| \otimes<T_2|\otimes ...\otimes<T_N| = \Sigma_{\boldsymbol{I}} (T_1\otimes T_2....\otimes T_N)_{\boldsymbol{I}} <u^{\boldsymbol{I}}| \qquad (6.6.8)$$

which is just the Dirac transpose (plus tilt reversal) of the equation $(5.6.13)_D$ of Chapter 5.

It is understood here that each bra space fits the rank of its tensor, and one could write $<T_i| = {}_{k_i}<T_i|$ to make this fact more explicit. Then (6.6.8) would read

$${}_{k_1}<T_1| \otimes {}_{k_2}<T_2 |\otimes ...\otimes {}_{k_N}<T_N| = \Sigma_{\boldsymbol{I}} (T_1\otimes T_2....\otimes T_N)_{\boldsymbol{I}} \; {}_{\kappa_N}<u^{\boldsymbol{I}}| \qquad \kappa_N = k_1+k_2+...+k_N . \qquad (6.6.9)$$

Consider now the following generic tensor product ket,

$$|v_1,v_2...v_{k_1}\rangle_{k_1} \otimes |v_{k_1+1},v_{k_1+2}...v_{k_1+k_2}\rangle_{k_2} \otimes ....$$

$$= |v_{I_1}\rangle_{k_1} \otimes |v_{I_2}\rangle_{k_2} \otimes .... \otimes |v_{I_N}\rangle_{k_N}$$

$$= |v_{\mathcal{I}}\rangle_{\kappa_N}$$

$$= |\, v_1,v_2,v_3,v_4 ....v_{\kappa_N}\rangle \quad . \qquad (6.6.10)$$

If we close the bra (6.6.9) with this ket, we obtain the simple rule for the tensor product of the corresponding tensor functions,

$$\langle T_1\otimes T_2....\otimes T_N \,|\, v_1,v_2 ...v_{\kappa_N}\rangle$$

$$= [\, {}_{k_1}\langle T_1| \otimes {}_{k_2}\langle T_2 \,|\otimes ...\otimes {}_{k_N}\langle T_N|\,] \;[\, |v_{I_1}\rangle_{k_1} \otimes |v_{I_2}\rangle_{k_2} \otimes .... \otimes |v_{I_N}\rangle_{k_N}\,]$$

$$= {}_{k_1}\langle T_1|v_{I_1}\rangle_{k_1} * \;\; {}_{k_2}\langle T_2|v_{I_2}\rangle_{k_2} \;\; .... * {}_{k_N}\langle T_N|v_{I_N}\rangle_{k_N}$$

$$= \mathcal{T}_1(v_{I_1})\, \mathcal{T}_2(v_{I_2}) \;..... \mathcal{T}_N(v_{I_N}) \qquad (6.6.11)$$

or

$$(\mathcal{T}_1\otimes\mathcal{T}_2....\otimes\mathcal{T}_N)(v_1,v_2 ....v_{\kappa_N}) = (\mathcal{T}_1\otimes\mathcal{T}_2....\otimes\mathcal{T}_N)(v_{I_1},v_{I_2} ...v_{I_N})$$

$$= \mathcal{T}_1(v_{I_1})\, \mathcal{T}_2(v_{I_2}) \;..... \mathcal{T}_N(v_{I_N}) \;. \qquad (6.6.12)$$

<u>Example:</u> For N = 2 and k_1= k and k_2 = k' : (this appears on Spivak p 75)

$$(\mathcal{T}\otimes\mathcal{S})(v_1,v_2....v_k, v_{k+1}....v_{k+k'}) = \mathcal{T}(v_1,v_2....v_k)\, \mathcal{S}(v_{k+1},v_{k+2}....v_{k+k'}) \;. \qquad (6.6.13)$$

<u>Example:</u> For N = 3 and k_1= k and k_2 = k' and k_3 = k" :

$$(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R})(v_1, v_2, ... v_{k+k'+k''})$$

$$\mathcal{T}(v_1,v_2....v_k)\, \mathcal{S}(v_{k+1},v_{k+2}....v_{k+k'})\mathcal{R}(v_{k+k'+1},v_{k+k'+2}....v_{k+k'+k''}) \;. \qquad (6.6.14)$$

Here is an alternate proof of (6.6.12), independent of Chapter 5, where we make use of the dense multiindex notation :

Let

$$\mathcal{T}_i = \text{tensor of rank } k_i \quad I_i = \text{multiindex range of } i_r \text{ values for tensor } \mathcal{T}_i$$
$$I_1 = \{i_1, i_2.....i_{k_1}\},\ \ I_2 = \{i_{k_1+1}, i_{k_1+2}.....i_{k_1+k_2}\}, \text{etc.} \qquad (6.6.15)$$

Then we have

$$\mathcal{T}_1 \otimes \mathcal{T}_2 \otimes ... \otimes \mathcal{T}_N = \Sigma_{I_1 I_2 ... I_N} (T_1)_{I_1} (T_2)_{I_2} ... (T_N)_{I_N}\ \lambda^{I_1} \otimes \lambda^{I_2} \otimes ... \otimes \lambda^{I_N} \qquad (6.6.16)$$

$$(\mathcal{T}_1 \otimes \mathcal{T}_2 \otimes ... \otimes \mathcal{T}_N)(v_{I_1}, v_{I_2} ... v_{I_N})$$

$$= \Sigma_{I_1 I_2 ... I_N} (T_1)_{I_1} (T_2)_{I_2} ... (T_N)_{I_N} (\lambda^{I_1} \otimes \lambda^{I_2} \otimes ... \otimes \lambda^{I_N})(v_{I_1}, v_{I_2} ... v_{I_N})$$

$$= \Sigma_{I_1 I_2 ... I_N} (T_1)_{I_1} (T_2)_{I_2} ... (T_N)_{I_N}\ \ (v_{I_1})^{I_1} (v_{I_2})^{I_2} ...\ (v_{I_N})^{I_N}$$

$$= [\Sigma_{I_1} (T_1)_{I_1} (v_{I_1})^{I_1}]\ [\Sigma_{I_2} (T_2)_{I_2} (v_{I_2})^{I_1}]\ ...\ [\Sigma_{I_N} (T_2)_{I_N} (v_{I_N})^{I_N}]$$

$$= \mathcal{T}_1(v_{I_1})\ \mathcal{T}_2(v_{I_2})\ ....\ \mathcal{T}_N(v_{I_N})\ . \qquad (6.6.17)$$

Operators on the tensor product space

Recall from (6.6.8) the following tensor product space vector,

$$< T_1, T_2 .... T_N| \ = < T_1 \otimes T_2 .... \otimes T_N\ | \ = \ <T_1| \otimes <T_2| \otimes ... \otimes <T_N| \qquad \textit{(6.6.8)}$$

which is an element of the tensor product space $V^{*k_1} \otimes V^{*k_2} \otimes ... \otimes V^{*k_N}$. The action of a linear operator $\mathcal{Q}$ on such a tensor product vector is defined in terms of its action in the spaces from which the tensor product is composed,

$$[\ <T_1| \otimes <T_2| \otimes ... \otimes <T_N|\ ]\ \mathcal{Q}\ =\ <T_1|\ \mathcal{Q}\ \otimes <T_2|\ \mathcal{Q} \otimes ... \otimes <T_N|\ \mathcal{Q} \qquad (6.6.18)$$

This equation is the transpose of (5.6.17) if we set $\mathcal{Q} = \mathcal{P}^T$.

7. The Wedge Product of k vectors : the vector spaces L^k and L(V)

Wedge products and the spaces L^k and L(V) to be defined below were developed by Hermann Grassmann (1809-1877) in the 1840's. The algebra of these spaces is now called the exterior algebra and the wedge products are alternately called exterior products. Grassmann more or less invented the notions of linear algebra and vector spaces -- the so-called "modern algebra" did not exist. Other people were involved, but he was a very major pioneer. His work, naturally, was unappreciated at that time.

7.1 Definition of the wedge product of k vectors

We wish to define the wedge product of k vectors $v_i \in V$,

$v_1 \wedge v_2 \wedge ..... \wedge v_k$.   // $|v_1\rangle \wedge |v_2\rangle \wedge ..... \wedge |v_k\rangle$

Wedge products of this form (and their linear combinations) inhabit a vector space we call L^k.

We now *impose* the requirement that this wedge product must change sign when any two vectors are swapped. This property is injected into the wedge product theory, it does not fall out from it.

One motivation for the requirement relates to geometry. We showed in (4.3.14) that $\mathbf{a} \wedge \mathbf{b} = \det(\mathbf{a},\mathbf{b})\, \mathbf{u}_1 \wedge \mathbf{u}_2$ where $\det(\mathbf{a},\mathbf{b})$ is the signed area of the 2-piped (parallelogram) spanned by **a** and **b**. Then $\mathbf{b} \wedge \mathbf{a} = [\, -\det(\mathbf{a},\mathbf{b})]\, \mathbf{u}_1 \wedge \mathbf{u}_2$ has the same area but of opposite sign. One associates this sign with the "orientation" of the area in exactly the same sense that **a** x **b** and **b** x **a** represent areas of opposite sign.

So $\mathbf{b} \wedge \mathbf{a} = - \mathbf{a} \wedge \mathbf{b}$ reflects the change in orientation, as suggested by these drawings from Suter p 7,

(7.1.1)

For R^3, as shown in (4.3.15), one associates $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c}$ with a 3-piped whose "orientation" is determined by the sign of the volume $\det(\mathbf{a},\mathbf{b},\mathbf{c})$, which one can associate with the "handedness" of the 3-piped. For a k-piped it is hard to imagine "handedness", but it is easy to talk about orientation as the sign of $\det(\mathbf{a},\mathbf{b},\mathbf{c}....)$ where swapping any two vectors changes the sign of the "volume".

This sign-change requirement leads to the following candidate definition for the wedge product of k vectors in V (the j_r are vector labels),

$$v_{j_1}\wedge v_{j_2}\wedge \;.....\wedge v_{j_k} = (1/k!)\ \Sigma_P\ (-1)^{S(P)}\ (v_{j_{P(1)}} \otimes v_{j_{P(2)}} \otimes \;.....\otimes v_{j_{P(k)}})$$

$$= (1/k!)\ [\ (v_{j_1} \otimes v_{j_2} \otimes \;.....\otimes v_{j_k}) \quad + \quad \text{all signed permutations}\ ]$$

$$= \text{Alt}(v_{j_1} \otimes v_{j_2} \otimes \;.....\otimes v_{j_k})\ . \qquad (7.1.2)$$

An explanation of the $\Sigma_P\ (-1)^{S(P)}$ notation is presented in Section A.1: the sum is over all permutations P of [1,2..k], S(P) is the number of index swaps required to get from [1,2..k] to P[1,2..k], and $(-1)^{S(P)}$ is the parity of permutation P.

The important Alt operator is described generically in Section A.2 and is then applied to tensors in Section A.5. The definition of the Alt operator on the last line in (7.1.2) is the expression on the right side of the first line. The Alt operator definition for all authors includes a (1/k!) factor so that Alt(f) = f if the object f is already totally antisymmetric, (A.2.16). However, the (1/k!) factor appearing on the first line of (7.1.2) in the definition of the wedge product varies from author to author.

From our viewpoint, this (1/k!) **normalization** factor is just a convention that many authors use. However, Benn & Tucker (p 11 bottom and p 5 footnote) and Conrad (p 13 top) argue that the (1/k!) is in fact the "correct" normalization to be consistent with more elegant methods of defining the wedge product, as briefly reviewed in our Chapter 9. For other authors like Spivak, the (1/k!) on the first line of (7.1.2) is replaced by 1, resulting in

$$v_{j_1}\wedge v_{j_2}\wedge \;.....\wedge v_{j_k} = \Sigma_P\ (-1)^{S(P)}\ (v_{j_{P(1)}} \otimes v_{j_{P(2)}} \otimes \;.....\otimes v_{j_{P(k)}})$$

$$= (v_{j_1} \otimes v_{j_2} \otimes \;.....\otimes v_{j_k}) \quad + \quad \text{all signed permutations}$$

$$= k!\ \text{Alt}(v_{j_1} \otimes v_{j_2} \otimes \;.....\otimes v_{j_k})\ . \qquad (7.1.2)_{\text{Spivak}}$$

The implications of this Spivak normalization are described in Section 7.9(g) below. Notice that when the (1/k!) is present, (7.1.2) gives $v_1 \wedge v_2 = (1/2)(\ v_1\otimes v_2 - v_2\otimes v_1)$ which is the form already assumed in (4.3.1) and (4.4.1). But Spivak would say $v_1\wedge v_2 = v_1\otimes v_2 - v_2\otimes v_1$.

Almost everything one does with the wedge product is unaffected by the normalization choice.

For the purposes of this section, we simplify (7.1.2) by taking $j_r \rightarrow r$ to get,

$$v_1\wedge v_2\wedge \;.....\wedge v_k = (1/k!)\ \Sigma_P\ (-1)^{S(P)}\ (\ v_{P(1)}\otimes v_{P(2)}\otimes \;.....\otimes v_{P(k)})$$

$$= (1/k!)\ [\ (v_1 \otimes v_2 \otimes \;.....\otimes v_k) \quad + \quad \text{all signed permutations}\ ]\ .$$

$$= \text{Alt}(v_1 \otimes v_2 \otimes \;.....\otimes v_k)$$

$$= (1/k!)\ \Sigma_{i_1 i_2 ... i_k}\ \varepsilon_{i_1 i_2 ... i_k}\ (v_{i_1} \otimes v_{i_2} \otimes \;.....\; \otimes v_{i_k}) \qquad i_r = 1 \text{ to } k\ . \qquad (7.1.3)$$

We have added a fourth line using the permutation tensor $\varepsilon_{i_1 i_2 ... i_k}$. This tensor is described in Section A.6 and the equivalence of the first and fourth forms above is shown in (A.6.8).

Our approach here is that $v_1\wedge v_2\wedge \ldots\wedge v_k$ is defined in terms of $v_1\otimes v_2\otimes \ldots\otimes v_k$. In Section 9.1 it is shown how $v_1\wedge v_2\wedge \ldots\wedge v_k$ can be defined perhaps more elegantly in the language of modern algebra.

Examples

$$v_1\wedge v_2 = (1/2!)\,\Sigma_{a,b=1}{}^2\ \varepsilon_{ab}\, v_a \otimes v_b \qquad // \ 2! = 2 \text{ terms}$$

$$= (v_1 \otimes v_2 - v_2 \otimes v_1)/2 \qquad // \text{ agrees with (4.3.1)} \qquad (7.1.4)$$

$$v_1 \wedge v_2 \wedge v_3 = (1/3!)\,\Sigma_{a,b,c=1}{}^3\ \varepsilon_{abc}\ v_a \otimes v_b \otimes v_c \qquad // \ 3! = 6 \text{ terms}$$

$$= (v_1\otimes v_2\otimes v_3 - v_1\otimes v_3\otimes v_2 + v_3\otimes v_1\otimes v_2 - v_3\otimes v_2\otimes v_1 + v_2\otimes v_3\otimes v_1 - v_2\otimes v_1\otimes v_3)/6\ . \qquad (7.1.5)$$

7.2 Properties of the wedge product of k vectors

1. The sums in (7.1.2) and (7.1.3) have k! terms. (7.2.1)

Since there are k! permutations P of {1,2..k} (including the identity permutation) there are k! terms in the Σ_P sums in (7.1.2) and (7.1.3). Because $\varepsilon_{i_1 i_2 \ldots i_k}$ vanishes whenever two or more indices are the same, the ε tensor has k! non-zero components (k for the first index, (k-1) for the second index, and so on). Thus, the second sum in (7.1.3) has k! terms (not k^k), just like the first sum.

2. The wedge product is k-multilinear. (7.2.2)

It is by-fiat axiom that the wedge product of k vectors is k-multilinear and therefore satisfies these rules,

$$v_1\wedge(v_2 + v'_2)\wedge v_3\wedge\ldots\wedge v_k = v_1\wedge v_2\wedge v_3\wedge \ldots\wedge v_k + v_1\wedge v'_2\wedge v_3\wedge \ldots\wedge v_k$$

$$v_1\wedge(sv_2)\wedge v_3\wedge \ldots\wedge v_k = s(v_1\wedge v_2\wedge v_3\wedge \ldots\wedge v_k) \qquad s,r = \text{scalar} \in K$$

or

$$v_1\wedge(rv_2 + sv'_2)\wedge v_3\wedge\ldots\wedge v_k = r(v_1\wedge v_2\wedge v_3\wedge \ldots\wedge v_k) + s(v_1\wedge v'_2\wedge v_3\wedge \ldots\wedge v_k)\ . \qquad (7.2.3)$$

Here we show the rules just for the 2 position, but k-multilinear means these rules must apply to all the vector positions. These rules cannot be derived from the similar tensor product rules (5.3.1) except in the case k = 2 as was shown in (4.3.4).

Our candidate expansions (7.1.2) and (7.1.3) satisfy (7.2.3) because they are k-multilinear:

$$v_1\wedge (rv_2 + sv'_2) \wedge \ldots\wedge v_k = (1/k!)\,\Sigma_P\,(-1)^{S(P)}\,(v_{P(1)}\otimes [rv_{P(2)} + sv'_{P(2)}] \otimes \ldots\otimes v_{P(k)})$$

$$= (1/k!)\,\Sigma_P\,(-1)^{S(P)}\,\{ r(v_{P(1)}\otimes v_{P(2)}\otimes \ldots\otimes v_{P(k)}) + s(v_{P(1)}\otimes v'_{P(2)}\otimes \ldots\otimes v_{P(k)})\} \qquad // \ (5.3.1)$$

$$= r(1/k!)\,\Sigma_P\,(-1)^{S(P)}(v_{P(1)}\otimes v_{P(2)}\otimes \ldots\otimes v_{P(k)}) + s(1/k!)\,\Sigma_P\,(-1)^{S(P)}(v_{P(1)}\otimes v'_{P(2)}\otimes \ldots\otimes v_{P(k)})$$

$= \; r(v_1 \wedge v_2 \wedge v_3 \wedge \; ..... \wedge v_k) + s(v_1 \wedge v'_2 \wedge v_3 \wedge \; ..... \wedge v_k)$.

or equivalently

$$\text{Alt}(v_1 \otimes (rv_2 + sv'_2) \otimes v_3 \wedge \; ..... \otimes v_k) = r\,\text{Alt}(v_1 \otimes v_2 \otimes v_3 \wedge \; ..... \otimes v_k) + s\text{Alt}(v_1 \otimes v'_2 \otimes v_3 \wedge \; ..... \otimes v_k) .$$

Above we have used the fact that the ⊗ product is k-multilinear (also by fiat) as declared in (5.3.1).

3. The wedge product changes sign if any vector pair is swapped. (7.2.4)

Consider the last line of (7.1.2) which in effect says,

$$T_{j_1 j_2 \ldots j_k} = [\text{Alt}(F)]_{j_1 j_2 \ldots j_k} \qquad \text{or} \qquad T = \text{Alt}(F)$$

where

$$T_{j_1 j_2 \ldots j_k} \equiv v_{j_1} \wedge v_{j_2} \wedge \; ..... \wedge \; v_{j_k}$$
$$F_{j_1 j_2 \ldots j_k} \equiv v_{j_1} \otimes v_{j_2} \otimes \; ..... \otimes v_{j_k} .$$

But we know from (A.5.9) that $T_{j_1 j_2 \ldots j_k} \equiv [\text{Alt}(F)]_{j_1 j_2 \ldots j_k}$ is totally antisymmetric in the indices. Therefore $v_{j_1} \wedge v_{j_2} \wedge \; ..... \wedge v_{j_k}$ is totally antisymmetric in the labels, and so changes sign if any pair of labels is swapped.

Comment: From (7.2.2) the pure vector wedge product $v_1 \wedge v_2 \wedge \; ..... \wedge v_k$ is k-multilinear in the v_i, and so is the underlying tensor product $v_1 \otimes v_2 \otimes \; ..... \otimes v_k$.

4. Wedge product of vectors vanishes if any two vectors are the same.

Given a sign change (7.2.4) for any pair swap of vectors in the wedge product, we know that

$$v_1 \wedge v_2 \wedge \; ..... \wedge v_k = 0 \text{ if any two (or more) vectors are the same.} \qquad (7.2.5)$$

Proof: For example,

$$a \equiv v_2 \wedge v_1 \wedge \; ..... \wedge v_k = - v_1 \wedge v_2 \wedge \; ..... \wedge v_k = -a; \quad \text{if } 1 = 2 \text{ then } a = -a \text{ so } a = 0 .$$

5. Wedge product vanishes if vectors are linearly dependent. (7.2.6)

It was just shown that the wedge product vanishes if any two vectors are the same. It is also true that the wedge product $v_1 \wedge v_2 \wedge \; ..... \wedge v_k$ vanishes if the vectors v_i are linearly dependent. Linear dependence means one can write at least one vector in the set as a linear combination of the others, so perhaps $v_2 = (\Sigma_{i \neq 2}\, a_i v_i)$. Then

$$v_1 \wedge v_2 \wedge \; ..... \wedge v_k = v_1 \wedge (\Sigma_{i \neq 2}\, a_i v_i) \wedge \; ..... \wedge v_k$$

$= \Sigma_{i\neq 2}\, a_i\ (v_1\wedge v_i \wedge\ .....\wedge v_k)$. // since ^ is k-multilinear, see (7.2.3)

The sum $\Sigma_{i\neq 2}$ requires that index i be some other index appearing in $(v_1\wedge v_i \wedge\ .....\wedge v_k)$, but then one has two indices the same and by (7.2.5) it follows that $(v_1\wedge v_i \wedge\ .....\wedge v_k) = 0$ for each term in the sum. QED

[Grassmann also invented the notion of linear independence.]

<u>6. Wedge product vanishes if k > n</u> . (7.2.7)

If dim(V) = n, there can be at most n linearly independent vectors in V. If k > n, any set of k vectors v_i must be linearly dependent. Thus, by (7.2.6) the wedge product of any set of k vectors must vanish if k > n. Therefore for a given vector space V of dimension n, the only wedge products of interest are those for k = 1,2,3....n. For example, for n = 2 and k = 3 one has $e_1 \wedge e_1 \wedge e_2 = 0$.

<u>7. Components.</u> From (7.1.2) we find

$$(v_{j_1}\wedge v_{j_2}\wedge\ .....\wedge v_{j_k})^{i_1 i_2 \ldots i_k} = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, (v_{j_{P(1)}} \otimes v_{j_{P(2)}} \otimes\ .....\otimes v_{j_{P(k)}})^{i_1 i_2 \ldots i_k}$$

$$= (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, (v_{j_{P(1)}})^{i_1}(v_{j_{P(2)}})^{i_2} \ldots (v_{j_{P(k)}})^{i_k} \qquad \text{// outer product form}$$

$$= (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, (v_{j_1})^{i_{P(1)}}(v_{j_2})^{i_{P(2)}} \ldots (v_{j_k})^{i_{P(k)}} \qquad \text{// (A.1.19) with } M_a{}^b = (v_{j_a})^{i_b}$$

$$= (1/k!)\ \det[\, (v_{j_*})^{i_*}] \,. \qquad \text{// (A.1.19)} \qquad (7.2.8)$$

As noted earlier, in the Spivak normalization the factor (1/k!) is replaced by 1.

Fact: $(v_{j_1}\wedge v_{j_2}\wedge ...\wedge v_{j_k})^{i_1 i_2 \ldots i_k}$ is totally antisymmetric in both the labels j_r and the indices i_r. (7.2.9)

<u>Proof</u>: We already know from (7.2.4) that $(v_{j_1}\wedge v_{j_2}\wedge\ .....\wedge v_{j_k})^{i_1 i_2 \ldots i_k}$ is totally antisymmetric in the labels j_r. For antisymmetry on the i_r, we give two arguments. We can take components of (7.1.2) to get

$$(v_{j_1}\wedge v_{j_2}\wedge\ .....\wedge v_{j_k})^{i_1 i_2 \ldots i_k} = [\ \mathrm{Alt}(v_{j_1} \otimes v_{j_2} \otimes\ .....\otimes v_{j_k})]^{i_1 i_2 \ldots i_k}$$

which we think of as saying $T^{i_1 i_2 \ldots i_k} \equiv [\mathrm{Alt}(F)]^{i_1 i_2 \ldots i_k}$. According to (A.5.9) $T^{i_1 i_2 \ldots i_k}$ is totally antisymmetric in the i^r indices. Alternatively, $(v_{j_1}\wedge v_{j_2}\wedge\ .....\wedge v_{j_k})^{i_1 i_2 \ldots i_k} = (1/k!)\ \det[\ (v_{j_*})^{i_*}]$ is totally antisymmetric on the i_r because the det changes sign when any two rows or columns are swapped.

<u>8. Associative Property of the wedge product.</u>

This topic is addressed below in (7.9.2) where the need first arises. The conclusion there is that the wedge product is fully associative. For example, $(v_1\wedge v_2)\wedge v_3 = v_1\wedge (v_2\wedge v_3) = v_1\wedge v_2\wedge v_3$.

7.3 The vector space L^k and its basis

L^k is the space whose elements are all linear combinations of wedge products of k vectors of V. (7.3.1)

A more precise name for this space is $L^k(V)$ but we call it L^k.

L^k is a vector space (7.3.2)

Fact (5.3.2) showed that V^k is a vector space where the 0 element could be any V^k element such as $v_1 \otimes 0 \otimes \ldots \otimes v_k$. L^k is a vector space by a similar argument. It is closed under addition, scalars work correctly according to the rules (7.2.3), and the 0 element can be any element such as $v_1 \wedge 0 \wedge \ldots \wedge v_k$ as the reader can verify looking for example at (7.1.5).

A key point is that it is the imposition of the k-multilinear wedge product rules (7.2.3) that makes L^k be a vector space. We had a similar situation in Chapter 5 where the imposition of the k-multilinear tensor product rules (5.3.1) made V^k be a vector space.

Basis elements for L^k

Consider the following objects in L^k obtained by wedging together k basis elements of V, where each u_{j_r} is selected from the set of n available for V (which has dimension n),

$$(u_{j_1} \wedge u_{j_2} \wedge \ldots \wedge u_{j_k}) . \qquad (7.3.3)$$

Of these putative n^k objects, only n*(n-1)*...*(n-k+1) = n!/(n-k)! are non-zero by (7.2.5) because all the others have at least two vectors the same. Thus we can assume that all the labels j_r are different.

Now there exists a unique permutation P of the all-different labels j_r such that

$$[\, j_1, j_2 \ldots j_k] = P[\, i_1, i_2 \ldots i_k] \quad \textit{where} \quad i_1 < i_2 < \ldots < i_k . \qquad (7.3.4)$$

If this permutation involves S(P) pairwise swaps of indices, then

$$(u_{j_1} \wedge u_{j_2} \wedge \ldots \wedge u_{j_k}) = (-1)^{S(P)} (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \qquad \text{where} \quad i_1 < i_2 < \ldots < i_k \qquad (7.3.5)$$

because from (7.2.4) each pairwise swap of vectors in a wedge product creates a minus sign. Since there are k! possible permutations P, there are k! equations like (7.3.5) which relate different objects to the same object $(u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k})$ which has $i_1 < i_2 < \ldots < i_k$ .Thus, if we want to count the number of independent basis elements of L^k, we have to divide our earlier count of n!/(n-k)! non-vanishing objects by k!. The conclusion is that there are $\binom{n}{k}$ independent basis elements for L^k and they can be taken to have this form,

$$(u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \qquad \text{where} \quad i_1 < i_2 < \ldots < i_k \qquad \binom{n}{k} \text{ basis elements} \qquad (7.3.6)$$

Examples: (7.3.7)

• For k = 3 and n ≥ 5, the following k! = 3! basis elements involving u_1, u_3 and u_5 are all equal to the one ordered element u_1^u_3^u_5 with a + or - sign :

u_1^u_3^u_5 = $(-1)^0$ u_1^u_3^u_5 = +u_1^u_3^u_5 135
u_1^u_5^u_3 = $(-1)^1$ u_1^u_3^u_5 = - u_1^u_3^u_5 153→135
u_3^u_1^u_5 = $(-1)^1$ u_1^u_3^u_5 = - u_1^u_3^u_5 315 →135
u_3^u_5^u_1 = $(-1)^2$ u_1^u_3^u_5 = +u_1^u_3^u_5 351→315→135
u_5^u_1^u_3 = $(-1)^2$ u_1^u_3^u_5 = +u_1^u_3^u_5 513→153→135
u_5^u_3^u_1 = $(-1)^3$ u_1^u_3^u_5 = - u_1^u_3^u_5 531→513→153→135

• For k = 2 and n = 3, the 3 basis elements are u_1^u_2, u_1^u_3, u_2^u_3 and $\binom{3}{2} = 3$.

Fact: $(u_{j_1} \wedge u_{j_2} \wedge .... \wedge u_{j_k}) = \mathrm{Alt}(u_{j_1} \otimes u_{j_2} \otimes .... \otimes u_{j_k})$ $u_{\wedge J} = \mathrm{Alt}(u_J)$ (7.3.8)

This is a special case of (7.1.2) with v→u. The right shows equivalent multiindex notation.

Components of the basis elements for L^k

Now reconsider the basis vectors of the vector space L^k ,

$$(u_{j_1} \wedge u_{j_2} \wedge .... \wedge u_{j_k}) . \qquad (7.3.3)$$

The components are given from (7.2.8) with v = u as [recall $(u_i)^j = \delta_i{}^j$] ,

$$(u_{j_1} \wedge u_{j_2} \wedge .....\wedge u_{j_k})^{i_1 i_2 ... i_k}$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)}\, \delta_{j_{P(1)}}{}^{i_1} \delta_{j_{P(2)}}{}^{i_2} .... \delta_{j_{P(k)}}{}^{i_k}$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)}\, \delta_{j_1}{}^{i_{P(1)}} \delta_{j_2}{}^{i_{P(2)}} ... \delta_{j_k}{}^{i_{P(k)}} \qquad // \text{(A.1.19) with } M_a{}^b = \delta_{j_a}{}^{i_b}$$

$$= (1/k!)\ \det[\, \delta_{j_*}{}^{i_*}] . \qquad (u_{\wedge J})^I = (1/k!) \det(\delta_J{}^I) \qquad (7.3.9)$$

Once again, in Spivak normalization (1/k!) → 1.

Then (7.2.9) applied to v = u shows that,

Fact: $(u_{j_1} \wedge u_{j_2} \wedge ... \wedge u_{j_k})^{i_1 i_2 ... i_k}$ is totally antisymmetric in both the labels j_r and the indices i_r. (7.3.10)

We saw an example of both antisymmetries for k = 2 back in equation (4.3.21),

$$(u_i \wedge u_j)^{rs} = -(u_i \wedge u_j)^{sr} = -(u_j \wedge u_i)^{rs} . \qquad \text{// two forms of antisymmetry} \qquad \textit{(4.3.21)}$$

Either form in (7.3.9) can be expressed in our usual informal notation,

$$(u_{j_1} \wedge u_{j_2} \wedge .... \wedge u_{j_k})^{i_1 i_2 \cdots i_k} = (1/k!) \left[\delta_{j_1}{}^{i_1} \delta_{j_2}{}^{i_2} ... \delta_{j_k}{}^{i_k} + \text{signed permutations} \right] . \qquad (7.3.11)$$

Example:

$$3!\, (u_{j_1} \wedge u_{j_2} \wedge u_{j_3})^{i_1 i_2 i_3} = \det \begin{pmatrix} \delta_{j_1}{}^{i_1} & \delta_{j_1}{}^{i_2} & \delta_{j_1}{}^{i_3} \\ \delta_{j_2}{}^{i_1} & \delta_{j_2}{}^{i_2} & \delta_{j_2}{}^{i_3} \\ \delta_{j_3}{}^{i_1} & f_{j_3}{}^{i_2} & f_{j_3}{}^{i_3} \end{pmatrix} = \det \begin{pmatrix} \delta_{j_1}{}^{i_1} & \delta_{j_2}{}^{i_1} & \delta_{j_3}{}^{i_1} \\ \delta_{j_1}{}^{i_2} & \delta_{j_2}{}^{i_2} & \delta_{j_3}{}^{i_2} \\ \delta_{j_1}{}^{i_3} & \delta_{j_2}{}^{i_3} & \delta_{j_3}{}^{i_3} \end{pmatrix}$$

$$\equiv \det(\delta_{j_1 j_2 j_3}{}^{i_1 i_2 i_3})$$

$$= \det(\delta_J{}^I) . \qquad \text{// in multiindex notation} \qquad (7.3.12)$$

7.4 Tensor Expansions for a tensor in L^k

Recall now the tensor expansion for a most-general tensor T in V^k,

$$T = \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} (u_{i_1} \otimes u_{i_2} \otimes ..... \otimes u_{i_k}) . \qquad T \in V^k \qquad \textit{(5.2.1)} \qquad (7.4.1)$$

It has been shown in (7.3.10) that $(u_{i_1} \wedge u_{i_2} \wedge ..... \wedge u_{i_k})^{j_1 j_2 \cdots j_k}$ is totally symmetric in both the i_r and j_r indices. This object then meets the conditions of Fact (A.8.27) which states that $Alt_I = Alt_J$ when applied to such an object. Consider then this most-general object in L^k which has a similar look to (7.4.1) and for which we explicitly display the tensor components,

$$\Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} (u_{i_1} \wedge u_{i_2} \wedge ..... \wedge u_{i_k})^{j_1 j_2 \cdots j_k}$$

$$= \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} Alt_I[(u_{i_1} \otimes u_{i_2} \otimes ..... \otimes u_{i_k})^{j_1 j_2 \cdots j_k}] \qquad \text{// (7.3.8)}$$

$$= \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} Alt_I[\delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k}] \qquad \text{// (5.1.4)}$$

$$= \Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} Alt_J[\delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k}] \qquad \text{// (A.8.30)}$$

$$= Alt_J[\Sigma_{i_1 i_2 .... i_k} T^{i_1 i_2 .... i_k} \delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k} \qquad \text{// (A.5.10) Alt is linear}$$

$$= Alt_J(T^{j_1 j_2 \cdots j_k}) = Alt(T^{j_1 j_2 \cdots j_k}) \qquad \text{// no ambiguity}$$

$$= (1/k!) \Sigma_P (-1)^{S(P)} T^{j_{P(1)} j_{P(2)} \cdots j_{P(k)}} \qquad \text{// def of Alt (A.5.3b)}$$

$$= [Alt(T)]^{j_1 j_2 \cdots j_k} \qquad \text{// (A.5.3c)}$$

$$\equiv [T_\wedge]^{j_1 j_2 \cdots j_k} \qquad (7.4.2)$$

where we define this notation,

$$T^\wedge \equiv \text{Alt}(T) . \tag{7.4.3}$$

From (7.4.2) we then have the following fully general element of L^k,

$$T^\wedge = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \wedge \ldots\wedge u_{i_k}) . \qquad T^\wedge \in L^k \tag{7.4.4}$$

We refer to this type of expansion as a symmetric expansion, and we know it is redundant since the symmetric sum includes each true basis vector k! times.

According to (A.5.9), we know from (7.4.3) that

Fact: $T^{\wedge i_1 i_2 \ldots i_k}$ is a totally antisymmetric tensor. (7.4.5)

Therefore,

Fact: The space L^k is the space of all totally antisymmetric rank-k tensors $T^\wedge$. To say that $T^\wedge$ is totally antisymmetric means that $T^{\wedge i_1 i_2 \ldots i_k}$ is totally antisymmetric. (7.4.6)

In contrast, the space V^k is the space of *all* rank-k tensors T, so $L^k \subset V^k$.

Since the set $(u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k})$ with $1 \le i_1 < i_2 < \ldots < i_k \le n$ forms a complete basis for L^k, as discussed below (7.3.3), it must be possible to express $T^\wedge$ in the following manner

$$T^\wedge = \Sigma_{1 \le i_1 < i_2 < \ldots < i_k \le n} A^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) . \tag{7.4.7}$$

Example: If n = 3 and k = 2, then

$$T^\wedge = \Sigma_{1 \le i_1 < i_2 \le 3} A^{i_1 i_2} (u_{i_1} \wedge u_{i_2}) = A^{12} (u_1 \wedge u_2) + A^{13} (u_1 \wedge u_3) + A^{23} (u_2 \wedge u_3) . \tag{7.4.8}$$

What then is the connection between the $A^{i_1 i_2 \ldots i_k}$ of (7.4.7) and the $T^{i_1 i_2 \ldots i_k}$ of (7.4.4)?

Start with the symmetric form (7.4.4),

$$T^\wedge = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \qquad i_r = 1,2..n$$

$$= \Sigma_{i_1 \ne i_2 \ne \ldots \ne i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) . \qquad // (7.2.5) \tag{7.4.9}$$

Partition the summation space as follows ($1 \le i_r \le n$),

$$\Sigma_{i_1 \ne i_2 \ne \ldots \ne i_k} = [\, \Sigma_{i_1 < i_2 < \ldots < i_k} + \Sigma_{i_2 < i_1 < \ldots < i_k} + \text{many similar reorderings} \,] . \tag{7.4.10}$$

The total sum can be written in this manner, using the permutation sum notation,

$$\Sigma_{i_1\neq i_2\neq\ldots\neq i_k} = \Sigma_P \ \Sigma_{i_{P(1)}<i_{P(2)}<\ldots<i_{P(k)}} \qquad (7.4.11)$$

where P are the k! permutations of the k integers [1,2,...k].

Using the form (7.4.11), the sum (7.4.9) may be rewritten as,

$$T^{\wedge} = \Sigma_P \ \Sigma_{i_{P(1)}<i_{P(2)}<\ldots<i_{P(k)}} \ T^{i_1 i_2\ldots i_k} (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \ . \qquad (7.4.12)$$

In (A.9.1) it is shown that,

$$\Sigma_P \ [\Sigma_{i_{P(1)}<i_{P(2)}<\ldots<i_{P(k)}}] f_{i_1 i_2\ldots i_k} = \Sigma_{i_1<i_2<\ldots<i_k} [\Sigma_P \ f_{i_{P(1)})i_{P(2)}\ldots i_{P(k)}}] \ . \qquad \textit{(A.9.1)}$$

That is to say, in the Σ_P permutation sum, the permutation operators P can be moved from the summation index subscripts to the summand index subscripts. One then has from (7.4.12),

$$T^{\wedge} = \Sigma_{i_1<i_2<\ldots<i_k} \ \Sigma_P [\ T^{i_{P(1)} i_{P(2)}\ldots i_{P(k)}} (u_{i_{P(1)}} \wedge u_{i_{P(2)}} \wedge \ldots \wedge u_{i_{P(k)}}) \ ] \ . \qquad (7.4.13)$$

But we know from (7.3.5) that

$$(u_{i_{P(1)}} \wedge u_{i_{P(2)}} \wedge \ldots \wedge u_{i_{P(k)}}) = (-1)^{S(P)} (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \qquad (7.4.14)$$

where S(P) is the number of swaps associated with permutation P. Thus,

$$T^{\wedge} = \Sigma_{i_1<i_2<\ldots<i_k} [\Sigma_P \ (-1)^{S(P)} \ T^{i_{P(1)} i_{P(2)}\ldots i_{P(k)}}] (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \qquad (7.4.15)$$

which we can compare with the ordered sum (7.4.7),

$$T^{\wedge} = \Sigma_{i_1<i_2<\ldots.<i_k} \ A^{i_1 i_2\ldots i_k} \ (u_{i_1} \wedge u_{i_2} \wedge \ldots \wedge u_{i_k}) \ . \qquad \textit{(7.4.7)}$$

Thus, since the basis is complete, the relation between the A and T coefficients is given by

$$A^{i_1 i_2\ldots i_k} = \ \Sigma_P (-1)^{S(P)} \ T^{i_{P(1)} i_{P(2)}\ldots i_{P(k)}} \qquad i_1 < i_2 < \ldots < i_k$$

$$= [\ T^{i_1 i_2\ldots i_k} + \text{ all signed permutations }] \qquad \text{// k! terms}$$

$$= k! \ [\mathrm{Alt}(T)]^{i_1 i_2\ldots i_k} . \qquad \text{// (A.5.3) def of Alt}$$

or

$$A = k!\mathrm{Alt}(T) = k! \ T^{\wedge}. \qquad \text{// (7.4.3)} \qquad (7.4.16)$$

The $A^{i_1i_2...i_k}$ appear in the expansion (7.4.7) only for index values $1 \le i_1 < i_2 < ..... < i_k \le n$, but we can interpret (7.4.16) as defining $A^{i_1i_2...i_k}$ for all index values. Since $A = k!\, T_\wedge$, (7.4.5) shows that

Fact: $A^{i_1i_2...i_k}$ and $T_\wedge{}^{i_1i_2...i_k}$ are both totally antisymmetric tensors. (7.4.17)

Comment: $T^{i_1i_2...i_k}$ and $A^{i_1i_2...i_k}$ are both rank-k tensors, see (5.5.3).

Examples: (relating the A and T coefficients)

$$A^{ab} = T^{ab} - T^{ba} \qquad \text{// as in (4.3.10)} \qquad k = 2$$

$$A^{abc} = T^{abc} - T^{acb} + T^{cab} - T^{cba} + T^{bca} - T^{bac} \qquad k = 3 \qquad (7.4.18)$$

Vector Case. For k = 1, we find that

$$T = \Sigma_{i_1} T^{i_1}\, u_{i_1} \qquad \text{// (5.2.1)}$$
$$T_\wedge = \Sigma_{i_1} T^{i_1}\, u_{i_1} \qquad \text{// (7.4.4)} \qquad \Rightarrow \qquad T_\wedge = T \qquad (7.4.19)$$

so for a vector there is no distinction between $T_\wedge$ and T (and in fact $V^1 = L^1$).

7.5 Various expansions for the wedge product of k vectors

The *symmetric* expansion for the wedge product of k vectors is very straightforward. Let

$$T^{i_1i_2...i_k} = (v_1)^{i_1} (v_2)^{i_2} ... (v_k)^{i_k} = (v_1 \otimes v_2 \otimes .....\otimes v_k)^{i_1i_2...i_k}$$
or (7.5.1)
$$T = (v_1 \otimes v_2 \otimes .....\otimes v_k) \, .$$

Then the symmetric expansion (7.4.4) gives,

$$T_\wedge = \Sigma_{i_1i_2...i_k} T^{i_1i_2...i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) \qquad \textit{(7.4.4)}$$

$$= \Sigma_{i_1i_2...i_k} (v_1)^{i_1} (v_2)^{i_2} ... (v_k)^{i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) \qquad \text{// (7.5.1)} \qquad (7.5.2)$$

$$= [\Sigma_{i_1}(v_1)^{i_1} u_{i_1}] \wedge [\Sigma_{i_2}(v_2)^{i_2} u_{i_2}] \wedge .... \wedge [\Sigma_{i_k} (v_k)^{i_k} u_{i_k}] \qquad \text{// } \wedge \text{ is multilinear}$$

$$= v_1 \wedge v_2 \wedge ... \wedge v_k \, . \qquad (7.5.3)$$

This pure tensor $T_\wedge = v_1 \wedge v_2 \wedge ... \wedge v_k$ is an element of L^k .

Expressing $v_1 \wedge v_2 \wedge ... \wedge v_k$ in terms of the *ordered* expansion is more complicated. One must first compute the tensor A as in (7.4.16),

$$(1/k!)\ A = \text{Alt}(T) = \text{Alt}(v_1 \otimes v_2 \otimes .....\otimes v_k) = T_{\wedge} = (v_1 \wedge v_2 \wedge ... \wedge v_k)\ . \tag{7.5.4}$$

Then the ordered expansion (7.4.7) can be written in a battery of ways,

$$v_1 \wedge v_2 \wedge ... \wedge v_k = \qquad // \ v_1 \wedge v_2 \wedge ... \wedge v_k \in L^k$$

(a) $= \Sigma_{i_1<i_2<....<i_k} A^{i_1 i_2 ... i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$ // (7.4.7)

(b) $= \Sigma_{i_1<i_2<....<i_k} k!\ [\text{Alt}(v_1 \otimes v_2 \otimes ... \otimes v_k)]^{i_1 i_2 ... i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$ // (7.5.4)

(c) $= \Sigma_{i_1<i_2<....<i_k} k!\ [v_1 \wedge v_2 \wedge ... \wedge v_k]^{i_1 i_2 ... i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$ // (7.1.3)

(d) $= \Sigma_{i_1<i_2<....<i_k} \det[\ (v_*)^{i*}]\ (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$ // (7.2.8) with $j_r \rightarrow r$

(e) $= \Sigma_{i_1<i_2<....<i_k} \det \begin{pmatrix} (v_1)^{i_1} & (v_1)^{i_2} & ... & (v_1)^{i_k} \\ (v_2)^{i_1} & (v_2)^{i_2} & ... & (v_2)^{i_k} \\ ... & ... & ... & ... \\ (v_k)^{i_1} & (v_k)^{i_2} & ... & (v_k)^{i_k} \end{pmatrix} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$

and with a transposes matrix

(f) $= \Sigma_{i_1<i_2<....<i_k} \det \begin{pmatrix} (v_1)^{i_1} & (v_2)^{i_1} & ... & (v_k)^{i_1} \\ (v_1)^{i_2} & (v_2)^{i_2} & ... & (v_k)^{i_2} \\ ... & ... & ... & ... \\ (v_1)^{i_k} & (v_2)^{i_k} & ... & (v_k)^{i_k} \end{pmatrix} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$

(g) $= \Sigma_{i_1 i_2 ... i_k} (v_1)^{i_1} (v_2)^{i_2} ... (v_k)^{i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k})$ // (7.5.2) (7.5.5)

where we throw in the symmetric sum at the end. Remember that, since generally dim(V) = n > k, the determinant in (f) is a full-width minor of matrix $M = [\mathbf{v}_1, \mathbf{v}_2.....\mathbf{v}_k]$. If k = n, the minor is the full matrix.

<u>Example</u> : Suppose k = n = 3. Then the following sum (form (f)) has only one term in which $i_1 = 1$, $i_2 = 2$ and $i_3 = 3$,

$$\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \mathbf{v}_3 = \Sigma_{1 \le i_1<i_2<i_3<3} \det \begin{pmatrix} (v_1)^{i_1} & (v_2)^{i_1} & (v_3)^{i_1} \\ (v_1)^{i_2} & (v_2)^{i_2} & (v_3)^{i_2} \\ (v_1)^{i_3} & (v_2)^{i_3} & (v_3)^{i_3} \end{pmatrix} (\mathbf{u}_{i_1} \wedge \mathbf{u}_{i_2} \wedge \mathbf{u}_{i_3})$$

$$= \det[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]\ (\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \mathbf{u}_3) \tag{7.5.6}$$

as quoted in (4.3.15).

Example : Here are the above expressions for k = 2 and general n ≥ k :

(a) $v_1 \wedge v_2 = \Sigma_{i_1<i_k} A^{i_1 i_2} (u_{i_1} \wedge u_{i_2})$

(b) $= \Sigma_{i_1<i_2} 2! [\text{Alt}(v_1 \otimes v_2)]^{i_1 i_2} (u_{i_1} \wedge u_{i_2})$

(c) $= \Sigma_{i_1<i_2} 2! (v_1 \wedge v_2)^{i_1 i_2} (u_{i_1} \wedge u_{i_2})$

(d) $= \Sigma_{i_1<i_2} \det[(v_*)^{i*}] (u_{i_1} \wedge u_{i_2})$

(e) $= \Sigma_{i_1<i_2} \det\begin{pmatrix} (v_1)^{i_1} & (v_1)^{i_2} \\ (v_2)^{i_1} & (v_2)^{i_2} \end{pmatrix} (u_{i_1} \wedge u_{i_2})$

(f) $= \Sigma_{i_1<i_2} \det\begin{pmatrix} (v_1)^{i_1} & (v_2)^{i_1} \\ (v_1)^{i_2} & (v_2)^{i_2} \end{pmatrix} (u_{i_1} \wedge u_{i_2}) = \Sigma_{i_1<i_2} [(v_1)^{i_1}(v_2)^{i_2} - (v_2)^{i_1}(v_1)^{i_2}] (u_{i_1} \wedge u_{i_2})$

(g) $= \Sigma_{i_1 i_2} (v_1)^{i_1} (v_2)^{i_2} (u_{i_1} \wedge u_{i_2}) = [\Sigma_{i_1}(v_1)^{i_1} u_{i_1}] \wedge [\Sigma_{i_2}(v_2)^{i_2} u_{i_2}] = v_1 \wedge v_2$ (7.5.7)

Result (f) matches that shown in (4.3.12),

$$a \wedge b = \Sigma_{ij} a^i b^j (u_i \wedge u_j) = \Sigma_{i<j} (a^i b^j - a^j b^i) (u_i \wedge u_j) = \Sigma_{i<j} A^{ij} (u_i \wedge u_j)$$

$$= \Sigma_{i<j} \det\begin{pmatrix} a^i & b^i \\ a^j & b^j \end{pmatrix} (u_i \wedge u_j) \qquad A^{ij} = (a^i b^j - a^j b^i) = \det\begin{pmatrix} a^i & b^i \\ a^j & b^j \end{pmatrix} . \qquad \textit{(4.3.12)}$$

7.6 Number of elements in L^k compared with V^k.

We know from (5.1.5) and (7.3.6) that,

$\dim(V^k) = n^k$ // number of basis elements of V^k *(5.1.5)*

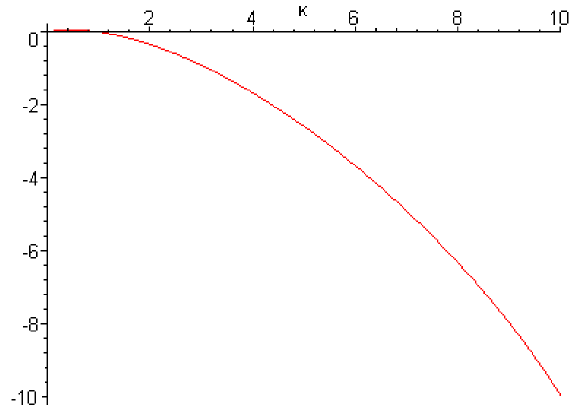
$\dim(L^k) = \binom{n}{k}$ // number of basis elements of L^k *(7.3.6)*

If the number of elements of field K is N (N → ∞ for K= reals), then the generalization of (4.3.11) is,

$$\text{ratio} = \frac{\#\text{ elements of } L^k}{\#\text{ elements of } V^k} = \frac{\binom{n}{k} N}{n^k N} = \frac{\binom{n}{k}}{n^k} = \binom{n}{k} / n^k . \qquad (7.6.1)$$

For a given n, this is a strongly decreasing function of k. For example, for n = 10 we can plot the log of the ratio for k = 0 to 10,

```
ratio := binomial(n,k)/(n^k):
n := 10:
plot(log10(ratio),k = 0..10);
```



(7.6.2)

For example, when k = n = 10, ratio = $\binom{10}{10}$ / $10^{10} = 10^{-10}$ and L^{10} has only one non-zero element.

7.7 Multiindex notation

In this section, multiindex versions of equations are shown in red.

Multiindexing is done in two different ways. First, for the symmetric expansion (7.4.4) :

$$T_{\wedge} = \Sigma_{i_1 i_2 ... i_k} T^{i_1 i_2 ... i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) \qquad (7.4.4)$$

$T_{\wedge} = \Sigma_I T^I u_{\wedge I}$ where $u_{\wedge I} \equiv u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}$ $\qquad T^I \equiv T^{i_1 i_2 ... i_k}$

and $I \equiv \{i_1, i_2, ... i_k\}$ with $1 \le i_r \le n$ = *ordinary* multiindex, $n = \dim(V)$. (7.7.1)

The more significant notation involves the ordered expansion (7.4.7) which has only one term for each linearly independent basis element. Note our use of Σ'_I (prime) to indicate an ordered multiindex summation :

$$T_{\wedge} = \Sigma_{i_1 < i_2 < .... < i_k} A^{i_1 i_2 ... i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) \qquad (7.4.7)$$

$T_{\wedge} = \Sigma'_I A^I u_{\wedge I}$ where $u_{\wedge I} \equiv u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}$ $\qquad A^I \equiv A^{i_1 i_2 ... i_k}$

and $I \equiv \{i_1, i_2, ... i_k\}$ with $1 \le i_1 < i_2 < .... < i_k \le n$ = *ordered* multiindex, $n = \dim(V)$. (7.7.2)

Here are some unofficial multiindex notations for other equations developed above:

$$T_\wedge = v_1 \wedge v_2 \wedge ... \wedge v_k \qquad T_\wedge = (\wedge v_Z) \qquad (7.5.3)$$

$$T^{i_1 i_2 ... i_k} = (v_1)^{i_1} (v_2)^{i_2} ... (v_k)^{i_k} \equiv (v_Z)^I \qquad T^I = (v_Z)^I \qquad (7.5.1)$$

with the idea that Z = 1,2...k . Continuing on,

$$T_\wedge = \Sigma_{i_1 i_2 ... i_k} (v_1)^{i_1} (v_2)^{i_2} ... (v_k)^{i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) \qquad T_\wedge = \Sigma_I (v_Z)^I u_{\wedge I} \qquad (7.5.2)$$

$$A = k!\, \text{Alt}(v_1 \otimes v_2 \otimes ... \otimes v_k) \qquad A = k!\, \text{Alt}(\otimes v_Z) \qquad (7.5.4)$$

$$v_1 \wedge v_2 \wedge ... \wedge v_k = \Sigma_{i_1 < i_2 < .... < i_k} k![\text{Alt}(v_1 \otimes v_2 \otimes ... \otimes v_k)]^{i_1 i_2 ... i_k} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) \quad (7.5.5b)$$

$$(\wedge v_Z) = \Sigma'_I\ k!\, \text{Alt}(\otimes v_Z)^I u_{\wedge I}$$

$$A^{i_1 i_2 ... i_k} = \det \begin{pmatrix} (v_1)^{i_1} & (v_2)^{i_1} & ... & (v_k)^{i_1} \\ (v_1)^{i_2} & (v_2)^{i_2} & ... & (v_k)^{i_2} \\ ... & ... & ... & ... \\ (v_1)^{i_k} & (v_2)^{i_k} & ... & (v_k)^{i_k} \end{pmatrix} \qquad A^I = \det(v_Z{}^I) \qquad (7.5.5a+f)$$

$$v_1 \wedge v_2 \wedge \ .....\wedge v_k = \Sigma_{i_1 < i_2 < .... < i_k} \det \begin{pmatrix} (v_1)^{i_1} & (v_2)^{i_1} & ... & (v_k)^{i_1} \\ (v_1)^{i_2} & (v_2)^{i_2} & ... & (v_k)^{i_2} \\ ... & ... & ... & ... \\ (v_1)^{i_k} & (v_2)^{i_k} & ... & (v_k)^{i_k} \end{pmatrix} (u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}). \quad (7.5.5f)$$

$$(\wedge v_Z) = \Sigma'_I\ \det(v_Z{}^I)\, u_{\wedge I}$$

7.8 The Exterior Algebra L(V)

We now construct the graded algebra L(V) in analogy with that of T(V) in (5.4.1).

Define a large vector space of the form (this is "the exterior algebra on V")

$$L(V) \equiv L^0 \oplus L^1 \oplus L^2 \oplus L^3 + ... \qquad // \ L(V) = \Sigma^{\oplus}{}_{k=0}{}^{\infty} L^k \qquad (7.8.1)$$

Here L^0 = the space of scalars, $L^1 = V$ the space of vectors, $L^2 = V \wedge V \subset V^2$ the space of antisymmetric rank-2 tensors (7.4.6), and so on. The most general element of the space L(V) would have the form

$$X = s \oplus \Sigma_i T^i u_i \oplus \Sigma_{ij} T^{ij} u_i \wedge u_j \oplus \Sigma_{ijk} T^{ijk} u_i \wedge u_j \wedge u_k + ... \qquad // \text{ symmetric}$$

or

$$X = s \oplus \Sigma_i T^i u_i \oplus \Sigma_{i<j} A^{ij} u_i \wedge u_j \oplus \Sigma_{i<j<k} A^{ijk} u_i \wedge u_j \wedge u_k + ... \qquad // \text{ ordered} \qquad (7.8.2)$$

The direct sum $\oplus$ is described in Appendix B.

Associativity of the Wedge Product

We have carefully managed to avoid this topic in all that has transpired above. Nothing so far has been assumed concerning associativity of the ^ operator. In (2.8.21) it was stated that the ⊗ operator is associative, and this was "proved" in our outer product approach to ⊗, but for the formal approaches of Chapter 1 it is an axiom that ⊗ is associative.

Once we define the space L(V) above, we must face the issue of wedge products of the form $(u_i \wedge u_j) \wedge u_k$ and more generally $(v_1 \wedge v_2) \wedge v_3$. These products arise when we multiply an element of L^2 by an element of L^1. Notice that our grandiose expansion (7.1.3) says *nothing* about $(v_1 \wedge v_2) \wedge v_3$. All it says is this:

$$v_1 \wedge v_2 \wedge v_3 = (v_1 \otimes v_2 \otimes v_3 - v_1 \otimes v_3 \otimes v_2 + v_3 \otimes v_1 \otimes v_2 - v_3 \otimes v_2 \otimes v_1 + v_2 \otimes v_3 \otimes v_1 - v_2 \otimes v_1 \otimes v_3)/6$$

$$v_1 \wedge v_2 = (v_1 \otimes v_2 - v_2 \otimes v_1)/2 \; . \qquad (7.8.3)$$

The product $(v_1 \wedge v_2) \wedge v_3 = (1/2)(v_1 \otimes v_2 - v_2 \otimes v_1) \wedge v_3$ is the wedge product of an antisymmetric rank-2 tensor and a vector and up to this point we have no idea how to evaluate such an creature.

Now is the time, then, to add a new axiom to the wedge product theory. We declare that,

Fact: The wedge product of k *vectors* $v_1 \wedge v_2 \wedge \,.....\wedge v_k$ can be "associated" in any manner without altering the meaning of the product. By this we mean that parentheses can be added in any manner without altering the object. (7.8.4)

What this in effect does is *define* an array of new objects to be the same as $v_1 \wedge v_2 \wedge \,.....\wedge v_6$. For example,

$$\begin{aligned}
&(v_1 \wedge v_2) \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 && \equiv v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \\
&v_1 \wedge (v_2 \wedge v_3) \wedge v_4 \wedge v_5 \wedge v_6 && \equiv v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \\
&v_1 \wedge (v_2 \wedge v_3 \wedge v_4) \wedge v_5 \wedge v_6 && \equiv v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \\
&v_1 \wedge (v_2 \wedge v_3 \wedge v_4) \wedge (v_5 \wedge v_6) && \equiv v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \qquad \text{// multiple ()} \\
&(v_1 \wedge v_2 \wedge v_3) \wedge (v_4 \wedge v_5 \wedge v_6) && \equiv v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \\
&(v_1 \wedge v_2) \wedge (v_3 \wedge v_4) \wedge (v_5 \wedge v_6) && \equiv v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \; .
\end{aligned} \qquad (7.8.5)$$

Given these definitions, it follows that *nested* parenthesis are also allowed. For example,

$$v_1 \wedge (v_2 \wedge (v_3 \wedge v_4) \wedge v_5) \wedge v_6 = v_1 \wedge (v_2 \wedge v_3 \wedge v_4 \wedge v_5) \wedge v_6 = v_1 \wedge v_2 \wedge v_3 \wedge v_4 \wedge v_5 \wedge v_6 \; . \qquad (7.8.6)$$

Since tensors like $T_\wedge$ can be expanded on $(e_{j_1} \wedge e_{j_2} \wedge .... \wedge e_{j_k})$, and since one may associate this wedge product arbitrarily as claimed in (7.8.4), one easily shows that :

Fact: The wedge product of N general *tensors* $A_\wedge \wedge B_\wedge \wedge C_\wedge ....$ can be "associated" in any manner without altering the meaning of the product. By this we mean that parentheses can be added in any manner without altering the object. (7.8.7)

This fact then extends the claim (7.8.4) made for N vectors, and is exactly analogous to the similar axiomatic statement for ⊗ associativity made in (2.8.21).

Example: In (7.7.1) multiindex notation, consider three L(V) tensors $T\hat{}$, $S\hat{}$, $R\hat{}$ of rank k,k',k" :

$$(T\hat{} \wedge S\hat{}) \wedge R\hat{} = (\ (\Sigma_I T^I u\hat{}_I) \wedge (\Sigma_J S^J u\hat{}_J)\ ) \wedge (\Sigma_K R^K u\hat{}_K)$$

$$= \Sigma_I T^I \Sigma_J S^J\ \{ (u\hat{}_I \wedge u\hat{}_J) \wedge (\Sigma_K R^K u\hat{}_K) \} \qquad \text{// rules (7.2.3)}$$

$$= \Sigma_I T^I \Sigma_J S^J \Sigma_K R^K\ (u\hat{}_I \wedge u\hat{}_J) \wedge (u\hat{}_K) \qquad \text{// rules (7.2.3) again}$$

$$= \Sigma_{IJK} T^I S^J R^K\ (u\hat{}_I \wedge u\hat{}_J \wedge u\hat{}_K) \qquad \text{// detail shown below}$$

$$= (\Sigma_I T^I u\hat{}_I) \wedge (\Sigma_J S^J u\hat{}_J) \wedge (\Sigma_K R^K u\hat{}_K) \qquad \text{// rules (7.2.3) again}$$

$$= T\hat{} \wedge S\hat{} \wedge R\hat{}\ .$$

Our example shows that for arbitrary L(V) tensors, $(T\hat{} \wedge S\hat{}) \wedge R\hat{} = T\hat{} \wedge S\hat{} \wedge R\hat{}$.

Here we illuminate the key detail above:

$$\begin{aligned}
(u_I \wedge u_J) \wedge u_K &= (\ (u_{i_1} \wedge u_{i_1} \wedge ... \wedge u_{i_k}) \wedge (u_{j_1} \wedge u_{j_1} \wedge ... \wedge u_{j_{k'}}\ )\ ) \wedge (u_{k_1} \wedge u_{k_1} \wedge ... \wedge u_{k_{k''}}\ ) \\
&= (\ u_{i_1} \wedge u_{i_1} \wedge ... \wedge u_{i_k} \wedge u_{j_1} \wedge u_{j_1} \wedge ... \wedge u_{j_{k'}}\ ) \wedge (u_{k_1} \wedge u_{k_1} \wedge ... \wedge u_{k_{k''}}\ ) \\
&= u_{i_1} \wedge u_{i_1} \wedge ... \wedge u_{i_k} \wedge u_{j_1} \wedge u_{j_1} \wedge ... \wedge u_{j_{k'}} \wedge u_{k_1} \wedge u_{k_1} \wedge ... \wedge u_{k_{k''}} \\
&= (u_{i_1} \wedge u_{i_1} \wedge ... \wedge u_{i_k}) \wedge (u_{j_1} \wedge u_{j_1} \wedge ... \wedge u_{j_{k'}}) \wedge (u_{k_1} \wedge u_{k_1} \wedge ... \wedge u_{k_{k''}}) \\
&= (u_I \wedge u_J \wedge u_K)\ .
\end{aligned}$$

In each step above the rule (7.8.4) for vectors (applied to basis vectors) is used.

Having faced up to the issue of associativity, we now resume the discussion of L(V).

Fact: This large space L(V) is in fact itself a vector space. (7.8.8)

We know this is true since $L(V) = \Sigma^{\oplus}{}_{k=0}{}^{\infty} L^k$ and we showed in (7.3.2) that each L^k is a vector space. For example, the "0" element in L(V) is the direct sum of the "0" elements of all the L^k. See Appendix B for more detail.

To show that L(V) is an algebra, we must show that it is closed under both addition and multiplication. It should be clear to the reader that L(V) is closed under addition and has the right scalar rule. For example, if k_1 and s are scalars,

$$k_1 \oplus a \oplus b \wedge c \oplus f \wedge g \wedge h\ = \text{sum of 4 elements of L(V)} = \text{an element of L(V)}$$

$$s(k_1 \oplus a \oplus b \wedge c \oplus f \wedge g \wedge h) = (sk_1) \oplus (sb) \wedge c \oplus f \wedge (sg) \wedge h\ = \text{element of L(V)} \qquad (7.8.9)$$

This additive closure is of course necessary for L(V) be a vector space.

The space is also closed under the multiplication operation ^. For example

$$(b\wedge c)\wedge(f\wedge g\wedge h) = b\wedge c\wedge f\wedge g\wedge h \ = \in \ L^5 = \in \ L(V) \ . \qquad // \ (b\wedge c) \in L^2 \quad (f\wedge g\wedge h) \in L^3 \qquad (7.8.10)$$

Here we have used the associative property (7.8.4). This closure claim is stated more generally below (7.9.a.6).

One then makes the following definitions with regard to the space L(V), where n = dim(V) :

Object	Name	any blade lincomb:	Grade(rank):	Space	
s	0-blade	scalar $\in$ K	0	L^0	
a	1-blade	vector	1	L^1	
a^b	2-blade	bivector	2	L^2	
a^b^c	3-blade	trivector	3	L^3	
a^b^c^d	4-blade	quadvector	4	L^4	
.....					
a^b^c^d^....	k-blade	k-vector	k	L^k	
....					
a^b^c^d^....	n-blade	n-vector	n	L^n	
arbitrary element of L(V)		multivector	mixed	L(V)	(7.8.11)

Since L(V) is closed under the operations ⊕ and ^, it is "an algebra" (the space L^k alone is not an algebra because it is not closed under ^). The L(V) algebra is different from that of the reals due to its definition as a sum of vector spaces. The elements of L(V) have different "grades" as shown above, so L(V) is a "graded algebra". Sometimes L(V) is called "the exterior tensor algebra" over V.

A **k-blade** is a pure wedge product of k vectors, whereas a **k-vector** is any *linear combination* of k-blades. A **multivector** is any linear combination of k-vectors for any mixed values of k.

Note that

$$s_1(a\wedge b) \ \oplus s_2(c\wedge d) \ = (s_1 a)\wedge b \ \oplus \ (s_2 c)\wedge d \ = (a'\wedge b) \ \oplus (c'\wedge d) \qquad // \text{ 2-blades}$$
$$s_1(a\wedge b) \ \oplus s_2(c\wedge d\wedge e) \ = (s_1 a)\wedge b \ \oplus \ (s_2 c)\wedge d\wedge e \ = (a'\wedge b) \ \oplus (c'\wedge d\wedge e) \qquad // \text{ multivector}$$

so it is also correct to say that a k-vector is any *sum* of k-blades, and a multivector is any *sum* of k-vectors. That is, any linear combination can be written as a sum as shown in the above examples.

Unlike in Tensor World, in Wedge World the above list (7.8.11) is finite for a given n = dim(V). For k = n there is exactly one linearly independent basis vector which is the ordered wedge product of all the basis vectors of V. For k > n, all wedge products vanish since the vectors in the wedge product are linearly dependent, see (7.2.6). The dimensionality of the space L(V) is as follows, based on (7.8.1) and (B.10)',

$$\dim[L(V)] = \dim[L^0 \ \oplus L^1 \oplus L^2 \oplus L^3 \ + ...] \ = \dim(L^0) + \dim(L^1) + \dim(L^2) + \dim(L^3) + ...$$

but for dim(V) = n this series truncates with L^n and we find from (7.3.6),

$$\dim[L(V)] = 1 + n + \binom{n}{2} + \binom{n}{3} + ... + \binom{n}{n} = \Sigma_{k=0}{}^n \binom{n}{k} = 2^n = \text{a finite number} \qquad (7.8.12)$$

7.9 The Wedge Product of two or more tensors in L(V)

(a) Wedge Product of two tensors $T_\wedge$ and $S_\wedge$

Here we shall mimic the developmental approach used in Section 5.6 for the tensor product. As before, we quietly "break in" the multiindex notation.

The symmetric expansions (7.4.4) of $T^\wedge$ and $S^\wedge$ are given by,

$$T^\wedge = \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k}) \qquad \text{rank } k, \quad T^\wedge \in L^k \qquad (7.9.a.1)$$
$$\Sigma_I T^I u^\wedge{}_I$$

$$S^\wedge = \Sigma_{j_1 j_2 \ldots j_{k'}} S^{j_1 j_2 \ldots j_{k'}} (u_{j_1} \wedge u_{j_2} \ldots \wedge u_{j_{k'}}) \qquad \text{rank } k', \quad S^\wedge \in L^{k'} . \qquad (7.9.a.2)$$
$$\Sigma_J S^J u^\wedge{}_J$$

We form the wedge product of these two tensors in a manner similar to (5.6.3) :

$$T^\wedge \wedge S^\wedge = [\Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k})] \wedge [\Sigma_{j_1 j_2 \ldots j_{k'}} S^{j_1 j_2 \ldots j_{k'}} (u_{j_1} \wedge u_{j_2} \ldots \wedge u_{j_{k'}})]$$
$$[\Sigma_I T^I u^\wedge{}_I] \qquad \wedge \qquad [\Sigma_J T^J u^\wedge{}_J]$$

(a) $$= \Sigma_{i_1 i_2 \ldots i_k} \Sigma_{j_1 j_2 \ldots j_{k'}} T^{i_1 i_2 \ldots i_k} S^{j_1 j_2 \ldots j_{k'}} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k}) \wedge (u_{j_1} \wedge u_{j_2} \ldots \wedge u_{j_{k'}})$$
$$\Sigma_{I,J} T^I S^J (u^\wedge{}_I) \wedge (u^\wedge{}_J)$$

(b) $$= \Sigma_{i_1 i_2 \ldots i_k j_1 j_2 \ldots j_{k'}} T^{i_1 i_2 \ldots i_k} S^{j_1 j_2 \ldots j_{k'}} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k} \wedge u_{j_1} \wedge u_{j_2} \ldots \wedge u_{j_{k'}})$$
$$\Sigma_{I,J} T^I S^J (u^\wedge{}_I \wedge u^\wedge{}_J)$$

(c) $$= \Sigma_{i_1 i_2 \ldots i_k i_{k+1} i_{k+2} \ldots i_{k+k'}} [T^{i_1 i_2 \ldots i_k} S^{i_{k+1} i_{k+2} \ldots i_{k+k'}}] (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_{k+k'}})$$
$$\Sigma_{I,I'} T^I S^{I'} (u^\wedge{}_I \wedge u^\wedge{}_{I'})$$

(d) $$= \Sigma_{i_1 i_2 \ldots i_k i_{k+1} i_{k+2} \ldots i_{k+k'}} [T \otimes S]^{i_1 i_2 \ldots i_k \, i_{k+1} i_{k+2} \ldots i_{k+k'}} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_{k+k'}})$$
$$\Sigma_{I,I'} [T \otimes S]^{I,I'} (u^\wedge{}_I \wedge u^\wedge{}_{I'})$$

(e) $$= \Sigma_{i_1 i_2 \ldots i_{k+k'}} [T \otimes S]^{i_1 i_2 \ldots i_{k+k'}} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_{k+k'}}) . \qquad (7.9.a.3)$$
$$\Sigma_{\boldsymbol{I}} (T \otimes S)^{\boldsymbol{I}} \, u^\wedge{}_{\boldsymbol{I}}$$

Comparing lines one sees that

$$I \equiv i_1, i_2...i_k \qquad I' \equiv i_{k+1}, i_{k+2}, ...i_{k+k'} \qquad \mathcal{I} \equiv I, I' = i_1,i_2...i_{k+k'}$$
$$u_{\wedge I} \equiv (u_{i_1} \wedge u_{i_2} .... \wedge u_{i_k}) \qquad u_{\wedge I'} \equiv (u_{i_{k+1}} \wedge ... \wedge u_{i_{k+k'}}) \qquad u_{\wedge \mathcal{I}} \equiv (u_{i_1} \wedge u_{i_2} .... \wedge u_{i_{k+k'}}) \qquad (7.9.a.4)$$

Notice that the (7.8.4) vector associativity of ^ is used going from (a) to (b).

The conclusion is that

$$T_\wedge \wedge S_\wedge = \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ u_{\wedge \mathcal{I}} \qquad \mathcal{I} \equiv I, I' = i_1,i_2...i_{k+k'}, \ \ u_{\wedge \mathcal{I}} \equiv (u_{i_1} \wedge u_{i_2} .... \wedge u_{i_{k+k'}}) . \qquad (7.9.a.5)$$

Since the $u_{\wedge \mathcal{I}}$ are basis vectors in $L^{k+k'}$, we have shown that:

$$T_\wedge \in L^k \text{ and } S_\wedge \in L^{k'} \quad \Rightarrow \quad T_\wedge \wedge S_\wedge \ \in\ L^{k+k'} \ \subset\ L(V) \ . \qquad (7.9.a.6)$$

Thus we have strengthened the claim made in (7.8.10) that L(V) is closed under the operation ^.

Recall now from (7.3.8) the relationship between $u_{\wedge \mathcal{I}}$ and $u_{\mathcal{I}}$,

$$(u_{i_1} \wedge u_{i_2} \wedge .... \wedge u_{i_k}) = \text{Alt}(u_{i_1} \otimes u_{i_2} \otimes .... \otimes u_{i_k})$$
$$u_{\wedge I} = \text{Alt}(u_I) \qquad \textit{(7.3.8)}$$

and (5.6.5) for the expansion of the tensor product T⊗S,

$$T\otimes S = \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ u_{\mathcal{I}} \qquad \mathcal{I} \equiv I, I' = i_1,i_2...i_{k+k'}, \ \ u_{\mathcal{I}} \equiv (u_{i_1} \otimes u_{i_2} .... \otimes u_{i_{k+k'}}) . \qquad \textit{(5.6.5)}$$

Applying Alt to this last equation (with component indices $\mathcal{J}$),

$$[\text{Alt}(T\otimes S)]^{\mathcal{J}} = \text{Alt}_{\mathcal{J}}\ [(T\otimes S)^{\mathcal{J}}] \qquad // \text{(A.5.3c)}$$

$$= \text{Alt}_{\mathcal{J}}\ [\ \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}} (u_{\mathcal{I}})^{\mathcal{J}}\ ] \qquad // \text{ component } \mathcal{J} \text{ of (5.6.5) quoted just above}$$

$$= \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ \text{Alt}_{\mathcal{J}}\ [(u_{\mathcal{I}})^{\mathcal{J}}] \qquad // \text{ (A.5.10) that Alt is linear}$$

$$= \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ \text{Alt}_{\mathcal{I}}\ [(u_{\mathcal{I}})^{\mathcal{J}}] \qquad // \text{ (A.8.31), } (u_{\mathcal{I}})^{\mathcal{J}} \text{ has factored form } (u_{i_1})^{j1}(u_{i_2})^{j2} ...$$

$$= \Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ (u_{\wedge \mathcal{I}})^{\mathcal{J}} \qquad // \text{ (7.3.8) quoted just above}$$

$$= [\Sigma_{\mathcal{I}} (T\otimes S)^{\mathcal{I}}\ (u_{\wedge \mathcal{I}})]^{\mathcal{J}}$$

$$= (T_\wedge \wedge S_\wedge)^{\mathcal{J}} \qquad // \text{ (7.9.a.5)}$$

so we end up with the following elegant and compact way to write the wedge product of two tensors,

$$T_\wedge \wedge S_\wedge = \text{Alt}(T\otimes S) . \qquad // \text{ see Sec (g) below for this result in Spivak normalization} \qquad (7.9.a.7)$$

Concealing the Alt_I and Alt_J details one can get the correct result with this sequence,

$$\text{Alt}(T\otimes S) = \text{Alt}(\Sigma_I (T\otimes S)^I u_I) = \Sigma_I (T\otimes S)^I \text{Alt}(u_I) = \Sigma_I (T\otimes S)^I (u^\wedge{}_I) = T^\wedge \wedge S^\wedge .$$

The components of (7.9.a.7) are,

$$[T^\wedge \wedge S^\wedge]^J = [\text{Alt}(T\otimes S)]^J$$

$$= \frac{1}{(k+k')!} \Sigma_P(-1)^{S(P)} (T\otimes S)^{P(J)} \qquad // \text{ (A.5.3a)}$$

$$= \frac{1}{(k+k')!} \Sigma_P(-1)^{S(P)} T^{P(J)} S^{P(J')} . \qquad // \text{ see e.g. (5.6.15)} \qquad (7.9.a.8)$$

This last line is an explicit instruction for computing the components of the tensor $T^\wedge \wedge S^\wedge$. We have added this new notation,

$$T^{P(I)} \equiv T^{i_{P(1)} i_{P(2)} \ldots i_{P(k)}} \qquad \text{for } I = i_1, i_2 \ldots i_k \qquad (7.9.a.9)$$

Example: Let S and T both be rank-2 tensors so $k = k' = 2$. Then

$$[T^\wedge \wedge S^\wedge]^I = [T^\wedge \wedge S^\wedge]^{i_1 i_2 i_3 i_4} = (1/4!) \Sigma_P(-1)^{S(P)} T^{i_{P(1)} i_{P(2)}} S^{i_{P(3)} i_{P(4)}}$$

$$= (1/24) [T^{i_1 i_2} S^{i_3 i_4} - T^{i_2 i_1} S^{i_3 i_4} + T^{i_2 i_3} S^{i_1 i_4} - T^{i_2 i_3} S^{i_4 i_1} + 20 \text{ more terms }] . \qquad (7.9.a.10)$$

Here as elsewhere we show in red the indices to be swapped to make the next term. From (7.9.c.6) below,

$$T^\wedge \wedge S^\wedge = (-1)^{2*2} S^\wedge \wedge T^\wedge = S^\wedge \wedge T^\wedge . \qquad (7.9.a.11)$$

(b) Special cases of the wedge product $T^\wedge \wedge S^\wedge$

Assume $T^\wedge$ and $S^\wedge$ have rank k and k'.
If $S = \kappa' \in K$ = a scalar, then rank(S) = k' = 0 and (7.9.a.3) (b) reads,

$$T^\wedge \wedge S^\wedge = \Sigma_{i_1 i_2 \ldots i_k j_1 j_2 \ldots j_{k'}} T^{i_1 i_2 \ldots i_k} S^{j_1 j_2 \ldots j_{k'}} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k} \wedge u_{j_1} \wedge u_{j_2} \ldots \wedge u_{j_{k'}})$$

$$\rightarrow \Sigma_{i_1 i_2 \ldots i_k} T^{i_1 i_2 \ldots i_k} (\kappa') (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k}) = \kappa' T \qquad (7.9.b.1)$$

and

$$S^\wedge \wedge T^\wedge = \Sigma_{j_1 j_2 \ldots j_{k'} i_1 i_2 \ldots i_k} S^{j_1 j_2 \ldots j_{k'}} T^{i_1 i_2 \ldots i_k} (u_{j_1} \wedge u_{j_2} \ldots \wedge u_{j_{k'}} \wedge u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k})$$

$$\rightarrow \Sigma_{i_1 i_2 \ldots i_k} (\kappa') T^{i_1 i_2 \ldots i_k} (u_{i_1} \wedge u_{i_2} \ldots \wedge u_{i_k}) = \kappa' T \qquad (7.9.b.2)$$

so we find that $T\hat{}S = S\hat{}T = \kappa' T$.

If $T = \kappa$ and $S = \kappa'$, the result above would be $T_\wedge \hat{} S_\wedge = \kappa\kappa'$ and $S_\wedge \hat{} T_\wedge = \kappa'\kappa$ and so $T_\wedge \hat{} S_\wedge = S_\wedge \hat{} T_\wedge = \kappa\kappa'$. Thus,

$$T_\wedge \hat{} S_\wedge = \kappa \hat{} S_\wedge = S_\wedge \hat{} T_\wedge = S_\wedge \hat{} \kappa = \kappa S_\wedge \qquad \text{if } T_\wedge = \kappa \in V^0$$

$$T_\wedge \hat{} S_\wedge = T_\wedge \hat{} \kappa' = S_\wedge \hat{} T_\wedge = \kappa' \hat{} T_\wedge = \kappa' T_\wedge \qquad \text{if } S_\wedge = \kappa' \in V^0$$

$$T_\wedge \hat{} S_\wedge = \kappa \hat{} \kappa' = S_\wedge \hat{} T_\wedge = \kappa' \hat{} \kappa = \kappa\kappa' \qquad \text{if } T_\wedge, S_\wedge = \kappa, \kappa' \in V^0 \qquad (7.9.b.3)$$

These special case results are the same as those for $T \otimes S$ shown in (5.6.16). When T is rank-0 or rank-1 we can write $T_\wedge = T$ according to (7.4.19), but we continue to use $T_\wedge$.

(c) Commutivity Rule for the Wedge Product of two tensors $T_\wedge$ and $S_\wedge$

Recall the expansion of $T_\wedge \hat{} S_\wedge$ from (7.9.a.3) item (b),

$$T_\wedge \hat{} S_\wedge = \Sigma_{i_1 i_2 \dots i_k j_1 j_2 \dots j_{k'}} T^{i_1 i_2 \dots i_k} S^{j_1 j_2 \dots j_{k'}} (u_{i_1} \hat{} u_{i_2} \dots \hat{} u_{i_k} \hat{} u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}})$$
$$\Sigma_{I,J} T^I S^J \,(u_{\wedge I} \hat{} u_{\wedge J}) \qquad (7.9.c.1)$$

Swapping $T \leftrightarrow S$, $k \leftrightarrow k'$ and $i \leftrightarrow j$ gives the following form for the wedge product $S_\wedge \hat{} T_\wedge$,

$$S_\wedge \hat{} T_\wedge = \Sigma_{j_1 j_2 \dots j_{k'} i_1 i_2 \dots i_k} S^{j_1 j_2 \dots j_{k'}} T^{i_1 i_2 \dots i_k} (u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}} \hat{} u_{i_1} \hat{} u_{i_2} \dots \hat{} u_{i_k})$$
$$\Sigma_{J,I} S^J T^I \,(u_{\wedge J} \hat{} u_{\wedge I})$$

$$= \Sigma_{i_1 i_2 \dots i_k j_1 j_2 \dots j_{k'}} T^{i_1 i_2 \dots i_k} S^{j_1 j_2 \dots j_{k'}} (u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}} \hat{} u_{i_1} \hat{} u_{i_2} \dots \hat{} u_{i_k})$$
$$\Sigma_{I,J} T^I S^J \,(u_{\wedge J} \hat{} u_{\wedge I}) \ . \qquad (7.9.c.2)$$

Equations (7.9.c.1) and (7.9.c.2) are identical *except for* the last factor involving the basis vectors. Consider the basis vector factor appearing in (7.9.c.2),

$$(u_{\wedge J} \hat{} u_{\wedge I}) = (u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}} \hat{} u_{i_1} \hat{} u_{i_2} \dots \hat{} u_{i_k}) \ . \qquad (7.9.c.3)$$

To make this match the basis factor in (7.9.c.1), we have to slide all the red basis vectors to the left through all the black basis vectors. Each time a red passes through a black, we pick up a minus sign due to the rule (7.2.4). Thus,

$$(u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}} \hat{} u_{i_1} \hat{} u_{i_2} \dots \hat{} u_{i_k}) = (-1)^{k'} u_{i_1} \hat{} (u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}} \hat{} u_{i_2} \dots \hat{} u_{i_k})$$

$$= (-1)^{k'} (-1)^{k'} u_{i_1} \hat{} u_{i_2} \hat{} (u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}} \dots \hat{} u_{i_k}) = \text{etc.} =$$

$$= [(-1)^{k'}]^k \,(u_{i_1} \hat{} u_{i_2} \dots \hat{} u_{i_k} \hat{} u_{j_1} \hat{} u_{j_2} \dots \hat{} u_{j_{k'}}) \ . \qquad (7.9.c.4)$$

Therefore,

$$(u^{\wedge}_J \wedge u^{\wedge}_I) = (-1)^{kk'}\ (u^{\wedge}_I \wedge u^{\wedge}_J)\ . \tag{7.9.c.5}$$

Inserting this result into (7.9.c.2) gives

$$S^{\wedge} \wedge T^{\wedge} = (-1)^{kk'} T^{\wedge} \wedge S^{\wedge} \qquad \text{ranks of the two tensors are k and k'}\ . \tag{7.9.c.6}$$

Since the commutivity sign is a function of the ranks (grades) of the tensors, this statement is sometimes referred to as "graded commutivity". The wedge product of two tensors commutes if kk' is even, and anticommutes if kk' is odd.

Using (7.9.a.7) the above becomes.

$$\text{Alt}(S\otimes T) = (-1)^{kk'} \text{Alt}(T\otimes S)\ . \tag{7.9.c.7}$$

Example: If k = k' = 1, $(-1)^{kk'}$ = -1 and we recover the simple rule for vectors,

$$S^{\wedge} \wedge T^{\wedge} = - T^{\wedge} \wedge S^{\wedge} \qquad \text{// S and T are rank-1 tensors (vectors)} \tag{7.9.c.8}$$

as first stated in (4.3.2). One must keep in mind that the result $S^{\wedge} \wedge T^{\wedge} = - T^{\wedge} \wedge S^{\wedge}$ is not valid for arbitrary tensors $S^{\wedge}$ and $T^{\wedge}$.

Examples:
If k = 0 so T = κ, rule (7.9.c.6) says $S^{\wedge} \wedge T^{\wedge} = T^{\wedge} \wedge S^{\wedge}$, consistent with (7.9.b.3) line 1.
If k=k'=0 so T = κ and S = κ', rule (7.9.c.6) again says $S^{\wedge} \wedge T^{\wedge} = T^{\wedge} \wedge S^{\wedge}$, consistent with (7.9.b.3) line 3.
(7.9.c.9)

(d) Wedge Product of three or more tensors

Mimicking (5.6.7) we write

$$T^{\wedge} \wedge S^{\wedge} \wedge R^{\wedge} = [\Sigma_I T^I u^{\wedge}_I] \wedge [\Sigma_J\ S^J u^{\wedge}_J] \wedge [\Sigma_K\ R^K u^{\wedge}_K]$$

(a) $\quad = \Sigma_{I,J,K}\ T^I S^J R^K\ (u^{\wedge}_I) \wedge (u^{\wedge}_J) \wedge (u^{\wedge}_K)$

(b) $\quad = \Sigma_{I,J,K}\ T^I S^J R^K\ (u^{\wedge}_I \wedge u^{\wedge}_J \wedge u^{\wedge}_K)$ // associative of ^ used here

(d) $\quad = \Sigma_{I,I',I''}\ T^I S^{I'} R^{I''}\ (u^{\wedge}_I \wedge u^{\wedge}_{I'} \wedge u^{\wedge}_{I''})$ // rename multiindices J→I',K→I"

$I \equiv i_1, i_2 ... i_k$	$I' \equiv i_{k+1}, i_{k+2}, ... i_{k+k'}$	$I'' \equiv i_{k+k'+1}, i_{k+k'+2}, ... i_{k+k'+k''}$
$u^{\wedge}_I \equiv (u_{i_1} \wedge .... \wedge u_{i_k})$	$u^{\wedge}_{I'} \equiv (u_{i_{k+1}} \wedge ... \wedge u_{i_{k+k'}})$	$u^{\wedge}_{I''} \equiv (u_{i_{k+k'+1}} \wedge ... \wedge u_{i_{k+k'+k''}})$

(e) $\quad = \Sigma_{\mathcal{I}}\ (T\otimes S\otimes R)^{\mathcal{I}}\ u^{\wedge}_{\mathcal{I}} \qquad u^{\wedge}_{\mathcal{I}} \equiv (u_{i_1} \wedge .... \wedge u_{i_{k+k'+k''}}) \qquad \mathcal{I} \equiv I, I', I'' = i_1, i_2 ... i_{k+k'+k''}$ (7.9.d.1)

The outer product form is $T^I S^{I'} R^{I''} = (T\otimes S\otimes R)^{I,I',I''} = (T\otimes S\otimes R)^{\mathcal{I}}$.

The conclusion is then,

$$T\hat{}\wedge S\hat{}\wedge R\hat{} = \Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}} u\hat{}_{\mathcal{I}} \qquad \mathcal{I} \equiv I, I',I'' = i_1,i_2...i_{k+k'+k''} , \ u\hat{}_{\mathcal{I}} \equiv (u_{i_1}\wedge....\wedge u_{i_{k+k'+k''}}) \qquad (7.9.d.2)$$

Since the $u\hat{}_{\mathcal{I}}$ are basis vectors in $L^{k+k'+k''}$, we have shown that:

$$T\hat{} \in L^k \text{ and } S\hat{} \in L^{k'} \text{ and } R\hat{} \in L^{k''} \qquad \Rightarrow \qquad T\hat{}\wedge S\hat{}\wedge R\hat{} \in L^{k+k'+k''} \subset L(V) . \qquad (7.9.d.3)$$

We now mimic the sequence of steps above (7.9.a.7) :

$[\mathrm{Alt}(T\otimes S\otimes R)]^{\mathcal{J}} = \mathrm{Alt}_{\mathcal{J}} [(T\otimes S\otimes R)^{\mathcal{J}}]$ // (A.5.3b)

$= \mathrm{Alt}_{\mathcal{J}} [\Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}}(u_{\mathcal{I}})^{\mathcal{J}}]$ // component J of (5.6.8) $T\otimes S\otimes R = \Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}} u_{\mathcal{I}}$

$= \Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}} \mathrm{Alt}_{\mathcal{J}} [(u_{\mathcal{I}})^{\mathcal{J}}]$ // (A.5.10) that Alt is linear

$= \Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}} \mathrm{Alt}_{\mathcal{I}} [(u_{\mathcal{I}})^{\mathcal{J}}]$ // (A.8.31) since $(u_{\mathcal{I}})^{\mathcal{J}}$ has factored form

$= \Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}} (u\hat{}_{\mathcal{I}})^{\mathcal{J}}$ // (7.3.8)

$= [\Sigma_{\mathcal{I}} (T\otimes S\otimes R)^{\mathcal{I}} (u\hat{}_{\mathcal{I}})]^{\mathcal{J}}$

$= (T\hat{}\wedge S\hat{}\wedge R\hat{})^{\mathcal{J}}$ // (7.9.d.2)

so

$$T\hat{}\wedge S\hat{}\wedge R\hat{} = \mathrm{Alt}(T\otimes S\otimes R) \qquad (7.9.d.4)$$

and then

$[T\hat{}\wedge S\hat{}\wedge R\hat{}]^{\mathcal{I}} = [\mathrm{Alt}(T\otimes S\otimes R)]^{\mathcal{I}}$

$= \frac{1}{(k+k'+k'')!} \Sigma_P (-1)^{S(P)} (T\otimes S\otimes R)^{P(\mathcal{I})}$ // (A.5.3)

$$= \frac{1}{(k+k'+k'')!} \Sigma_P (-1)^{S(P)} T^{P(I)} S^{P(I')} R^{P(I'')} \qquad (7.9.d.5)$$

which gives instructions for how to compute the components of $T\hat{}\wedge S\hat{}\wedge R\hat{}$.

Using the systematic notation outlined in (5.6.10) through (5.6.12), and generalizing the above development for the wedge product of three tensors, we find the following expansion for the wedge product of N tensors of L(V),

$$(T_1)_\wedge\hat{}(T_2)_\wedge\hat{}...\hat{}(T_N)_\wedge \;=\; \Sigma_{\boldsymbol{I}}\ (T_1{}^{I_1}T_2{}^{I_2}\ ....\ T_N{}^{I_N})\, u_{\wedge\boldsymbol{I}} \;=\Sigma_{\boldsymbol{I}}\,(T_1{\otimes}T_2....{\otimes}T_N)^{\boldsymbol{I}}\ u_{\wedge\boldsymbol{I}}$$

where $\qquad u_{\wedge\boldsymbol{I}} = u_{i_1}\hat{}\ u_{i_2}\ .....\hat{}\ u_{i_{k_1+k_2+...+k_N}} \;=\; u_{i_1}\hat{}\ u_{i_2}\ .....\hat{}\ u_{i_\kappa}, \qquad \kappa = \Sigma_{i=1}{}^N k_i$

and $\qquad (T_1{\otimes}T_2....{\otimes}T_N)^{\boldsymbol{I}} = T_1{}^{I_1}T_2{}^{I_2}\ ....\ T_N{}^{I_N}\quad .$ (7.9.d.6)

The rank of this product tensor is then $\kappa = \Sigma_{i=1}{}^N k_i$ and the tensor is an element of $L^\kappa \subset L(V)$. Notice that if $\kappa > n$, the tensor product (7.9.d.6) *vanishes* since there are then > n factors in $u_{\wedge I}$ so one or more are then duplicated,

$$(T_1)_\wedge\hat{}(T_2)_\wedge\hat{}...\hat{}(T_N)_\wedge = 0 \qquad\qquad \text{if } \kappa = \Sigma_{i=1}{}^N k_i \ge n{+}1\ . \tag{7.9.d.7}$$

For example, if all the tensors are the same tensor $T_\wedge$ of rank k, then

$$T_\wedge{}^N \equiv\ T_\wedge\hat{}T_\wedge\hat{}...\hat{}T_\wedge = 0 \qquad\qquad \text{if } Nk\ \ge n{+}1 \quad\text{or}\quad N \ge (n{+}1)/k\ . \tag{7.9.d.8}$$

If $N \ge (n+1)$, then $N \ge (n+1)/k$ for any $k \ge 1$. Thus

$$T_\wedge{}^N\ = 0 \qquad\qquad \text{for any } N \ge n{+}1 \quad \text{assuming } k \ne 0. \tag{7.9.d.9}$$

Recall (5.6.13),

$$T_1{\otimes}T_2{\otimes}...{\otimes}T_N \;=\; \Sigma_{\boldsymbol{I}}\ (T_1{}^{I_1}T_2{}^{I_2}\ ....\ T_N{}^{I_N})\, u_{\boldsymbol{I}} \;=\Sigma_{\boldsymbol{I}}\,(T_1{\otimes}T_2....{\otimes}T_N)^{\boldsymbol{I}}\ u_{\boldsymbol{I}} \quad . \qquad \textit{(5.6.13)}$$

Repeating the sequence above (7.9.d.4) for a longer product, we find that

$$(T_1)_\wedge\hat{}(T_2)_\wedge\hat{}...\hat{}(T_N)_\wedge \;= \mathrm{Alt}(T_1{\otimes}T_2{\otimes}...{\otimes}T_N)\ . \tag{7.9.d.10}$$

Components of this tensor are computed as follows:

$$[(T_1)_\wedge\hat{}(T_2)_\wedge\hat{}...\hat{}(T_N)_\wedge]^{\boldsymbol{I}} \;= [\mathrm{Alt}(T_1{\otimes}T_2{\otimes}...{\otimes}T_N)]^{\boldsymbol{I}}$$

$$= \frac{1}{\kappa!}\,\Sigma_P(-1)^{S(P)}\ (T_1{\otimes}T_2{\otimes}...{\otimes}T_N)^{P(\boldsymbol{I})} \qquad\qquad // \text{(A.5.3)},\ \kappa = \Sigma_{i=1}{}^N k_i$$

$$= \frac{1}{\kappa!}\,\Sigma_P(-1)^{S(P)}\ T_1{}^{P(I_1)}T_2{}^{P(I_2)}\ ....T_N{}^{P(I_N)} \tag{7.9.d.11}$$

where $\qquad T_1{}^{P(I_1)} \;\equiv\; T_1{}^{i_{P(1)}i_{P(2)}...i_{P(k)}} \qquad\qquad$ for $I_1 = i_1, i_2...i_{\kappa_1}$

$\qquad\qquad T_2{}^{P(I_2)} \;\equiv\; T_2{}^{i_{P(\kappa_1+1)}i_{P(\kappa_1+2)}...i_{P(\kappa_2)}} \qquad$ for $I_2 = \{i_{\kappa_1+1}, i_{\kappa_1+2}.....i_{\kappa_2}\}$

etc. // see (5.6.10 thru 12) for details

In the Dirac notation of Section 2.11 one can write (7.9.d.10) as

$$|(T_1)^\wedge\rangle \wedge |(T_2)^\wedge\rangle \wedge \ ... \wedge |(T_N)^\wedge\rangle \ = \ \text{Alt}\,(\,|T_1\rangle \otimes |T_2\rangle \otimes \ ... \otimes |T_N\rangle\,)\ . \qquad (7.9.d.12)$$

It is shown in (C.4.14) that "pre-antisymmetrization makes no difference", so the above may also be written

$$|(T_1)^\wedge\rangle \wedge |(T_2)^\wedge\rangle \wedge \ ... \wedge |(T_N)^\wedge\rangle \ = \ \text{Alt}\,(\,|(T_1)^\wedge\rangle \otimes |(T_2)^\wedge\rangle \otimes \ ... \otimes |(T_N)^\wedge\rangle\,)\ . \qquad (7.9.d.13)$$

Both sides of this equation are elements of the wedge product space $L^{k_1+k_2+..+k_N}$, but they are also both elements of the larger tensor product space $V^{k_1} \otimes V^{k_2} \otimes ... \otimes V^{k_N}$. The action of linear operator $\mathcal{P}$ on a tensor product space vector is defined in the obvious manner, as in (5.6.17),

$$\mathcal{P}\,[\,|(T_1)^\wedge\rangle \otimes |(T_2)^\wedge\rangle \otimes \ ... \otimes |(T_2)^\wedge\rangle\,] \ = \mathcal{P}\,|(T_1)^\wedge\rangle \otimes \mathcal{P}\,|(T_2)^\wedge\rangle \otimes \ ... \otimes \mathcal{P}\,|(T_2)^\wedge\rangle\ . \qquad (7.9.d.14)$$

In other words, the action of $\mathcal{P}$ on the larger space is defined in terms of its action on the spaces which make up the tensor product. This result holds as well for the wedge product of N tensors,

$$\mathcal{P}\,[\,|(T_1)^\wedge\rangle \wedge |(T_2)^\wedge\rangle \wedge \ ... \wedge |(T_2)^\wedge\rangle\,] \ = \mathcal{P}\,|(T_1)^\wedge\rangle \wedge \mathcal{P}\,|(T_2)^\wedge\rangle \wedge \ ... \wedge \mathcal{P}\,|(T_2)^\wedge\rangle \qquad (7.9.d.15)$$

Proof: $\mathcal{P}\,[\,|(T_1)^\wedge\rangle \wedge |(T_2)^\wedge\rangle \wedge \ ... \wedge |(T_2)^\wedge\rangle\,] \ = \ \mathcal{P}\,[\text{Alt}\,(\,|(T_1)^\wedge\rangle \otimes |(T_2)^\wedge\rangle \otimes \ ... \otimes |(T_2)^\wedge\rangle\,)\,]$

$= \text{Alt}\,[\mathcal{P}\,(\,|(T_1)^\wedge\rangle \otimes |(T_2)^\wedge\rangle \otimes \ ... \otimes |(T_2)^\wedge\rangle\,)\,]$

$= \text{Alt}\,[\,(\,\mathcal{P}\,|(T_1)^\wedge\rangle \otimes \mathcal{P}\,|(T_2)^\wedge\rangle \otimes \ ... \otimes \mathcal{P}\,|(T_2)^\wedge\rangle\,)\,]$

$= \mathcal{P}\,|(T_1)^\wedge\rangle \wedge \mathcal{P}\,|(T_2)^\wedge\rangle \wedge \ ... \wedge \mathcal{P}\,|(T_2)^\wedge\rangle$.

(e) Commutativity Rule for product of N tensors

Consider an example where we have a wedge product of 9 tensors. The $u^\wedge_I$ basis function groups are

$$u^\wedge_{I_1} \wedge u^\wedge_{I_2} \wedge u^\wedge_{I_3} \wedge u^\wedge_{I_4} \wedge u^\wedge_{I_5} \wedge u^\wedge_{I_6} \wedge u^\wedge_{I_7} \wedge u^\wedge_{I_8} \wedge u^\wedge_{I_9} \qquad (7.9.e.1)$$

which goes with

$$(T_1)^\wedge \wedge (T_2)^\wedge \wedge (T_3)^\wedge \wedge (T_4)^\wedge \wedge (T_5)^\wedge \wedge (T_6)^\wedge \wedge (T_7)^\wedge \wedge (T_8)^\wedge \wedge (T_9)^\wedge\ . \qquad (7.9.e.2)$$

The sign caused by swapping $(T_3)^\wedge \leftrightarrow (T_7)^\wedge$ will be the same as the sign swapping $u^\wedge_{I_3} \leftrightarrow u^\wedge_{I_7}$ in the basis function. We do it one step at a time, first sliding the group $u^\wedge_{I_7}$ to the left using (7.9.c.5),

$$u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_3 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_7 \wedge u^\wedge I_8 \wedge u^\wedge I_9$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_3 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_7 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{k_6k_7}$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_3 \wedge u^\wedge I_4 \wedge u^\wedge I_7 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{(k_6+k_5)k_7}$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_3 \wedge u^\wedge I_7 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{(k_6+k_5+k_4)k_7}$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_7 \wedge u^\wedge I_3 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{(k_6+k_5+k_4+k_3)k_7} . \tag{7.9.e.3}$$

Now with this as a starting point, we slide $u^\wedge I_3$ to the right, one group at a time,

$$u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_7 \wedge u^\wedge I_3 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_7 \wedge u^\wedge I_4 \wedge u^\wedge I_3 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{k_3k_4}$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_7 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_3 \wedge u^\wedge I_6 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{k_3(k_4+k_5)}$$

$$= u^\wedge I_1 \wedge u^\wedge I_2 \wedge u^\wedge I_7 \wedge u^\wedge I_4 \wedge u^\wedge I_5 \wedge u^\wedge I_6 \wedge u^\wedge I_3 \wedge u^\wedge I_8 \wedge u^\wedge I_9 \quad (-1)^{k_3(k_4+k_5+k_6)} \tag{7.9.e.4}$$

and now we have successfully swapped $u^\wedge I_3 \leftrightarrow u^\wedge I_7$ so also $(T_3)^\wedge \leftrightarrow (T_7)^\wedge$. The total sign is

$$\text{sign} = (-1)^m \qquad \text{where } m = (k_6+ k_5+ k_4+ k_3)k_7 + (k_4+k_5+k_6)k_3$$

$$= (k_4+k_5+k_6)(k_3+k_7)+ k_3k_7 . \tag{7.9.e.5}$$

Based on this result, we claim that :

Fact: In a product of tensors $(T_1)^\wedge \wedge (T_2)^\wedge \wedge (T_3)^\wedge$.... of rank k_1, k_2, k_3 ... , if two tensors are swapped $(T_r)^\wedge \leftrightarrow (T_s)^\wedge$ (with $r < s$), the resulting tensor incurs the following sign relative to the starting tensor,

$$\text{sign} = (-1)^m \qquad \text{where } m = (k_{r+1}+k_{r+2} ...+k_{s-1})(k_r+k_s) + k_rk_s . \tag{7.9.e.6}$$

Corollary: If the sum of the ranks of the two swapped tensor is even, in effect $m = k_rk_s$. (7.9.e.7)

<u>Example 1:</u>

$$(T_1)^\wedge \wedge (T_2)^\wedge \wedge (T_3)^\wedge = (-1)^m (T_2)^\wedge \wedge (T_1)^\wedge \wedge (T_3)^\wedge \qquad r = 1 \quad s = 2$$

$$m = (0)(k_1+k_2) + k_1k_2 = k_1k_2 \qquad (-1)^m = (-1)^{k_1k_2} \tag{7.9.e.8}$$

which is consistent with (7.9.c.6) saying $T_1 \wedge T_2 = (-1)^{k_1k_2} T_2 \wedge T_1$.

Example 2:

$(T_1)_\wedge \wedge (T_2)_\wedge \wedge (T_3)_\wedge = (-1)^m (T_3)_\wedge \wedge (T_2)_\wedge \wedge (T_1)_\wedge$ $\qquad r = 1 \quad s = 3$

$$m = (k_2)(k_1+k_3) + k_1k_3 = k_1k_2 + k_1k_3 + k_2k_3 \qquad (-1)^m = (-1)^{k_1k_2+k_1k_3+k_2k_3} \tag{7.9.e.9}$$

This result can be obtained as well by direct pairwise swapping using (7.9.c.6) and the associativity of ^,

$$(T_1)_\wedge \wedge (T_2)_\wedge \wedge (T_3)_\wedge = (-1)^{k_1k_2} (T_2)_\wedge \wedge (T_1)_\wedge \wedge (T_3)_\wedge = (-1)^{k_1k_2} (-1)^{k_1k_3} (T_2)_\wedge \wedge (T_3)_\wedge \wedge (T_1)_\wedge$$

$$= (-1)^{k_1k_2} (-1)^{k_1k_3} (-1)^{k_2k_3} (T_3)_\wedge \wedge (T_2)_\wedge \wedge (T_1)_\wedge \tag{7.9.e.10}$$

Example 3: Suppose all the tensors are vectors with rank = 1. Then the sum of the ranks of any two tensors is 2, which is even, so the Corollary above says $m = k_rk_s = 1*1 = 1$, so swapping any two of these tensors produces a minus sign,

$$\text{phase} = (-1)^m = -1 \qquad \text{where } m \approx k_rk_s = 1*1 = 1$$

in agreement with the basic vector swap rule (7.2.4). (7.9.e.11)

(f) Theorems from Appendix C : pre-antisymmetrization makes no difference

We showed above that one can form wedge products of elements of L(V) in this manner (in Spivak normalization, the right sides of these equations incur factorials as shown in Section (g) below),

$$T_\wedge \wedge S_\wedge = \text{Alt}(T\otimes S) \; . \tag{7.9.a.7}$$

$$T_\wedge \wedge S_\wedge \wedge R_\wedge = \text{Alt}(T\otimes S\otimes R) \tag{7.9.d.4}$$

$$(T_1)_\wedge \wedge (T_2)_\wedge \wedge ... \wedge (T_N)_\wedge = \text{Alt}(T_1\otimes T_2\otimes ...\otimes T_N) \tag{7.9.d.7}$$

where the operator Alt acts on the tensor indices which are not displayed in the above compact notation.

For example

$$T_\wedge \wedge S_\wedge = \text{Alt}(T\otimes S)$$

means, in multiindex notation,

$$(T_\wedge \wedge S_\wedge)^I = \text{Alt}_I [(T\otimes S)^I] = \text{Alt}_I [\, T^I S^{I'}] = \frac{1}{(k+k')!} \Sigma_P (-1)^{S(P)}\, T^{P(I)} S^{P(I')} \; .$$

A very simple case is the following (recall for vectors that $a = a_\wedge$)

$$(a \wedge b)^{i_1 i_2} = \text{Alt}\,[(a\otimes b)^{i_1 i_2}] = \text{Alt}\,[a^{i_1}b^{i_2}] = \frac{1}{(1+1)!}\,\Sigma_P(-1)^{S(P)}\, a^{i_{P(1)}}\, b^{i_{P(2)}}$$

$$= (1/2)\,[\,a^{i_1}b^{i_2} - a^{i_2}b^{i_1}] = (1/2)\,[(a\otimes b)^{i_1 i_2} - (b\otimes a)^{i_1 i_2}\,]$$

$$= \{\,(1/2)\,[(a\otimes b) - (b\otimes a)\,]\}^{i_1 i_2}$$

which replicates our Chapter 4 statement that

$$a \wedge b = [a\otimes b - b\otimes a]/2\ . \qquad\qquad (4.3.1)$$

Appendix C uses the rearrangement theorem in three separate Theorems to show that

$$T_\wedge \wedge S_\wedge = \text{Alt}(T\otimes S) = \text{Alt}(T_\wedge\otimes S) = \text{Alt}(T\otimes S_\wedge) = \text{Alt}(T_\wedge\otimes S_\wedge)\ . \qquad (7.9.f.1)$$

Theorem One   Theorem Two   Theorem Three

Recall that

$$T_\wedge \equiv \text{Alt}(T) \qquad\qquad (7.4.3)$$

so that $T_\wedge$ is a totally antisymmetric tensor. What (7.9.f.1) says is that Alt(T⊗S) provides total antisymmetrization on all tensor indices, so pre-antisymmetrizing either or both tensors makes no difference. A similar statement applies to working with totally symmetric tensors. So we have,

$$\text{Alt}[T\otimes S] = \text{Alt}[T_\wedge\otimes S] = \text{Alt}[T\otimes S_\wedge] = \text{Alt}[T_\wedge\otimes S_\wedge]$$
$$\text{where} \quad T_\wedge = \text{Alt}(T) \quad S_\wedge = \text{Alt}(S) \qquad\qquad (C.4.1)$$

$$\text{Sym}[T\otimes S] = \text{Sym}[T_s\otimes S] = \text{Sym}[T\otimes S_s] = \text{Sym}[T_s\otimes S_s]$$
$$\text{where} \quad T_s = \text{Sym}(T) \quad S_s = \text{Sym}(S)\ . \qquad\qquad (C.4.2)$$

These can of course be rewritten as

$$\text{Alt}[T\otimes S] = \text{Alt}[\text{Alt}(T)\otimes S] = \text{Alt}[T\otimes \text{Alt}(S)] = \text{Alt}[\text{Alt}(T)\otimes \text{Alt}(S)] \qquad (C.4.3)$$

$$\text{Sym}[T\otimes S] = \text{Sym}[\text{Sym}(T)\otimes S] = \text{Sym}[T\otimes \text{Sym}(S)] = \text{Sym}[\text{Sym}(T)\otimes \text{Sym}(S)]\ . \qquad (C.4.4)$$

Similarly Appendix C shows that

$$T_\wedge\wedge S_\wedge\wedge R_\wedge = \text{Alt}(T\otimes S\otimes R) = \text{Alt}(T_\wedge\otimes S\otimes R) = \text{Alt}(T\otimes S_\wedge\otimes R) = \text{Alt}(T\otimes S\otimes R_\wedge)$$
$$= \text{Alt}(T_\wedge\otimes S_\wedge\otimes R) = \text{Alt}(T_\wedge\otimes S\otimes R_\wedge) = \text{Alt}(T\otimes S_\wedge\otimes R_\wedge)$$
$$= \text{Alt}(T_\wedge\otimes S_\wedge\otimes R_\wedge)\ . \qquad\qquad (7.9.f.2)$$

Adding ^ subscripts inside an Alt expression changes nothing. Here is another example:

$$T_\wedge\wedge S_\wedge\wedge R_\wedge = \text{Alt}(T\otimes S\otimes R) = \text{Alt}((T\otimes S)\otimes R) = \text{Alt}((T\otimes S)_\wedge\otimes R) = \text{Alt}(\text{Alt}(T\otimes S)\otimes R)\ . \qquad (7.9.f.3)$$

(g) Spivak Normalization

Spivak's definition of the Alt operator is the same as ours and the same as Benn & Tucker's, but the latter authors write the Alt operator in an elaborate script font as $\mathcal{ALT}$. Our wedge product normalization is the same as Benn & Tucker's but differs from that of Spivak, a difference which we now explore.

Suppose we were to omit the (1/k!) normalization factor in the definition of the wedge product of k vectors, so that (7.1.2) would become

$$v_{j_1} \wedge v_{j_2} \wedge \,.....\wedge v_{j_k} = 1\, \Sigma_P (-1)^{S(P)} (v_{j_{P(1)}} \otimes v_{j_{P(2)}} \otimes \,.....\otimes v_{j_{P(k)}})$$

$$= 1\, [\, (v_{j_1} \otimes v_{j_2} \otimes \,.....\otimes v_{j_k}) \quad + \quad \text{all signed permutations}\,]$$

$$= k!\, \text{Alt}(v_{j_1} \otimes v_{j_2} \otimes \,.....\otimes v_{j_k}) \,. \qquad (7.1.2)_S$$

A convenient way to understand this change is that everything stays the same but Spivak's wedge products are "bigger than" ours.

In particular,

$$a \wedge b = 1[\, a\otimes b - b\otimes a] \,. \qquad \text{// no factor of 1/2} \qquad (4.3.1)_S$$

We show all factors that are *different* from our normalization in red. Earlier equations converted to Spivak normalization are shown below with a subscript S added to the earlier equation number. For example, equation (7.2.8) becomes

$$(v_{j_1} \wedge v_{j_2} \wedge \,.....\wedge v_{j_k})^{i_1 i_2 ... i_k} = 1\, \det[\, (v_{j_*})^{i_*}] \qquad (7.2.8)_S$$

and correspondingly

$$(u_{j_1} \wedge u_{j_2} \wedge \,.....\wedge u_{j_k})^{i_1 i_2 ... i_k} = 1\, \det[\, \delta_{j_*}{}^{i_*}] = 1\, \det(\delta_J{}^I) \,. \qquad (7.3.9)_S$$

Our L^k basis vectors of (7.3.8) become

$$(u_{i_1} \wedge u_{i_2} \wedge \,.....\wedge u_{i_k}) = k!\, \text{Alt}(u_{i_1} \otimes u_{i_2} \otimes \,.....\otimes u_{i_k})$$

or

$$u_{\wedge I} = k!\text{Alt}(u_I) \,. \qquad (7.3.8)_S$$

where recall that the Alt operator (A.5.3) always contains an internal factor (1/k!) which is required so AltT = T if the tensor T is already totally antisymmetric.

The tensor expansion for $T^\wedge \in L^k$ is still given by (7.4.4),

$$T^\wedge = \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k} (u_{i_1} \wedge u_{i_2} \wedge ..... \wedge u_{i_k}) \quad . \qquad (7.4.4)_s$$

and of course the corresponding *tensor* expansion of $T \in V^k$ is also unaltered,

$$T = \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k} (u_{i_1} \otimes u_{i_2} ..... \otimes u_{i_k}) \quad . \qquad (7.4.1)$$

The reader is thus reminded of the difference between tensors $T^\wedge$ and T in our notation.

Then the new (7.4.2) is,

$$\Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k} (u_{i_1} \wedge u_{i_2} \wedge ..... \wedge u_{i_k})^{j_1j_2...j_k}$$

$$= \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k}\, k!\, \mathrm{Alt}_I[(u_{i_1} \otimes u_{i_2} \otimes ..... \otimes u_{i_k})^{j_1j_2...j_k}] \qquad // (7.3.8)$$

$$= \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k}\, k!\, \mathrm{Alt}_I[\, \delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k}] \qquad // (5.1.4)$$

$$= k!\, \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k}\, \mathrm{Alt}_J[\delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k}] \qquad // (A.8.30)$$

$$= k!\, \mathrm{Alt}_J[\, \Sigma_{i_1i_2....i_k} T^{i_1i_2....i_k} \quad \delta_{i_1}{}^{j_1} \delta_{i_2}{}^{j_2} .... \delta_{i_k}{}^{j_k} \qquad // (A.5.10) \text{ Alt is linear}$$

$$= k!\, \mathrm{Alt}_J(T^{j_1j_2...j_k}) \;= \mathrm{Alt}(T^{j_1j_2...j_k}) \qquad // \text{ no ambiguity}$$

$$= \;1\; \Sigma_P (-1)^{S(P)}\, T^{j_{P(1)}j_{P(2)}...j_{P(k)}} \qquad // \text{ def of Alt } (A.5.3b)$$

$$= \;k!\, [\mathrm{Alt}(T)]^{j_1j_2...j_k} \qquad // (A.5.3c)$$

$$\equiv \;[T^\wedge]^{j_1j_2...j_k} \qquad (7.4.2)_s$$

with the result (rank T = k, rank S = k')

$$T^\wedge \equiv k!\mathrm{Alt}(T) \qquad \text{and} \qquad S^\wedge \equiv k'!\mathrm{Alt}(S) \; . \qquad (7.4.3)_s$$

The wedge product development of Section 7.9 (a) goes through with no change to give the result

$$T^\wedge \wedge S^\wedge = \Sigma_{\mathcal{I}} (T \otimes S)^{\mathcal{I}}\; u^\wedge_{\mathcal{I}} \qquad \mathcal{I} \equiv I, I' = i_1,i_2...i_{k+k'}, \;\; u^\wedge_{\mathcal{I}} \equiv (u_{i_1} \wedge u_{i_2} .... \wedge u_{i_{k+k'}}) \; . \qquad (7.9.a.5)_s$$

But then we find

$$\mathrm{Alt}(T\otimes S)^{\boldsymbol{J}} = \mathrm{Alt}_{\boldsymbol{J}}(T\otimes S)^{\boldsymbol{J}} = \Sigma_{\boldsymbol{I}}\, (T\otimes S)^{\boldsymbol{I}}\mathrm{Alt}_{\boldsymbol{J}}(u_{\boldsymbol{I}})^{\boldsymbol{J}} \qquad \text{// (5.6.5) and (A.5.10) that Alt is linear}$$

$$= \Sigma_{\boldsymbol{I}}\, (T\otimes S)^{\boldsymbol{I}}\mathrm{Alt}_{\boldsymbol{I}}(u_{\boldsymbol{I}})^{\boldsymbol{J}} \qquad \text{// use (A.8.27) since } (u_{\boldsymbol{I}})^{\boldsymbol{J}} \text{ is totally antisymmetric in } I \text{ and } J$$

$$= \Sigma_{\boldsymbol{I}}\, (T\otimes S)^{\boldsymbol{I}} \frac{1}{(k+k')!} (u_{\wedge\boldsymbol{I}})^{\boldsymbol{J}} \qquad \text{// } (7.3.8)_{\mathbf{S}} \text{ above with } k \rightarrow k+k'$$

$$= \frac{1}{(k+k')!} (T_\wedge \wedge S_\wedge)^{\boldsymbol{J}} \qquad \text{// } (7.9.a.5)_{\mathbf{S}} \text{ above}$$

so

$$T_\wedge \wedge S_\wedge = (k+k')!\ \mathrm{Alt}(T\otimes S)\ . \qquad (7.9.a.7)_{\mathbf{S}}$$

The fact that "pre-antisymmetrizing makes no difference" is unaltered, so we still have

$$\mathrm{Alt}(T\otimes S) = \mathrm{Alt}(\mathrm{Alt}(T)\otimes\mathrm{Alt}(S))\ . \qquad \textit{(C.4.3)}$$

Then using $T_\wedge \equiv k!\mathrm{Alt}(T)$ and $S_\wedge \equiv k'!\mathrm{Alt}(T)$ we end up with

$$T_\wedge \wedge S_\wedge = (k+k')!\ \mathrm{Alt}(T\otimes S) = (k+k')!\ \mathrm{Alt}(\mathrm{Alt}(T)\otimes\mathrm{Alt}(S))$$

$$= \frac{(k+k')!}{k!\ k'!}\ \mathrm{Alt}(T_\wedge\otimes S_\wedge) \quad T_\wedge \in L^k \ \text{ and } \ S_\wedge \in L^{k'}\ . \qquad (7.9.g.1)$$

By the same analysis, we would find for a triple product in the Spivak normalization,

$$T_\wedge \wedge S_\wedge \wedge R_\wedge = (k+k'+k")!\ \mathrm{Alt}(T\otimes S\otimes R)$$

$$= \frac{(k+k'+k")!}{k!\ k'!k"!}\ \mathrm{Alt}(T_\wedge\otimes S_\wedge\otimes R_\wedge) \quad T_\wedge \in L^k,\ S_\wedge \in L^{k'},\ R_\wedge \in L^{k"}\ . \qquad (7.9.g.2)$$

Spivak uses lower-case Greek letters for elements of L^k, so the above two equations appear as

$$\omega \wedge \eta = \frac{(k+l)!}{k!\ l!}\mathrm{Alt}(\omega \otimes \eta).$$

Spivak page 79

$$(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta)$$
$$= \frac{(k+l+m)!}{k!\,l!\,m!} \mathit{Alt}(\omega \otimes \eta \otimes \theta).$$

Spivak page 80

Actually, Spivak never talks about rank-k tensors and L^k, only rank-k *tensor functions* which he calls "k-tensors" and which we will associate with the dual space Λ^k in Chapter 8, and that is what the Greek objects are in the above. But if he did talk about rank-k tensors and L^k, the above in red would be his normalization. We shall of course reprise this topic in Chapter 8.

One advantage of the Spivak normalization is that vector wedge products don't have the annoying 1/k! so that, for example, there is no overall 1/3! in the following,

$$v_1 \wedge v_2 \wedge v_3 = v_1\otimes v_2\otimes v_3 - v_1\otimes v_3\otimes v_2 + v_3\otimes v_1\otimes v_2 - v_3\otimes v_2\otimes v_1 + v_2\otimes v_3\otimes v_1 - v_2\otimes v_1\otimes v_3 \qquad (7.1.5)_S$$

The disadvantage, which seems a large one to us, is all the extra factorials in the wedge products of multiple tensors, and the fact that $T_\wedge = k!\mathrm{Alt}(T)$ instead of the simpler $T_\wedge = \mathrm{Alt}(T)$.

8. The Wedge Product of k dual vectors : the vector spaces Λ^k and $\Lambda(V)$

Comment: This Chapter 8 is a *partial* copy, paste and edit version of Chapter 7 -- a translation from non-dual to dual. See our similar comment at the start of Chapter 6. Since Chapter 7 is so long, here in Chapter 8 we shall delete all material that is basically unchanged from the non-dual Chapter 7. We also delete most "comments" and examples. Just as in going from Chapter 5 to Chapter 6, the notion of tensor components is replaced by the notion of tensor functions. The equation numbers for Chapter 8 match those of Chapter 7, and deletions thus cause "holes" in the sequence for Chapter 8.

8.1 Definition of the wedge product of k dual vectors

We wish to define the wedge product of k dual vectors $\alpha_i \in V^*$,

$$\alpha_1 \wedge \alpha_2 \wedge \ .....\wedge \alpha_k \quad . \qquad\qquad // \ \langle\alpha_1| \wedge \langle\alpha_2| \wedge .... \wedge \langle\alpha_k|$$

Wedge products of this form (and their linear combinations) inhabit a vector space we call $\Lambda^k(V)$ or Λ^k.

We now *impose* the requirement that this wedge product must change sign when any two vectors are swapped. This property is injected into the wedge product theory, it does not fall out from it.

This sign-change requirement leads to the following candidate definition for the wedge product of k vectors in V (the j_r are vector labels),

$$\alpha_{j_1} \wedge \alpha_{j_2} \wedge \ .....\wedge \alpha_{j_k} = (1/k!)\ \Sigma_P (-1)^{S(P)} \left(\alpha_{P(j_1)} \otimes \alpha_{P(j_2)} \otimes \ .....\otimes \alpha_{P(j_k)} \right)$$

$$= (1/k!)\ [\ (\alpha_{j_1} \otimes \alpha_{j_2} \otimes \ .....\otimes \alpha_{j_k}) \quad + \quad \text{all signed permutations}\ ]$$

$$= \mathrm{Alt}(\alpha_{j_1} \otimes \alpha_{j_2} \otimes \ .....\otimes \alpha_{j_k}) \ . \qquad (8.1.2)$$

For the purposes of this section, we simplify things by taking $j_r \rightarrow r$ so (8.1.2) becomes,

$$\alpha_1 \wedge \alpha_2 \wedge \ .....\wedge \alpha_k = (1/k!)\ \Sigma_P (-1)^{S(P)} \left(\alpha_{P(1)} \otimes \alpha_{P(2)} \otimes \ .....\otimes \alpha_{P(k)} \right)$$

$$= (1/k!)\ [\ (\alpha_1 \otimes \alpha_2 \otimes \ .....\otimes \alpha_k) \quad + \quad \text{all signed permutations}\ ] \ .$$

$$= \mathrm{Alt}(\alpha_1 \otimes \alpha_2 \otimes \ .....\otimes \alpha_k)$$

$$= (1/k!)\ \Sigma_{i_1 i_2 \ldots i_k}\, \varepsilon_{i_1 i_2 \ldots i_k} (\alpha_{i_1} \otimes \alpha_{i_2} \otimes \ ..... \ \otimes \alpha_{i_k}) \qquad i_r = 1 \text{ to } k \qquad (8.1.3)$$

where (1/k!) is a normalization factor. In Spivak normalization, the (1/k!) factors above are all replaced by 1, see discussion in Section 8.9 (g) below.

8.2 Properties of the wedge product of k dual vectors

Where there is no comment on an item, see the the parallel item in Chapter 7.

1. The sums in (8.1.2) and (8.1.3) have k! terms. (8.2.1)

2. The wedge product is k-multilinear. (8.2.2)

It is by-fiat axiom that the wedge product of k vectors is k-multilinear and therefore satisfies these rules,

$$\alpha_1\wedge(\alpha_2 + \alpha'_2)\wedge\alpha_3\wedge.....\wedge\alpha_k = \alpha_1\wedge\alpha_2\wedge\alpha_3\wedge .....\wedge\alpha_k + \alpha_1\wedge\alpha'_2\wedge\alpha_3\wedge .....\wedge\alpha_k$$
$$\alpha_1\wedge(s\alpha_2)\wedge\alpha_3\wedge .....\wedge \alpha_k = s(\alpha_1\wedge\alpha_2\wedge\alpha_3\wedge .....\wedge\alpha_k) \qquad s,r= \text{scalar} \in K$$

or

$$\alpha_1\wedge(r\alpha_2 + s\alpha'_2)\wedge\alpha_3\wedge.....\wedge\alpha_k = r(\alpha_1\wedge\alpha_2\wedge\alpha_3\wedge .....\wedge\alpha_k) + s(\alpha_1\wedge\alpha'_2\wedge\alpha_3\wedge .....\wedge\alpha_k) \ . \qquad (8.2.3)$$

Here we show the rules just for the 2 position, but k-multilinear means these rules must apply to all the vector positions. These rules cannot be derived from the similar tensor product rules (6.3.1).

3. The wedge product changes sign if any vector pair is swapped. (8.2.4)

4. Wedge product of vectors vanishes if any two vectors are the same.

Given a sign change (8.2.4) for any pair swap of vectors in the wedge product, we know that

$$\alpha_1\wedge \alpha_2\wedge .....\wedge \alpha_k = 0 \quad \text{if any two (or more) vectors are the same.} \qquad (8.2.5)$$

5. Wedge product vanishes if vectors are linearly dependent. (8.2.6)

6. Wedge product vanishes if $k > n$. (8.2.7)

7. Components. For the dual space, we consider tensor functions in place of tensor components, so

$$(\alpha_{j_1}\wedge \alpha_{j_2}\wedge .....\wedge \alpha_{j_k})(v_{i_1},v_{i_2}.....v_{i_k}) \qquad \alpha_{\wedge J}(v_I)$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)} (\alpha_{j_{P(1)}}\otimes \alpha_{j_{P(2)}}\otimes .....\otimes \alpha_{j_{P(k)}})(v_{i_1},v_{i_2}.....v_{i_k}) \qquad // (8.1.2)$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)} (\alpha_{j_{P(1)}})(v_{i_1}) (\alpha_{j_{P(2)}})(v_{i_2}) ... (\alpha_{j_{P(k)}})(v_{i_k}) \qquad // (6.6.17) \text{ for vectors}$$

$$= (1/k!) \det [\alpha_{j*}(v_{i*})] \ . \qquad // (A.1.17) \qquad (1/k!) \det [\alpha_J(v_I)] \qquad (8.2.8a)$$

Evaluating at the basis vectors then gives,

$$(\alpha_{j_1}\wedge \alpha_{j_2}\wedge .....\wedge \alpha_{j_k})(u_{i_1},u_{i_2}.....u_{i_k}) = (1/k!) \det [\alpha_{j*}(u_{i*})] \qquad \alpha_{\wedge J} (u_I)$$

$$= (1/k!) \det [(\alpha_{j*})_{i*}] \ . \qquad // \text{ see } (2.11.c.9) \qquad (1/k!) \det [(\alpha_J)_I] \qquad (8.2.8b)$$

In the last equation, the α_r are rank-1 functionals. We know that each such functional is associated with a unique vector $\boldsymbol{\alpha}_r$ in V which appears in (2.11.a.4), $\alpha_r(\mathbf{v}) = \langle\boldsymbol{\alpha}_r \mid \mathbf{v}\rangle = \boldsymbol{\alpha}_r \bullet \mathbf{v}$. Thus, we can form a tensor in V^k called $(\boldsymbol{\alpha}_{j_1}\wedge \boldsymbol{\alpha}_{j_2}\wedge \; .....\wedge \boldsymbol{\alpha}_{j_k})$. For this rank-k tensor we have

$$(\boldsymbol{\alpha}_{j_1}\wedge \boldsymbol{\alpha}_{j_2}\wedge \; .....\wedge \boldsymbol{\alpha}_{j_k})_{i_1 i_2 .. i_k} = (\alpha_{j_1}\wedge \alpha_{j_2}\wedge \; .....\wedge \alpha_{j_k})(\mathbf{u}_{i_1},\mathbf{u}_{i_2}.....\mathbf{u}_{i_k}) \quad // (6.5.1)$$

$$= (1/k!) \det [\, (\boldsymbol{\alpha}_{j*})_{i*}] \; . \qquad // (8.2.8b) \qquad (8.2.8c)$$

In the Spivak normalization the factor (1/k!) in equations (8.2.8) is replaced by 1.

Fact: $(\alpha_{j_1}\wedge\alpha_{j_2}\wedge...\wedge\alpha_{j_k})(v_{i_1},v_{i_2}....v_{i_k})$ is totally antisymmetric in both the labels j_r and the labels i_r. (8.2.9)

Proof: Antisymmetry on the j_r follows from (8.2.4), while antisymmetry on i_r follows from (8.2.8a) (determinant changes sign if any two rows or columns are swapped).

8. Associative Property of the wedge product.

For example, $(\alpha_1\wedge \alpha_2)\wedge \alpha_3 = \alpha_1\wedge (\alpha_2\wedge \alpha_3) = \alpha_1\wedge \alpha_2\wedge a_3$.

8.3 The vector space Λ^k and its basis

Λ^k is the space whose elements are all linear combinations of wedge products of k vectors of V*. (8.3.1)

A more precise name for this space is $\Lambda^k(V)$ but we just call it Λ^k.

It seems useful at this point to compare our vector space names with those of Spivak:

Names of spaces. [TA = totally antisymmetric = alternating]

us	Spivak		
		tensor product spaces	
V^k	--	space of rank-k tensors, $T = \|T\rangle$, $T^{i_1 i_2 .... i_k}$	
V^{*k}	--	dual space of k-multilinear tensor functionals on V, $\mathcal{T} = \langle T\|$	
$V^{*k}{}_f$	$\mathcal{T}^k(V)$	space of k-multilinear tensor functions on V, $\mathcal{T}(\mathbf{v}) = \langle T\|\mathbf{v}\rangle$	
		wedge product spaces	
L^k	--	space of TA rank-k tensors, $T_\wedge$, $T_\wedge{}^{i_1 i_2 .... i_k}$	
Λ^k	--	dual space of TA k-multilinear tensor functionals on V, $\mathcal{T}_\wedge$	
$\Lambda^k{}_f$	$\Lambda^k(V)$	space of TA k-multilinear tensor functions on V, $\mathcal{T}_\wedge(\mathbf{v})$	(8.3.1a)

Sjamaar refers to the last space as $A^k V$ (2006) and $A^k(V)$ in his 2015 update. We wanted to end up with the name Λ^k for the last space to agree with Spivak, Benn & Tucker, Conrad and others, and this led to the non-Greek L^k for the corresponding non-dual wedge space.

We now go down the list of items in Section 7.3, adapting them as needed. Again, where there is no comment on an item below, please see the the parallel item in Chapter 7. Multiindex versions of some equations appear on the right below in red.

Λ^k is a vector space (8.3.2)

Basis elements for Λ^k

Consider the following objects in Λ^k obtained by wedging together k basis elements of V*, where each λ^i is selected from the set of n available for V* (which has dimension n),

$$(\lambda^{j_1} \wedge \lambda^{j_2} \wedge .... \wedge \lambda^{j_k}) \ . \qquad (8.3.3)$$

There are $\binom{n}{k}$ independent basis elements for Λ^k and they all have this form

$$(\lambda^{i_1} \wedge \lambda^{i_2} \wedge .... \wedge \lambda^{i_k}) \qquad \text{where} \quad i_1 < i_2 < ..... < i_k \qquad \binom{n}{k} \text{ basis elements} \qquad (8.3.6)$$

Fact: $(\lambda^{i_1} \wedge \lambda^{i_2} \wedge .... \wedge \lambda^{i_k}) = \text{Alt}(\lambda^{i_1} \otimes \lambda^{i_2} \otimes .... \otimes \lambda^{i_k})$ $\qquad \lambda_{\wedge}{}^{I} = \text{Alt}(\lambda^I)$ (8.3.8)

This is just a special case of (8.1.2).

Components of the basis elements for Λ^k .

Now reconsider the basis vectors of the vector space Λ^k ,

$$(\lambda^{j_1} \wedge \lambda^{j_2} \wedge .... \wedge \lambda^{j_k}) \ . \qquad \textit{(8.3.3)}$$

For this dual space Λ^k, we consider tensor functions in place of tensor components, so we then have these special cases of (8.2.8a) and (8.2.8b),

$$(\lambda^{j_1} \wedge \lambda^{j_2} \wedge .... \wedge \lambda^{j_k})(v_{i_1},v_{i_2}.....v_{i_k}) \qquad \lambda_{\wedge J}(v_I)$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)} (\lambda^{j_{P(1)}} \otimes \lambda^{j_{P(2)}} \otimes .....\otimes \lambda^{j_{P(k)}})(v_{i_1},v_{i_2}.....v_{i_k})$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)} (\lambda^{j_{P(1)}})(v_{i_1}) (\lambda^{j_{P(2)}})(v_{i_2}) ... (\lambda^{j_{P(k)}})(v_{i_k}) \qquad \text{// (6.6.17) for vectors}$$

$$= (1/k!)\, \Sigma_P (-1)^{S(P)} (v_{i_1})^{j_{P(1)}} (v_{i_2})^{j_{P(2)}} ... (v_{i_k})^{j_{P(k)}} \qquad \text{// (2.11.c.5)}$$

$$= (1/k!) \det [\, (v_{i_*})^{j_*} \,] \ . \qquad \text{// (A.1.19)} \qquad (1/k!) \det [\, (v_I)^J \,] \qquad (8.3.9a)$$

Evaluation at the basis vectors then gives,

$(\lambda^{j1} \wedge \lambda^{j2} \wedge .... \wedge \lambda^{jk})(u_{i_1},u_{i_2}.....u_{i_k}) \quad = \quad (1/k!) \det [(u_{i_*})^{j*}] \qquad \lambda_\wedge{}^J (v_I)$

$= (1/k!) \det [\delta_{i_*}{}^{j*}] \qquad$ // see (2.6.8) $(\mathbf{u}_m)^n = \delta_m{}^n \qquad (1/k!) \det [\delta_I{}^J] \qquad$ (8.3.9b)

Once again, in Spivak normalization $(1/k!) \rightarrow 1$ for equations (8.3.9). Looking at $\det [(v_{i_*})^{j*}]$ above we immediately conclude that,

Fact: $(\lambda^{j1} \wedge \lambda^{j2} \wedge .... \wedge \lambda^{jk})(v_{i_1},v_{i_2}....v_{i_k})$ is totally antisymmetric in both the labels j_r and the labels i_r. (8.3.10)

We saw an example of both antisymmetries for k = 2 back in equation (4.4.21),

$(\lambda^i \wedge \lambda^j)(v_r,v_s) = - (\lambda^i \wedge \lambda^j)(v_s,v_r) = - (\lambda^j \wedge \lambda^i)(v_r,v_s)$ // two forms of antisymmetry *(4.4.21)*

Equation (8.3.9b) can be expressed in our usual informal notation,

$(\lambda^{j1} \wedge \lambda^{j2} \wedge .... \wedge \lambda^{jk})(u_{i_1},u_{i_2}....u_{i_k}) = (1/k!) [\delta^{j1}{}_{i_1} \delta^{j2}{}_{i_2} ... \delta^{jk}{}_{i_k} + \text{signed permutations}]$ (8.3.11)

<u>Example</u>: [see (7.3.12)]

$3!(\lambda^{j1} \wedge \lambda^{j2} \wedge \lambda^{j3})(u_{i_1},u_{i_2},u_{i_3}) = 3! \lambda_\wedge{}^J(u_I) = \det(\delta^J{}_I) = \det (\delta^{j_1 j_2 j_3}{}_{i_1 i_2 i_3})$ (8.3.12)

8.4 Tensor Expansions for a dual tensor in Λ^k

Recall now the tensor expansion for a most-general tensor $\mathcal{T}$ in V^{*k} ,

$\mathcal{T} = \Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} (\lambda^{i1} \otimes \lambda^{i2} ..... \otimes \lambda^{ik}) \quad \mathcal{T} \in V^{*k}$ *(6.2.3)* (8.4.1)

where $T_{i_1 i_2 .... i_k}$ are some general coefficients.

Consider then the similar-looking most-general object in Λ^k, evaluated at $(v_{j_1},v_{j_2}....v_{j_k})$,

$\Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} (\lambda^{i1} \wedge \lambda^{i2} ..... \wedge \lambda^{ik})(v_{j_1},v_{j_2}....v_{j_k})$

$= \Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} \mathrm{Alt}_I(\lambda^{i1} \otimes \lambda^{i2} ..... \otimes \lambda^{ik})(v_{j_1},v_{j_2}....v_{j_k})$ // (8.3.8)

$= \Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} \mathrm{Alt}_J(\lambda^{i1} \otimes \lambda^{i2} ..... \otimes \lambda^{ik})(v_{j_1},v_{j_2}....v_{j_k})$ // (8.2.9) and (A.8.27)

$= \mathrm{Alt}_J[\Sigma_{i_1 i_2 .... i_k} T_{i_1 i_2 .... i_k} (\lambda^{i1} \otimes \lambda^{i2} ..... \otimes \lambda^{ik})(v_{j_1},v_{j_2}....v_{j_k})]$ // (A.5.10), Alt is linear

$= \mathrm{Alt}_J \mathcal{T}(v_{j_1},v_{j_2}....v_{j_k}) = [\mathrm{Alt}_J(\mathcal{T})](v_{j_1},v_{j_2}....v_{j_k}) = [\mathrm{Alt}(\mathcal{T})](v_{j_1},v_{j_2}....v_{j_k})$ // (8.4.1)

$\equiv \mathcal{T}_\wedge(v_{j_1},v_{j_2}....v_{j_k})$. (8.4.2)

Here we define this functional (dual-space) notation,

$$\mathcal{T}^\wedge \equiv \mathrm{Alt}(\mathcal{T}) \tag{8.4.3}$$

which is really this statement about tensor functions,

$$\mathcal{T}^\wedge(v_{j_1},v_{j_2}....v_{j_k}) \equiv [\mathrm{Alt}(\mathcal{T})]\,(v_{j_1},v_{j_2}....v_{j_k}) \,. \tag{8.4.3a}$$

From (8.4.2) we then have the following fully general element of Λ^k,

$$\mathcal{T}^\wedge = \Sigma_{i_1 i_2 ....i_k}\ T_{i_1 i_2 ....i_k} (\lambda^{i_1} \wedge \lambda^{i_2} .....\wedge \lambda^{i_k}) \ . \tag{8.4.4}$$

We refer to this type of expansion as a symmetric expansion, and we know it is redundant since the symmetric sum includes each true basis vector k! times.

According to (A.8.9), we know from (8.4.3a) and (8.2.2) that

Fact: $\mathcal{T}^\wedge(v_{i_1},v_{i_2}....v_{i_k})$ is a totally antisymmetric k-multilinear tensor function. (8.4.5)

Therefore,

Fact: The space Λ^k is the space of all *totally antisymmetric* k-multilinear rank-k tensors $\mathcal{T}^\wedge$. To say that $\mathcal{T}^\wedge$ is totally antisymmetric means that $\mathcal{T}^\wedge(v_{i_1},v_{i_2}....v_{i_k})$ is a totally antisymmetric tensor function. (8.4.6)

In contrast, the space V^{*k} is the space of *all* k-multilinear rank-k tensors $\mathcal{T}$, so $\Lambda^k \subset V^{*k}$.

Since the set $(\lambda^{i_1}\wedge \lambda^{i_2} .....\wedge \lambda^{i_k})$ with $1 \le i_1 < i_2 < ..... < i_k \le n$ forms a complete basis for Λ^k, as discussed above in (8.3.6), it must be possible to express $\mathcal{T}^\wedge$ in the following manner

$$\mathcal{T}^\wedge = \Sigma_{1\le i_1<i_2<....<i_k\le n}\ A_{i_1 i_2 ...i_k} (\lambda^{i_1} \wedge \lambda^{i_2} .....\wedge \lambda^{i_k}) \,. \tag{8.4.7}$$

What then is the connection between the $A^{i_1 i_2 ...i_k}$ of (8.4.7) and the $T^{i_1 i_2 ...i_k}$ of (8.4.4)?

The discussion goes exactly as in Chapter 7 and the result is,

$$A^{i_1 i_2 ...i_k} = \Sigma_P (-1)^{S(P)}\ T^{i_{P(1)} i_{P(2)} ...i_{P(k)}} \qquad\qquad i_1 < i_2 < ..... < i_k$$

$$= [\ T^{i_1 i_2 ...i_k} + \text{ all signed permutations }] \qquad \text{// k! terms}$$

$$= k!\ [\mathrm{Alt}(T)]^{i_1 i_2 ...i_k} \,. \qquad \text{// (A.5.3a) def of Alt}$$

or

$$A = k!\mathrm{Alt}(T) = k!\ T^\wedge \tag{8.4.16}$$

where $T_\wedge$ = Alt(T) as shown in (7.4.3), not to be confused with tensor functional $\mathcal{T}_\wedge \equiv$ Alt($\mathcal{T}$) in (8.4.3).

Since A = k! $T_\wedge$, (7.4.5) shows that

Fact: $A^{i_1 i_2 \ldots i_k}$ and $T_\wedge{}^{i_1 i_2 \ldots i_k}$ are both totally antisymmetric tensors. (8.4.17)

Vector Case. For k = 1, we find that

$$\mathcal{T} = \Sigma_{i_1} T_{i_1} \lambda^{i_1} \qquad // (6.2.3)$$

$$\mathcal{T}_\wedge = \Sigma_{i_1} T_{i_1} \lambda^{i_1} \qquad // (8.4.4) \qquad \Rightarrow \qquad \mathcal{T}_\wedge = \mathcal{T} \qquad (8.4.19)$$

so for a vector there is no distinction between $\mathcal{T}_\wedge$ and $\mathcal{T}$ (and in fact $V^{*1} = \Lambda^1$).

8.5 Various expansions for the wedge product of k dual vectors

We have generally stopped bolding vectors in V, but in this section we bold the vectors $\boldsymbol{\alpha}_r$ to distinguish them from the corresponding functionals α_r, where recall (2.11.a.4) that $\alpha_r(\mathbf{v}) = \boldsymbol{\alpha}_r \bullet \mathbf{v}$.

The *symmetric* expansion is very straightforward. First, consider this rank-k tensor in V^k ,

$$T_{i_1 i_2 \ldots i_k} = (\boldsymbol{\alpha}_1)_{i_1} (\boldsymbol{\alpha}_2)_{i_2} \ldots (\boldsymbol{\alpha}_k)_{i_k} = (\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes \ldots\ldots \otimes \boldsymbol{\alpha}_k)_{i_1 i_2 \ldots i_k} \qquad // \boldsymbol{\alpha}_r \in V$$

or (8.5.1a)

$$T = (\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes \ldots\ldots \otimes \boldsymbol{\alpha}_k) \qquad \text{or} \qquad |T\rangle = |\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_k\rangle . \qquad // T \in V^k$$

The corresponding rank-k tensor functional is,

$$\mathcal{T} = (\alpha_1 \otimes \alpha_2 \otimes \ldots\ldots \otimes \alpha_k) \qquad // \alpha_r \in V^*, \ \mathcal{T} \in V^{*k}$$

or (8.5.1b)

$$\langle T| = \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots \boldsymbol{\alpha}_k| \ .$$

Then the symmetric expansion (8.4.4) gives,

$$\mathcal{T}_\wedge = \Sigma_{i_1 i_2 \ldots i_k} T_{i_1 i_2 \ldots i_k} (\lambda^{i_1} \wedge \lambda^{i_2} \ldots\ldots \wedge \lambda^{i_k}) \qquad \textit{(8.4.4)}$$

$$= \Sigma_{i_1 i_2 \ldots i_k} (\boldsymbol{\alpha}_1)_{i_1} (\boldsymbol{\alpha}_2)_{i_2} \ldots (\boldsymbol{\alpha}_k)_{i_k} (\lambda^{i_1} \wedge \lambda^{i_2} \ldots\ldots \wedge \lambda^{i_k}) \qquad // (8.5.1a) \qquad (8.5.2)$$

$$= [\Sigma_i (\boldsymbol{\alpha}_1)_{i_1} \lambda^{i_1}] \wedge [\Sigma_{i_2} (\boldsymbol{\alpha}_2)_{i_2} \lambda^{i_2}] \wedge \ldots\ldots \wedge [\Sigma_{i_k} (\boldsymbol{\alpha}_k)_{i_k} \lambda^{i_k}]$$

$$= \alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_k \ . \qquad (8.5.3)$$

This pure tensor functional $\mathcal{T}_\wedge = (\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_k)$ is an element of Λ^k .

The corresponding element of L^k is $T_\wedge = (\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge .... \wedge \boldsymbol{\alpha}_k) = \mathrm{Alt}(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes ... \otimes \boldsymbol{\alpha}_k)$ as in (7.4.3).

Expressing $\alpha_1 \wedge \alpha_2 \wedge ... \wedge \alpha_k$ in terms of the *ordered* expansion is more complicated. One must first compute the tensor A as in (8.4.16) or (7.4.3),

$$(1/k!)\, A = \mathrm{Alt}(T) = \mathrm{Alt}(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes ..... \otimes \boldsymbol{\alpha}_k) = T_\wedge = (\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge ... \wedge \boldsymbol{\alpha}_k) \quad \in \Lambda^k \qquad (8.5.4)$$

Then the ordered expansion (8.4.7) can be written in a battery of ways,

$$\alpha_1 \wedge \alpha_2 \wedge ... \wedge \alpha_k = \qquad // \ \alpha_1 \wedge \alpha_2 \wedge ... \wedge \alpha_k \in \Lambda^k$$

(a) $$= \Sigma_{1 \le i_1 < i_2 < .... < i_k \le n} \ A_{i_1 i_2 ... i_k} (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \qquad // \ (8.4.7)$$

(b) $$= \Sigma_{i_1 < i_2 < .... < i_k} k! \, [\mathrm{Alt}(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes ... \otimes \boldsymbol{\alpha}_k)]_{i_1 i_2 ... i_k} (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \ . \qquad // \ (8.5.4)$$

(c) $$= \Sigma_{i_1 < i_2 < .... < i_k} k! \, [\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \wedge ... \wedge \boldsymbol{\alpha}_k]_{i_1 i_2 ... i_k} (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \ . \qquad // \ (8.1.3)$$

(d) $$= \Sigma_{i_1 < i_2 < .... < i_k} \det[\, (\boldsymbol{\alpha}_*)_{i_*}] \ (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \ . \qquad // \ (8.2.8c) \text{ with } j_r \to r$$

(e) $$= \Sigma_{i_1 < i_2 < .... < i_k} \det \begin{pmatrix} (\boldsymbol{\alpha}_1)_{i_1} & (\boldsymbol{\alpha}_1)_{i_2} & ... & (\boldsymbol{\alpha}_1)_{i_k} \\ (\boldsymbol{\alpha}_2)_{i_1} & (\boldsymbol{\alpha}_2)_{i_2} & ... & (\boldsymbol{\alpha}_2)_{i_k} \\ ... & ... & ... & ... \\ (\boldsymbol{\alpha}_k)_{i_1} & (\boldsymbol{\alpha}_k)_{i_2} & ... & (\boldsymbol{\alpha}_k)_{i_k} \end{pmatrix} (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \ .$$

(f) $$= \Sigma_{i_1 < i_2 < .... < i_k} \det \begin{pmatrix} (\boldsymbol{\alpha}_1)_{i_1} & (\boldsymbol{\alpha}_2)_{i_1} & ... & (\boldsymbol{\alpha}_k)_{i_1} \\ (\boldsymbol{\alpha}_1)_{i_2} & (\boldsymbol{\alpha}_2)_{i_2} & ... & (\boldsymbol{\alpha}_k)_{i_2} \\ ... & ... & ... & ... \\ (\boldsymbol{\alpha}_1)_{i_k} & (\boldsymbol{\alpha}_2)_{i_k} & ... & (\boldsymbol{\alpha}_k)_{i_k} \end{pmatrix} (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \ .$$

(g) $$= \Sigma_{i_1 i_2 ... i_k} \ (\boldsymbol{\alpha}_1)_{i_1} (\boldsymbol{\alpha}_2)_{i_2} ... \ (\boldsymbol{\alpha}_k)_{i_k} \ (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \qquad // \ (8.5.2) \qquad (8.5.5)$$

where we throw in the symmetric sum at the end. Remember that, since generally dim(V) = n > k, the determinant in (f) is a full-width minor of matrix $M = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 ..... \boldsymbol{\alpha}_k]$. If k = n, the minor is the full matrix.

<u>Example</u> : Suppose k = n = 3. Then the following sum (form (f)) has only one term,

$$\alpha_1 \wedge \alpha_2 \wedge \alpha_3 = \Sigma_{1 \le i_1 < i_2 < i_3 < 3} \det[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3] \, (\lambda^{i_1} \wedge \lambda^{i_2} \wedge \lambda^{i_3})$$

$$= \det[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3] \, (\lambda^{i_1} \wedge \lambda^{i_2} \wedge \lambda^{i_3}) \qquad (8.5.6)$$

as quoted in (4.4.15).

Here are the above expressions for k = 2 and general n ≥ k :

(a) $\alpha_1 \wedge \alpha_2 = \Sigma_{i_1<i_2} A_{i_1 i_2} (\lambda^{i_1} \wedge \lambda^{i_2})$

(b) $= \Sigma_{i_1<i_2} 2! [\mathrm{Alt}(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2)]_{i_1 i_2} (\lambda^{i_1} \wedge \lambda^{i_2})$

(c) $= \Sigma_{i_1<i_2} 2! (\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2)_{i_1 i_2} (\lambda^{i_1} \wedge \lambda^{i_2})$

(d) $= \Sigma_{i_1<i_2} \det[(\boldsymbol{\alpha}_*)_{i_*}] (\lambda^{i_1} \wedge \lambda^{i_2})$

(e) $= \Sigma_{i_1<i_2} \det \begin{pmatrix} (\boldsymbol{\alpha}_1)_{i_1} & (\boldsymbol{\alpha}_1)_{i_2} \\ (\boldsymbol{\alpha}_2)_{i_1} & (\boldsymbol{\alpha}_2)_{i_2} \end{pmatrix} (\lambda^{i_1} \wedge \lambda^{i_2})$

(f) $= \Sigma_{i_1<i_2} \det \begin{pmatrix} (\boldsymbol{\alpha}_1)_{i_1} & (\boldsymbol{\alpha}_2)_{i_1} \\ (\boldsymbol{\alpha}_1)_{i_2} & (\boldsymbol{\alpha}_2)_{i_2} \end{pmatrix} (\lambda^{i_1} \wedge \lambda^{i_2}) = \Sigma_{i_1<i_2}[(\boldsymbol{\alpha}_1)_{i_1}(\boldsymbol{\alpha}_2)_{i_2} - (\boldsymbol{\alpha}_2)_{i_1}(\boldsymbol{\alpha}_1)_{i_2}](\lambda^{i_1} \wedge \lambda^{i_2})$

(g) $= \Sigma_{i_1 i_2} (\boldsymbol{\alpha}_1)_{i_1} (\boldsymbol{\alpha}_2)_{i_2} (\lambda^{i_1} \wedge \lambda^{i_2}) = [\Sigma_{i_1}(\boldsymbol{\alpha}_1)_{i_1}\lambda^{i_1}] \wedge [\Sigma_{i_2}(\boldsymbol{\alpha}_2)_{i_2} \lambda^{i}] = \alpha_1 \wedge \alpha_2$ (8.5.7)

Result (f) matches that shown in (4.4.12),

$$\alpha \wedge \beta = \Sigma_{ij}\, \alpha_i\beta_j\, (\lambda^i \wedge \lambda^j) = \Sigma_{i<j}\, (\alpha_i\beta_j - \alpha_j\beta_i)\, (\lambda^i \wedge \lambda^j) = \Sigma_{i<j}\, A_{ij}\, (\lambda^i \wedge \lambda^j)$$

$$= \Sigma_{i<j} \det \begin{pmatrix} \alpha_i & \beta_i \\ \alpha_j & \beta_j \end{pmatrix} (\lambda^i \wedge \lambda^j) \qquad A_{ij} = (\alpha_i\beta_j - \alpha_j\beta_i) = \det \begin{pmatrix} \alpha_i & \beta_i \\ \alpha_j & \beta_j \end{pmatrix} . \qquad \textit{(4.4.12)}$$

8.6 Number of elements in Λ^k compared with V^{*k}.

We know from (6.1.5) and (8.3.6) that,

$\dim(V^{*k}) = n^k$ // number of basis elements of V^{*k} *(6.1.5)*

$\dim(\Lambda^k) = \binom{n}{k}$ // number of basis elements of Λ^k *(8.3.6)*

If the number of elements of field K is N (N → ∞ for K= reals), then

$$\text{ratio} = \frac{\#\text{ elements of } \Lambda^k}{\#\text{ elements of } V^{*k}} = \frac{\binom{n}{k} N}{n^k N} = \frac{\binom{n}{k}}{n^k} = \binom{n}{k} / n^k . \qquad (8.6.1)$$

For a given n, this is a strongly decreasing function of k, see graph in (7.6.2).

8.7 Multiindex notation

In this section, multiindex versions of equations are shown in red.

Multiindexing is done in two different ways. First, for the symmetric expansion (8.4.4) :

$$\mathcal{T}^\wedge = \Sigma_{i_1 i_2 \ldots i_k} T_{i_1 i_2 \ldots i_k} (\lambda^{i1} \wedge \lambda^{i2} \ldots\ldots \wedge \lambda^{ik}) \qquad (8.4.4)$$

$$\mathcal{T}^\wedge = \Sigma_I T_I \lambda^{\wedge I} \qquad \text{where } \lambda^{\wedge I} \equiv \lambda^{i1} \wedge \lambda^{i2} \wedge \ldots\ldots \wedge \lambda^{ik} \qquad T_I \equiv T_{i_1 i_2 \ldots i_k}$$

and $I \equiv \{i_1, i_2, \ldots\ldots i_k\}$ with $1 \leq i_r \leq n$ = *ordinary* multiindex, $n = \dim(V^*)$. (8.7.1)

The more significant notation involves the ordered expansion (8.4.7) which has only one term for each linearly independent basis element. Note our use of Σ'_I (prime) to indicate an ordered multiindex summation :

$$\mathcal{T}^\wedge = \Sigma_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} A_{i_1 i_2 \ldots i_k} (\lambda^{i1} \wedge \lambda^{i2} \ldots\ldots \wedge \lambda^{ik}) \,. \qquad (8.4.7)$$

$$\mathcal{T}^\wedge = \Sigma'_I A_I \lambda^{\wedge I} \qquad \text{where } \lambda^{\wedge I} \equiv \lambda^{i1} \wedge \lambda^{i2} \wedge \ldots\ldots \wedge \lambda^{ik} \qquad A_I \equiv A_{i_1 i_2 \ldots i_k}$$

and $I \equiv \{i_1, i_2, \ldots\ldots i_k\}$ with $1 \leq i_1 < i_2 < \ldots\ldots < i_k \leq n$ = *ordered* multiindex, $n = \dim(V^*)$. (8.7.2)

Here are some unofficial multiindex notations for other equations developed above:

$$\mathcal{T}^\wedge = \alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_k \qquad \mathcal{T}^\wedge = (\wedge \alpha_Z) \qquad (8.5.3)$$

$$T_{i_1 i_2 \ldots i_k} = (\boldsymbol{\alpha}_1)_{i_1} (\boldsymbol{\alpha}_2)_{i_2} \ldots (\boldsymbol{\alpha}_k)_{i_k} \equiv (\boldsymbol{\alpha}_Z)_I \qquad T_I = (\boldsymbol{\alpha}_Z)_I \qquad (8.5.1a)$$

with the idea that Z = 1,2...k . Continuing on,

$$\mathcal{T}^\wedge = \Sigma_{i_1 i_2 \ldots i_k} (\boldsymbol{\alpha}_1)_{i_1} (\boldsymbol{\alpha}_2)_{i_2} \ldots (\boldsymbol{\alpha}_k)_{i_k} (\lambda^{i1} \wedge \lambda^{i2} \ldots\ldots \wedge \lambda^{ik}) \qquad \mathcal{T}^\wedge = \Sigma_I (\boldsymbol{\alpha}_Z)_I \lambda^{\wedge I} \qquad (8.5.2)$$

$$A = k!\, \mathrm{Alt}(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes \ldots\ldots \otimes \boldsymbol{\alpha}_k) \qquad A = k!\, \mathrm{Alt}(\otimes \boldsymbol{\alpha}_Z) \qquad (8.5.4)$$

$$\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_k = \Sigma_{i_1 < i_2 < \ldots < i_k} k!\, [\mathrm{Alt}(\boldsymbol{\alpha}_1 \otimes \boldsymbol{\alpha}_2 \otimes \ldots \otimes \boldsymbol{\alpha}_k)]_{i_1 i_2 \ldots i_k} (\lambda^{i1} \wedge \lambda^{i2} \ldots\ldots \wedge \lambda^{ik}) \,. \qquad (8.5.5b)$$

$$(\wedge \alpha_Z) = \Sigma'_I \; k!\, \mathrm{Alt}(\otimes \boldsymbol{\alpha}_Z)_I \lambda^{\wedge I}$$

$$A_{i_1 i_2 \ldots i_k} = \det \begin{pmatrix} (\boldsymbol{\alpha}_1)_{i_1} & (\boldsymbol{\alpha}_2)_{i_1} & \ldots & (\boldsymbol{\alpha}_k)_{i_1} \\ (\boldsymbol{\alpha}_1)_{i_2} & (\boldsymbol{\alpha}_2)_{i_2} & \ldots & (\boldsymbol{\alpha}_k)_{i_2} \\ \ldots & \ldots & \ldots & \ldots \\ (\boldsymbol{\alpha}_1)_{i_k} & (\boldsymbol{\alpha}_2)_{i_k} & \ldots & (\boldsymbol{\alpha}_k)_{i_k} \end{pmatrix} \qquad A_I = \det[(\boldsymbol{\alpha}_Z)_I] \qquad (8.5.5a{+}f)$$

$$\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_k = \Sigma_{i_1 < i_2 < \ldots < i_k} \det \begin{pmatrix} (\boldsymbol{\alpha}_1)_{i_1} & (\boldsymbol{\alpha}_2)_{i_1} & \ldots & (\boldsymbol{\alpha}_k)_{i_1} \\ (\boldsymbol{\alpha}_1)_{i_2} & (\boldsymbol{\alpha}_2)_{i_2} & \ldots & (\boldsymbol{\alpha}_k)_{i_2} \\ \ldots & \ldots & \ldots & \ldots \\ (\boldsymbol{\alpha}_1)_{i_k} & (\boldsymbol{\alpha}_2)_{i_k} & \ldots & (\boldsymbol{\alpha}_k)_{i_k} \end{pmatrix} (\lambda^{i1} \wedge \lambda^{i2} \ldots\ldots \wedge \lambda^{ik}). \qquad (8.5.5f)$$

$(\hat{}\alpha_Z) = \Sigma'_I \ \det[(\boldsymbol{\alpha}_Z)_I]\lambda\hat{}^I$

8.8 The Exterior Algebra Λ(V)

We now construct the graded algebra Λ(V) in analogy with that of T(V) in (5.4.1).

Define a large vector space of the form (this is "the dual exterior algebra on V")

$$\Lambda(V) \equiv \Lambda^0 \oplus \Lambda^1 \oplus \Lambda^2 \oplus \Lambda^3 + ... \qquad // \ \Lambda(V) = \Sigma^{\oplus}{}_{k=0}{}^{\infty} \Lambda^k(V) \qquad (8.8.1)$$

Here Λ^0 = the space of scalars, Λ^1 the space of dual vectors, $\Lambda^2 = \Lambda \wedge \Lambda \subset V^{*2}$ the space of antisymmetric dual rank-2 tensors (8.4.6), and so on. The most general element of the space Λ(V) would have the form

$$X = s \oplus \Sigma_i T_i \lambda^i \oplus \Sigma_{ij} T_{ij} \lambda^i\wedge\lambda^j \oplus \Sigma_{ijk} T_{ijk} \lambda^i\wedge\lambda^j\wedge\lambda^k + ...$$

or

$$X = s \oplus \Sigma_i T_i \lambda^i \oplus \Sigma_{i<j} A_{ij} \lambda^i\wedge\lambda^j \oplus \Sigma_{i<j<k} A_{ijk} \lambda^i\wedge\lambda^j\wedge\lambda^k + ... \qquad (8.8.2)$$

The direct sum ⊕ is described in Appendix B.

Associativity of the Wedge Product

See discussion near (7.8.3) and replace $e_i \rightarrow \lambda^i$ and $v\rightarrow\alpha$. One then concludes that,

Fact: The wedge product of k *vectors* $\alpha_1\wedge \alpha_2\wedge .....\wedge \alpha_k$ can be "associated" in any manner without altering the meaning of the product. By this we mean that parentheses can be added in any manner without altering the object. (8.8.4)

What this in effect does is *define* an array of new objects to be the same as $\alpha_1\wedge \alpha_2\wedge .....\wedge \alpha_6$. For example,

$$(\alpha_1\wedge (\alpha_2\wedge \alpha_3 \wedge \alpha_4) \wedge (\alpha_5\wedge \alpha_6) \equiv \alpha_1\wedge \alpha_2\wedge \alpha_3\wedge \alpha_4\wedge \alpha_5\wedge \alpha_6$$
$$\alpha_1\wedge (\alpha_2\wedge (\alpha_3\wedge \alpha_4)\wedge \alpha_5)\wedge \alpha_6 \equiv \alpha_1\wedge \alpha_2\wedge \alpha_3\wedge \alpha_4\wedge \alpha_5\wedge \alpha_6 \qquad (8.8.5)$$

Since tensors like $\mathcal{T}\hat{}$ can be expanded on $(\lambda^{i_1} \wedge \lambda^{i_2} \wedge .... \wedge \lambda^{i_k})$, and since one may associate this wedge product arbitrarily as claimed in (8.8.4), one easily shows that :

Fact: The wedge product of N general dual *tensors* $\mathcal{A}\hat{}\wedge\mathcal{B}\hat{}\wedge\mathcal{C}\hat{}$.... can be "associated" in any manner without altering the meaning of the product. By this we mean that parentheses can be added in any manner without altering the object. (8.8.7)

This fact then extends the claim (8.8.4) made for N vectors, and is exactly analogous to the similar axiomatic statement for ⊗ associativity made in (2.8.21).

Example: In (8.7.1) multiindex notation, consider three Λ(V) tensors $\mathcal{T}^\wedge, \mathcal{S}^\wedge, \mathcal{R}^\wedge$ of rank k,k',k" :

$(\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge) \wedge \mathcal{R}^\wedge = ((\Sigma_I T_I \lambda^{\wedge I}) \wedge (\Sigma_J S_J \lambda^{\wedge J})) \wedge (\Sigma_K R_K \lambda^{\wedge K})$

$= \Sigma_I T_I \Sigma_J S_J \{ (\lambda^{\wedge I} \wedge \lambda^{\wedge J}) \wedge (\Sigma_K R_K \lambda^{\wedge K}) \}$ // rules (8.2.3)

$= \Sigma_I T_I \Sigma_J S_J \Sigma_K R_K (\lambda^{\wedge I} \wedge \lambda^{\wedge J}) \wedge (\lambda^{\wedge K})$ // rules (8.2.3) again

$= \Sigma_{IJK} T_I S_J R_K (\lambda^{\wedge I} \wedge \lambda^{\wedge J} \wedge \lambda^{\wedge K})$ // detail shown above (7.8.8) with $u_i \rightarrow \lambda^i$

$= (\Sigma_I T_I \lambda^{\wedge I}) \wedge (\Sigma_J S_J \lambda^{\wedge J}) \wedge (\Sigma_K R_K \lambda^{\wedge K})$ // rules (8.2.3) again

$= \mathcal{T}^\wedge \wedge \mathcal{S}^\wedge \wedge \mathcal{R}^\wedge$.

Our example shows that for arbitrary Λ(V) tensors, $(\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge) \wedge \mathcal{R}^\wedge = \mathcal{T}^\wedge \wedge \mathcal{S}^\wedge \wedge \mathcal{R}^\wedge$.

Fact: The large space Λ(V) is in fact itself a vector space. (8.8.8)

We know this is true since $\Lambda(V) = \Sigma^{\oplus}{}_{k=0}{}^{\infty} \Lambda^k$ and we showed in (8.3.2) that each Λ^k is a vector space. For example, the "0" element in Λ(V) is the direct sum of the "0" elements of all the Λ^k. See Appendix B for more detail.

To show that Λ(V) is an algebra, we must show that it is closed under both addition and multiplication. It should be clear to the reader that Λ(V) is closed under addition and has the right scalar rule. For example, if k_1 and s are scalars,

$k_1 \oplus \alpha \oplus \beta\wedge\kappa \oplus \rho\wedge\sigma\wedge\eta$ = sum of 4 elements of Λ(V) = an element of Λ(V)

$s(k_1 \oplus \alpha \oplus \beta\wedge\kappa \oplus \rho\wedge\sigma\wedge\eta) = (sk_1) \oplus (s\alpha) \oplus (s\beta)\wedge\kappa \oplus \rho\wedge(s\sigma)\wedge\eta$ = element of Λ(V) (8.8.9)

This additive closure is of course necessary for Λ(V) be a vector space.

The space is also closed under the multiplication operation ^. For example

$(\beta\wedge\kappa)\wedge(\rho\wedge\sigma\wedge\eta) = \beta\wedge\kappa\wedge\rho\wedge\sigma\wedge\eta \;=\in\; \Lambda^5 =\in\; \Lambda(V)$. // $(\beta\wedge\kappa) \in \Lambda^2$ $(\rho\wedge\sigma\wedge\eta) \in \Lambda^3$ (8.8.10)

Here we have used the associative property (8.8.4). This closure claim is stated more generally below (8.9.a.6).

One then makes the following definitions with regard to the space Λ(V), where n = dim(V*) = dim(V):
(the objects in this table are pure multilinear functionals)

Object	Name	any blade lincomb:	Grade(rank):	Space	
s	dual 0-blade	scalar ∈ K	0	Λ^0	
α	dual 1-blade	dual vector	1	Λ^1	
α^β	dual 2-blade	dual bivector	2	Λ^2	
α^β^γ	dual 3-blade	dual trivector	3	Λ^3	
α^β^γ^δ	dual 4-blade	dual quadvector	4	Λ^4	
.....					
α^β^γ^δ^....	dual k-blade	dual k-vector	k	Λ^k	
....					
α^β^γ^δ^....	dual n-blade	dual n-vector	n	Λ^n	
arbitrary element of Λ(V)		dual multivector	mixed	Λ(V)	(8.8.11)

Since Λ(V) is closed under the operations ⊕ and ^, it is "an algebra" (the space Λ^k alone is not an algebra because it is not closed under ^). The Λ(V) algebra is different from that of the reals due to its definition as a sum of vector spaces. The elements of Λ(V) have different "grades" as shown above, so Λ(V) is a "graded algebra". Sometimes Λ(V) is called "the dual exterior tensor algebra" over V.

A **k-blade** is a pure wedge product of k vectors, whereas a **k-vector** is any *linear combination* of k-blades. A **multivector** is any linear combination of k-vectors for any mixed values of k.

Note that

$$s_1(\alpha\text{^}\beta) \oplus s_2(\gamma\text{^}\delta) = (s_1\alpha)\text{^}\beta \oplus (s_2\gamma)\text{^}\delta = (\alpha'\text{^}\beta) \oplus (\gamma'\text{^}\delta) \qquad \text{// 2-blades}$$
$$s_1(\alpha\text{^}\beta) \oplus s_2(\gamma\text{^}\delta\text{^}\varepsilon) = (s_1\alpha)\text{^}\beta \oplus (s_2\gamma)\text{^}\delta\text{^}\varepsilon = (\alpha'\text{^}\beta) \oplus (\gamma'\text{^}\delta\text{^}\varepsilon) \qquad \text{// multivector}$$

so it is also correct to say that a k-vector is any *sum* of k-blades, and a multivector is any *sum* of k-vectors. That is, any linear combination can be written as a sum as shown in the above examples.

Unlike in Tensor World, in Wedge World the above list (8.8.11) is finite for a given n = dim(V). For k = n there is exactly one linearly independent basis vector which is the ordered wedge product of all the basis vectors of V*. For k > n, all wedge products vanish since the vectors in the wedge product are linearly dependent, see (8.2.6). The dimensionality of the space Λ(V) is as follows, based on (8.8.1) and (B.10)',

$$\dim[\Lambda(V)] = \dim[\Lambda^0 \oplus \Lambda^1 \oplus \Lambda^2 \oplus \Lambda^3 + ...] = \dim(\Lambda^0) + \dim(\Lambda^1) + \dim(\Lambda^2) + \dim(\Lambda^3) + ...$$

but for dim(V*) = n this series truncates with Λ^n and we find from (7.3.6),

$$\dim[\Lambda(V)] = 1 + n + \binom{n}{2} + \binom{n}{3} + ... + \binom{n}{n} = \Sigma_{k=0}{}^n \binom{n}{k} = 2^n \quad = \text{a finite number} \qquad (8.8.12)$$

Recall from the discussion above (4.4.34) that the space Λ^2 of rank-2 tensor functionals is isomorphic to the space $\Lambda^2{}_f$ of rank-2 tensor functions, where we added a subscript f to distinguish these two vector spaces. We apply this similar notation to the full space Λ(V) to obtain this tensor function version of (8.8.1),

$$\Lambda_f(V) \equiv \Lambda^0{}_f \oplus \Lambda^1{}_f \oplus \Lambda^2{}_f \oplus \Lambda^3{}_f + ... \qquad // \ \Lambda_f(V) = \Sigma^{\oplus}{}_{k=0}{}^{\infty} \Lambda^k{}_f(V) \qquad (8.8.13)$$

where now $\Lambda_f(V)$ is the space of all multilinear totally antisymmetric (alternating) functions of any number of vector arguments.

8.9 The Wedge Product of two or more dual tensors in Λ(V)

(a) Wedge Product of two dual tensors $\mathcal{T}_\wedge$ and $\mathcal{S}_\wedge$

Rather than translate the many details of this section from Chapter 7, we will skip these details and state the conclusions. The details may be obtained from Section 7.9 by making these simple replacements:

$$u_{i_1} \to \lambda^{i_1} \qquad u_I \to \lambda^I \qquad u_{\wedge I} \to \lambda_\wedge{}^I$$

$$T^{i_1 i_2 .... i_k} \to T_{i_1 i_2 .... i_k}, \quad T^I \to T_I \ , \quad T_\wedge \to \mathcal{T}_\wedge$$
$$S^{i_1 i_2 .... i_k} \to S_{i_1 i_2 .... i_k}, \quad S^I \to S_I \ , \quad S_\wedge \to \mathcal{S}_\wedge \ .$$

In subsection (d) below on the product of three tensors, more details are provided.

Here then are selected results:

Tensor product of two tensors:

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = \Sigma_{\mathbf{I}} (T\otimes S)_{\mathbf{I}} \lambda_\wedge{}^{\mathbf{I}} \quad \mathbf{I} \equiv I, I' = i_1, i_2 ... i_{k+k'}, \ \lambda_\wedge{}^I \equiv (\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) \ . \qquad (8.9.a.5)$$

Closure: $\mathcal{T}_\wedge \in \Lambda^k$ and $\mathcal{S}_\wedge \in \Lambda^{k'} \quad \Rightarrow \quad \mathcal{T}_\wedge \wedge \mathcal{S}_\wedge \in \Lambda^{k+k'} \subset \Lambda(V)$. (8.9.a.6)

Basis relation: $(\lambda^{i_1} \wedge \lambda^{i_2} ..... \wedge \lambda^{i_k}) = \mathrm{Alt}(\lambda^{i_1} \otimes \lambda^{i_2} \otimes .... \otimes \lambda^{i_k})$

$$\lambda_\wedge{}^I = \mathrm{Alt}(\lambda^I) \qquad (8.3.8)$$

$$\mathcal{T}\otimes\mathcal{S} == \Sigma_{\mathbf{I}} (T\otimes S)_{\mathbf{I}} \lambda^{\mathbf{I}} \quad \mathbf{I} \equiv I, I' = i_1, i_2 ... i_{k+k'}, \ u_{\mathbf{I}} \equiv (u_{i_1} \otimes u_{i_2} .... \otimes u_{i_{k+k'}}) \ . \qquad (5.6.5)$$

$\mathrm{Alt}(\mathcal{T}\otimes\mathcal{S})(v_{\mathbf{J}}) = \mathrm{Alt}_{\mathbf{J}}(\mathcal{T}\otimes\mathcal{S})(v_{\mathbf{J}}) = \Sigma_{\mathbf{I}} (T\otimes S)_{\mathbf{I}} \ \mathrm{Alt}_{\mathbf{J}} (\lambda^{\mathbf{I}}(v_{\mathbf{J}}))$ // (5.6.5) and (A.5.10) that Alt is linear

$= \Sigma_{\mathbf{I}} (T\otimes S)_{\mathbf{I}} \ \mathrm{Alt}_{\mathbf{I}} (\lambda^{\mathbf{I}}(v_{\mathbf{J}}))$ // (A.8.31), $\lambda^{\mathbf{I}}(v_{\mathbf{J}})$ has factored form $\lambda^{i_1}(v_{j_1}) \lambda^{i_2}(v_{j_2})$ ...

$= \Sigma_{\mathbf{I}} (T\otimes S)_{\mathbf{I}} \ \lambda_\wedge{}^{\mathbf{I}}(v_{\mathbf{J}})$ // (8.3.8)

$= (\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge)(v_{\mathbf{J}})$ // (8.9.a.5)

so

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = \mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}) \ . \qquad (8.9.a.7)$$

The "components" (tensor functions) are

$$(\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge)(v_{\mathcal{J}}) = \text{Alt}(\mathcal{T} \otimes \mathcal{S})(v_{\mathcal{J}})$$

$$= \frac{1}{(k+k')!} \Sigma_P (-1)^{S(P)} (\mathcal{T} \otimes \mathcal{S})(v_{P(\mathcal{J})})$$

$$= \frac{1}{(k+k')!} \Sigma_P (-1)^{S(P)} \mathcal{T}(v_{P(J)}) \mathcal{S}(v_{P(J')}) \tag{8.9.a.8}$$

where

$$J \equiv j_1, j_2 \dots j_k \qquad J' \equiv j_{k+1}, j_{k+2}, \dots j_{k+k'} \qquad \mathcal{J} \equiv J, J' = j_1, j_2 \dots j_{k+k'} \tag{7.9.a.4}$$

The above is an explicit instruction for computing the "components" of the tensor $\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge$. We have added this new notation,

$$\mathcal{T}(v_{P(J)}) \equiv \mathcal{T}(v_{j_{P(1)}}, v_{j_{P(2)}} \cdots v_{j_{P(k)}}) \qquad \text{for } J \equiv j_1, j_2 \dots j_k \ . \tag{8.9.a.9}$$

Example: Let $\mathcal{S}$ and $\mathcal{T}$ both be rank-2 dual tensors so k = k' = 2 . Then

$$(\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge)(v_I) = (\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge)(v_1,v_2,v_3,v_4) = (1/4!)\, \Sigma_P (-1)^{S(P)} \mathcal{T}(v_{i_{P(1)}}, v_{i_{P(2)}}) \mathcal{S}(v_{i_{P(3)}}, v_{i_{P(4)}})$$

$$= (1/24)\, [\, \mathcal{T}(v_{i_1},v_{i_2})\mathcal{S}(v_{i_3},v_{i_4}) - \mathcal{T}(v_{i_2},v_{i_1})\mathcal{S}(v_{i_3},v_{i_4}) + \mathcal{T}(v_{i_2},v_{i_3})\mathcal{S}(v_{i_1},v_{i_4}) + 21 \text{ more terms}] \tag{8.9.a.10}$$

Here as elsewhere we show in red the indices to be swapped to make the next term. From (8.9.c.6) below,

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = (-1)^{2*2} \mathcal{S}_\wedge \wedge \mathcal{T}_\wedge = \mathcal{S}_\wedge \wedge \mathcal{T}_\wedge . \tag{8.9.a.11}$$

(b) Special cases of the wedge product $\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge$

Same as Section 7.9 (b) with $T_\wedge \rightarrow \mathcal{T}_\wedge$ and $S_\wedge \rightarrow \mathcal{S}_\wedge$. Here are the conclusions :

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = \kappa \wedge \mathcal{S}_\wedge = \mathcal{S}_\wedge \wedge \mathcal{T}_\wedge = \mathcal{S}_\wedge \wedge \kappa = \kappa \mathcal{S}_\wedge \qquad \text{if } \mathcal{T}_\wedge = \kappa \in V^{*0} \ [\, V^{*0} = V^0 \,]$$

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = \mathcal{T}_\wedge \wedge \kappa' = \mathcal{S}_\wedge \wedge \mathcal{T}_\wedge = \kappa' \wedge \mathcal{T}_\wedge = \kappa' \mathcal{T}_\wedge \qquad \text{if } \mathcal{S}_\wedge = \kappa' \in V^{*0}$$

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = \kappa \wedge \kappa' = \mathcal{S}_\wedge \wedge \mathcal{T}_\wedge = \kappa' \wedge \kappa = \kappa\kappa' \qquad \text{if } \mathcal{T}_\wedge, \mathcal{S}_\wedge = \kappa, \kappa' \in V^{*0} \tag{8.9.b.3}$$

(c) Commutivity Rule for the Wedge Product of two dual tensors $\mathcal{T}_\wedge$ and $\mathcal{S}_\wedge$

Same as Section 7.9 (c) with $T_\wedge \rightarrow \mathcal{T}_\wedge$ and $S_\wedge \rightarrow \mathcal{S}_\wedge$ and u→λ. Here are some of the translated conclusions:

$$(\lambda_\wedge^J \wedge \lambda_\wedge^I) = (-1)^{kk'} (\lambda_\wedge^I \wedge \lambda_\wedge^J) . \qquad \text{dual basis vectors} \tag{8.9.c.5}$$

$$\mathcal{S}_\wedge \wedge \mathcal{T}_\wedge = (-1)^{kk'} \mathcal{T}_\wedge \wedge \mathcal{S}_\wedge \qquad \text{ranks of the two dual tensors are k and k' .} \tag{8.9.c.6}$$

(d) Wedge Product of three or more dual tensors

For this section we do a full translation of Section 7.9 (d) :

$$\mathcal{T}\hat{}\wedge\mathcal{S}\hat{}\wedge\mathcal{R}\hat{} = [\Sigma_I T_I\lambda\hat{}^{I}]\wedge[\Sigma_J\ S_J\lambda\hat{}^{J}]\wedge[\Sigma_K\ R_K\lambda\hat{}^{K}]$$

(a) $= \Sigma_{I,J,K}\ T_I S_J R_K\ (\lambda\hat{}^{I}) \wedge (\lambda\hat{}^{J}) \wedge (\lambda\hat{}^{K})$

(b) $= \Sigma_{I,J,K}\ T_I S_J R_K\ (\lambda\hat{}^{I} \wedge \lambda\hat{}^{J} \wedge \lambda\hat{}^{K})$ // associative of ^ used here

(d) $= \Sigma_{I,I',I''}\ T_I S_{I'} R_{I''}\ (\lambda\hat{}^{I} \wedge \lambda\hat{}^{I'} \wedge \lambda\hat{}^{I''})$ // rename multiindices J→I',K→I"

$I \equiv i_1, i_2 ... i_k$ $\qquad I' \equiv i_{k+1}, i_{k+2}, ...i_{k+k'}$ $\qquad I'' \equiv i_{k+k'+1}, i_{k+k'+2}, ...i_{k+k'+k''}$

$\lambda\hat{}^{I} \equiv (\lambda^{i_1}\wedge....\wedge \lambda^{i_k})$ $\qquad \lambda\hat{}^{I'} \equiv (\lambda^{i_{k+1}}\wedge...\wedge\lambda^{i_{k+k'}})$ $\qquad \lambda\hat{}^{I''} \equiv (\lambda^{i_{k+k'+1}}\wedge...\wedge\lambda^{i_{k+k'+k''}})$

(e) $= \Sigma_{\mathit{I}}\ (T\otimes S\otimes R)_{\mathit{I}}\ \lambda\hat{}^{\mathit{I}} \qquad \lambda\hat{}^{\mathit{I}} \equiv (\lambda^{i_1}\wedge....\wedge \lambda^{i_{k+k'+k''}}) \qquad \mathit{I} \equiv I, I',I'' = i_1,i_2...i_{k+k'+k''}$ (8.9.d.1)

The outer product form is $T_I S_{I'} R_{I''} = (T\otimes S\otimes R)_{I,I',I''} = (T\otimes S\otimes R)_{\mathit{I}}$.

The conclusion is this:

$$\mathcal{T}\hat{}\wedge\mathcal{S}\hat{}\wedge\mathcal{R}\hat{} = \Sigma_{\mathit{I}}\ (T\otimes S\otimes R)_{\mathit{I}}\lambda\hat{}^{\mathit{I}} \qquad \mathit{I} \equiv I, I',I'' = i_1,i_2...i_{k+k'+k''}\ ,\ \lambda\hat{}^{\mathit{I}} \equiv (\lambda^{i_1}\wedge....\wedge \lambda^{i_{k+k'+k''}}) \quad (8.9.d.2)$$

Since the $\lambda\hat{}^{\mathit{I}}$ are basis vectors in $\Lambda^{k+k'+k''}$, we have shown that:

$$\mathcal{T}\hat{} \in \Lambda^k \text{ and } \mathcal{S}\hat{} \in \Lambda^{k'} \text{ and } \mathcal{R}\hat{} \in \Lambda^{k''} \quad\Rightarrow\quad \mathcal{T}\hat{}\wedge\mathcal{S}\hat{}\wedge\mathcal{R}\hat{} \in \Lambda^{k+k'+k''} \subset \Lambda(V)\ . \quad (8.9.d.3)$$

Recalling the Chapter 6 result,

$$\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R} = \Sigma_{\mathit{I}}\ (T\otimes S\otimes R)_{\mathit{I}}\ \lambda^{\mathit{I}} \qquad \mathit{I} \equiv I, I',I'' = i_1,i_2...i_{k+k'+k''} \qquad \lambda^{I} \equiv (\lambda^{i_1}\otimes....\otimes \lambda^{i}{}_{k+k'+k''}) \quad \textit{(6.6.5)}$$

and (8.3.8) that $\lambda\hat{}^{\mathit{I}} = \mathrm{Alt}(\lambda^{\mathit{I}})$, we find,

$\mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}) = \Sigma_{\mathit{I}}\ (T\otimes S\otimes R)_{\mathit{I}}\ \mathrm{Alt}(\lambda^{\mathit{I}})$ // Alt is linear, see (7.9.d.4)

$= \Sigma_{\mathit{I}}\ (T\otimes S\otimes R)_{\mathit{I}}\ \lambda\hat{}^{\mathit{I}}$ // (8.3.8)

$= \mathcal{T}\hat{}\wedge\mathcal{S}\hat{}\wedge\mathcal{R}\hat{}$ // (8.9.d.2)

so

$$\mathcal{T}\hat{}\wedge\mathcal{S}\hat{}\wedge\mathcal{R}\hat{} = \mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}) \quad (8.9.d.4)$$

and then

$$[\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge \wedge \mathcal{R}^\wedge](v_I) = [\mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R})](v_I)$$

$$= \Sigma_P(-1)^{S(P)}\,(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R})(v_{P(I)}) \qquad // \text{(A.5.3)}$$

$$= \Sigma_P(-1)^{S(P)}\,\mathcal{T}(v_{P(I)})\mathcal{S}(v_{P(I')})\mathcal{R}(v_{P(I")}) \qquad (8.9.d.5)$$

which gives instructions for how to compute the "components" of $\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge \wedge \mathcal{R}^\wedge$.

Using the systematic notation outlined in (5.6.10) through (5.6.12), and generalizing the above development for the wedge product of three tensors, we find the following expansion for the wedge product of N tensors of Λ(V),

$$(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge ... \wedge (\mathcal{T}_N)^\wedge = \Sigma_I\ (T_1)_{I_1}(T_2)_{I_2} \ldots (T_N)_{I_N}\ \lambda^\wedge{}^I = \Sigma_I\ (T_1\otimes T_2 \ldots \otimes T_N)_I\ \lambda^\wedge{}^I$$

$$\text{where} \qquad \lambda^\wedge{}^I = \lambda^{i_1}\wedge\lambda^{i_2} \ldots \wedge \lambda^{i_{k_1+k_2+\ldots+k_N}} = \lambda^{i_1}\wedge\lambda^{i_2} \ldots \wedge \lambda^{i_\kappa} \qquad (8.9.d.6)$$

$$\text{and} \qquad (T_1\otimes T_2 \ldots \otimes T_N)_I = (T_1)_{I_1}(T_2)_{I_2} \ldots (T_N)_{I_N} \ .$$

The rank of this product tensor is then $\kappa = \Sigma_{i=1}^N k_i$ and the tensor is an element of $\Lambda^\kappa \subset \Lambda(V)$. Notice that if κ > n, the tensor product (8.9.d.6) *vanishes* since there are then > n factors in $\lambda^\wedge{}^I$ so one or more are then duplicated,

$$(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge ... \wedge (\mathcal{T}_N)^\wedge = 0 \qquad \text{if } \kappa = \Sigma_{i=1}^N k_i \ge n+1 \ . \qquad (8.9.d.7)$$

For example, if all the tensors are the same tensor $\mathcal{T}^\wedge$ of rank k, then

$$\mathcal{T}^\wedge{}^N \equiv \mathcal{T}^\wedge \wedge \mathcal{T}^\wedge \wedge ... \wedge \mathcal{T}^\wedge = 0 \qquad \text{if } Nk \ge n+1 \ \text{ or } \ N \ge (n+1)/k \ . \qquad (8.9.d.8)$$

If N ≥ (n+1), then N ≥ (n+1)/k for any k ≥ 1. Thus

$$\mathcal{T}^\wedge{}^N = 0 \qquad \text{for any } N \ge n+1 \ \text{ assuming } k \ne 0. \qquad (8.9.d.9)$$

Recall (6.6.16),

$$\mathcal{T}_1\otimes\mathcal{T}_2\otimes ... \otimes\mathcal{T}_N = \Sigma_I\ (T_1{}^{I_1}T_2{}^{I_2} \ldots T_N{}^{I_N})\ \lambda^I = \Sigma_I\ (T_1\otimes T_2 \ldots \otimes T_N)^I\ \lambda^I \ . \qquad \textit{(6.6.16)}$$

Applying Alt to both sides again with $\lambda^\wedge{}^I = \mathrm{Alt}(\lambda^I)$ shows that, as in (8.9.d.4),

$$(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge ... \wedge (\mathcal{T}_N)^\wedge = \mathrm{Alt}(\mathcal{T}_1\otimes\mathcal{T}_2\otimes ... \otimes\mathcal{T}_N) \ . \qquad (8.9.d.10)$$

"Components" (the tensor function) of this tensor are computed as follows:

$$[(\mathcal{T}_1)^{\wedge} \wedge (\mathcal{T}_2)^{\wedge} \wedge ... \wedge (\mathcal{T}_N)^{\wedge}](v_I) = [\mathrm{Alt}(\mathcal{T}_1 \otimes \mathcal{T}_2 \otimes ... \otimes \mathcal{T}_N)](v_I)$$

$$= \Sigma_P (-1)^{S(P)} (\mathcal{T}_1 \otimes \mathcal{T}_2 \otimes ... \otimes \mathcal{T}_N)(v_{P(I)}) \qquad // \text{ (A.5.3)}$$

$$= \Sigma_P (-1)^{S(P)} \mathcal{T}_1(v_{P(I_1)}) \mathcal{T}_2(v_{P(I_2)}) ... \mathcal{T}_N(v_{P(I_N)}) \qquad (8.9.d.11)$$

where $\mathcal{T}_1(v_{P(I_1)}) \equiv \mathcal{T}_1(v_{i_{P(1)}}, v_{i_{P(2)}} ... v_{i_{P(k)}})$ for $I_1 = i_1, i_2 ... i_{\kappa_1}$

$\mathcal{T}_2(v_{P(I_2)}) \equiv \mathcal{T}_2(v_{i_{P(\kappa_1+1)}}, v_{i_{P(\kappa_1+2)}} ... v_{i_{P(\kappa_2)}})$ for $I_2 = \{i_{\kappa_1+1}, i_{\kappa_1+2} ..... i_{\kappa_2}\}$

etc. // see (5.6.10 thru 12) for details

In the Dirac notation of Section 2.11 one can write (8.9.d.10) as

$$<(\mathcal{T}_1)^{\wedge}| \wedge <(\mathcal{T}_2)^{\wedge}| \wedge ... \wedge <(\mathcal{T}_N)^{\wedge}| = \mathrm{Alt}(<\mathcal{T}_1| \otimes <\mathcal{T}_2| \otimes ... \otimes <\mathcal{T}_N|) \ . \qquad (8.9.d.12)$$

It is shown in (C.4.14) that "pre-antisymmetrization makes no difference", so the above may also be written

$$<(\mathcal{T}_1)^{\wedge}| \wedge <(\mathcal{T}_2)^{\wedge}| \wedge ... \wedge <(\mathcal{T}_N)^{\wedge}| = \mathrm{Alt}(<(\mathcal{T}_1)^{\wedge}| \otimes <(\mathcal{T}_2)^{\wedge}| \otimes ... \otimes <(\mathcal{T}_N)^{\wedge}|) \ . \qquad (8.9.d.13)$$

Both sides of this equation are elements of the dual wedge product space $\Lambda^{k_1+k_2+..+k_N}$, but they are also both elements of the larger dual tensor product space $V^{*k_1} \otimes V^{*k_2} \otimes ... \otimes V^{*k_N}$. The action of linear operator $\mathcal{Q}$ on a dual tensor product space vector is defined in the obvious manner, as in (6.6.18),

$$[<(\mathcal{T}_1)^{\wedge}| \otimes <(\mathcal{T}_2)^{\wedge}| \otimes ... \otimes <(\mathcal{T}_N)^{\wedge}|] \mathcal{Q} = <(\mathcal{T}_1)^{\wedge}|\mathcal{Q} \otimes <(\mathcal{T}_2)^{\wedge}|\mathcal{Q} \otimes ... \otimes <(\mathcal{T}_N)^{\wedge}|\mathcal{Q} \ . \qquad (8.9.d.14)$$

In other words, the action of $\mathcal{Q}$ on the larger space is defined in terms of its action on the spaces which make up the tensor product. This result holds as well for the wedge product of N dual tensors,

$$[<(\mathcal{T}_1)^{\wedge}| \wedge <(\mathcal{T}_2)^{\wedge}| \wedge ... \wedge <(\mathcal{T}_N)^{\wedge}|] \mathcal{Q} = <(\mathcal{T}_1)^{\wedge}|\mathcal{Q} \wedge <(\mathcal{T}_2)^{\wedge}|\mathcal{Q} \wedge ... \wedge <(\mathcal{T}_N)^{\wedge}|\mathcal{Q} \qquad (8.9.d.15)$$

<u>Proof</u>: $[<(\mathcal{T}_1)^{\wedge}| \wedge <(\mathcal{T}_2)^{\wedge}| \wedge ... \wedge <(\mathcal{T}_N)^{\wedge}|] \mathcal{Q} = [\mathrm{Alt}(<(\mathcal{T}_1)^{\wedge}| \otimes <(\mathcal{T}_2)^{\wedge}| \otimes ... \otimes <(\mathcal{T}_N)^{\wedge}|)] \mathcal{Q}$

$$= \mathrm{Alt} [(<(\mathcal{T}_1)^{\wedge}| \otimes <(\mathcal{T}_2)^{\wedge}| \otimes ... \otimes <(\mathcal{T}_N)^{\wedge}|) \mathcal{Q}]$$

$$= \mathrm{Alt} [(<(\mathcal{T}_1)^{\wedge}|\mathcal{Q} \otimes <(\mathcal{T}_2)^{\wedge}|\mathcal{Q} \otimes ... \otimes <(\mathcal{T}_N)^{\wedge}|\mathcal{Q})]$$

$$= <(\mathcal{T}_1)^{\wedge}|\mathcal{Q} \wedge <(\mathcal{T}_2)^{\wedge}|\mathcal{Q} \wedge ... \wedge <(\mathcal{T}_N)^{\wedge}|\mathcal{Q} \ .$$

Equations (8.9.d.14,15) are the transposes of (7.9.d.14,15) if we set $Q = P^T$.

(e) Commutativity Rule for product of N dual tensors

The argument of Section 7.9 (e) can be repeated with u→λ. Here we just quote the conclusion.

Fact: In a product of tensors $(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge (\mathcal{T}_3)^\wedge$.... of rank k_1, k_2, k_3 ... , if two tensors are swapped $(\mathcal{T}_r)^\wedge \leftrightarrow (\mathcal{T}_s)^\wedge$ (with r < s), the resulting tensor incurs the following sign relative to the starting tensor,

$$\text{sign} = (-1)^m \qquad \text{where } m = (k_{r+1}+k_{r+2} ...+k_{s-1})(k_r+k_s) + k_r k_s \ . \tag{8.9.e.6}$$

Corollary: If the sum of the ranks of the two swapped tensor is even, in effect $m = k_r k_s$. (8.9.e.7)

Example:

$$(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge (\mathcal{T}_3)^\wedge = (-1)^m (\mathcal{T}_3)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge (\mathcal{T}_1)^\wedge \qquad r = 1 \quad s = 3$$

$$m = (k_2)(k_1+k_3) + k_1k_3 = k_1k_2 + k_1k_3 + k_2k_3 \qquad (-1)^m = (-1)^{k_1k_2+k_1k_3+k_2k_3} \tag{8.9.e.8}$$

(f) Theorems from Appendix C : pre-antisymmetrization makes no difference

We showed above that one can form wedge products of elements of Λ(V) in this manner,

$$\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge = \text{Alt}(\mathcal{T}\otimes\mathcal{S}) \ . \tag{8.9.a.7}$$

$$\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge \wedge \mathcal{R}^\wedge = \text{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}) \tag{8.9.d.4}$$

$$(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge ... \wedge (\mathcal{T}_N)^\wedge = \text{Alt}(\mathcal{T}_1\otimes\mathcal{T}_2\otimes...\otimes\mathcal{T}_N) \tag{8.9.d.7}$$

where the operator Alt acts on the vector arguments which are not displayed in the above compact functional notation. For example

$$\mathcal{T}^\wedge \wedge \mathcal{S}^\wedge = \text{Alt}(\mathcal{T}\otimes\mathcal{S})$$

means, in multiindex notation,

$$(\mathcal{T}^\wedge \wedge S^\wedge)(v_I) = \text{Alt}_I [(\mathcal{T}\otimes S)(v_I)] = \text{Alt}_I [\, T(v_I)S(v_{I'})] = \frac{1}{(k+k')!} \Sigma_P(-1)^{S(P)}\, T(v_{P(I)})S(v_{P(I')}) \ .$$

A very simple case is the following (recall for vectors that $\alpha = \alpha^\wedge$)

$$(\alpha \wedge \beta)(v_{i_1},v_{i_2}) = \text{Alt}\,(\alpha\otimes\beta)(v_{i_1},v_{i_2}) = \text{Alt}\,[\alpha(v_{i_1})\beta(v_{i_2})] = \frac{1}{(1+1)!}\Sigma_P(-1)^{S(P)}\, \alpha(v_{i_{P(1)}})\, \beta(v_{i_{P(2)}})$$

$$= (1/2)\,[\, \alpha(v_{i_1})\beta(v_{i_2}) - \alpha(v_{i_2})\beta(v_{i_1})] = (1/2)[\, (\alpha\otimes\beta)(v_{i_1},v_{i_2}) - (\beta\otimes\alpha)(v_{i_1},v_{i_2})\,]$$

$$= \{\, (1/2)\,[(\alpha\otimes\beta) - (\beta\otimes\alpha)\,]\}(v_{i_1},v_{i_2})$$

which replicates our Chapter 4 statement that

$$\alpha \wedge \beta = [\alpha\otimes\beta - \beta\otimes\alpha]/2 \quad . \qquad \textit{(4.4.1)}$$

The objects here are functionals in $\Lambda(V)$ which, when closed with a vector set, become tensor functions in $\Lambda_f(V)$.

Appendix C uses the rearrangement theorem in three separate Theorems to show that

$$\mathcal{T}_\wedge \wedge \mathcal{S}_\wedge = \text{Alt}(\mathcal{T}\otimes\mathcal{S}) \quad = \underset{\text{Theorem One}}{\text{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S})} \quad = \underset{\text{Theorem Two}}{\text{Alt}(\mathcal{T}\otimes\mathcal{S}_\wedge)} \quad = \underset{\text{Theorem Three}}{\text{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}_\wedge)} \quad . \qquad (8.9.f.1)$$

These three theorems are derived in a generic space with vectors |1,2...k> and so apply to both tensors and tensor functionals, T^I and $\mathcal{T}(v_I)$.

Recall that

$$\mathcal{T}_\wedge \equiv \text{Alt}(\mathcal{T}) \qquad \textit{(8.4.3)}$$

so that $\mathcal{T}_\wedge$ is a totally antisymmetric tensor functional. What (8.9.f.1) says is that $\text{Alt}(\mathcal{T}\otimes\mathcal{S})$ provides total antisymmetrization on all the (undisplayed) vector argument indices, so pre-antisymmetrizing either or both tensors makes no difference. A similar statement applies to working with totally symmetric tensors. So we have,

$$\text{Alt}[\mathcal{T}\otimes\mathcal{S}] = \text{Alt}[\mathcal{T}_\wedge\otimes\mathcal{S}] = \text{Alt}[\mathcal{T}\otimes\mathcal{S}_\wedge] = \text{Alt}[\mathcal{T}_\wedge\otimes\mathcal{S}_\wedge]$$
$$\text{where} \qquad \mathcal{T}_\wedge = \text{Alt}(\mathcal{T}) \quad \mathcal{S}_\wedge = \text{Alt}(\mathcal{S}) \qquad \textit{(C.4.1)}$$

$$\text{Sym}[\mathcal{T}\otimes\mathcal{S}] = \text{Sym}[\mathcal{T}_s\otimes\mathcal{S}] = \text{Sym}[\mathcal{T}\otimes\mathcal{S}_s] = \text{Sym}[\mathcal{T}_s\otimes\mathcal{S}_s]$$
$$\text{where} \qquad \mathcal{T}_s = \text{Sym}(\mathcal{T}) \quad \mathcal{S}_s = \text{Sym}(\mathcal{S}) \ . \qquad \textit{(C.4.2)}$$

These can of course be rewritten as

$$\text{Alt}[\mathcal{T}\otimes\mathcal{S}] = \text{Alt}[\text{Alt}(\mathcal{T})\otimes\mathcal{S}] = \text{Alt}[\mathcal{T}\otimes\text{Alt}(\mathcal{S})] = \text{Alt}[\text{Alt}(\mathcal{T})\otimes\text{Alt}(\mathcal{S})] \qquad \textit{(C.4.3)}$$

$$\text{Sym}[\mathcal{T}\otimes\mathcal{S}] = \text{Sym}[\text{Sym}(\mathcal{T})\otimes\mathcal{S}] = \text{Sym}[\mathcal{T}\otimes\text{Sym}(\mathcal{S})] = \text{Sym}[\text{Sym}(\mathcal{T})\otimes\text{Sym}(\mathcal{S})] \ . \qquad \textit{(C.4.4)}$$

Similarly Appendix C shows that

$$\begin{aligned} \mathcal{T}_\wedge \wedge \mathcal{S}_\wedge \wedge \mathcal{R}_\wedge &= \text{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}) = \text{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}\otimes\mathcal{R}) = \text{Alt}(\mathcal{T}\otimes\mathcal{S}_\wedge\otimes\mathcal{R}) = \text{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}_\wedge) \\ &= \text{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}_\wedge\otimes\mathcal{R}) = \text{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}\otimes\mathcal{R}_\wedge) = \text{Alt}(\mathcal{T}\otimes\mathcal{S}_\wedge\otimes\mathcal{R}_\wedge) \\ &= \text{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}_\wedge\otimes\mathcal{R}_\wedge) \quad . \qquad (8.9.f.2) \end{aligned}$$

Adding ^ subscripts inside an Alt expression changes nothing. Here is another example:

$$\mathcal{T}_\wedge\wedge\mathcal{S}_\wedge\wedge\mathcal{R}_\wedge = \mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}) = \mathrm{Alt}((\mathcal{T}\otimes\mathcal{S})\otimes\mathcal{R}) = \mathrm{Alt}((\mathcal{T}\otimes\mathcal{S})_\wedge\otimes\mathcal{R}) = \mathrm{Alt}(\mathrm{Alt}(\mathcal{T}\otimes\mathcal{S})\otimes\mathcal{R})$$

and

$$\mathcal{T}_\wedge\wedge\mathcal{S}_\wedge\wedge\mathcal{R}_\wedge = \mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}\otimes\mathcal{R}) = \mathrm{Alt}(\mathcal{T}\otimes(\mathcal{S}\otimes\mathcal{R})) = \mathrm{Alt}(\mathcal{T}\otimes(\mathcal{S}\otimes\mathcal{R})_\wedge) = \mathrm{Alt}(\mathcal{T}\otimes\mathrm{Alt}(\mathcal{S}\otimes\mathcal{R})) \ . \qquad (8.9.f.3)$$

This appears in Spivak p 80 as

$$(2)\ \ Alt(Alt(\omega \otimes \eta) \otimes \theta) = Alt(\omega \otimes \eta \otimes \theta) = Alt(\omega \otimes Alt(\eta \otimes \theta)). \qquad (C.4.8)$$

(g) Spivak Normalization

We won't repeat the discussion of Section 7.9 (g), but the reader can do the translation with the usual rules

$$\mathbf{v} \rightarrow \alpha \qquad \mathbf{e}_i \rightarrow \lambda^i \qquad T^I \rightarrow T_I \qquad \text{etc.}$$

Here are the results, where factors shown in red show changes caused by the Spivak notation in which the normalization factor in (8.1.2) is changed from (1/k!) to 1,

$$\alpha_{j_1}\wedge \alpha_{j_2}\wedge \ .....\wedge \alpha_{j_k} = 1\ \Sigma_P (-1)^{S(P)} (\ \alpha_{P(j_1)}\otimes \alpha_{P(j_2)}\otimes \ .....\otimes \alpha_{P(j_k)})$$

$$= 1\ [\ (\alpha_{j_1} \otimes \alpha_{j_2} \otimes \ .....\otimes \alpha_{j_k}) \quad + \quad \text{all signed permutations}\ ]$$

$$= k!\ \mathrm{Alt}(\alpha_{j_1} \otimes \alpha_{j_2} \otimes \ .....\otimes \alpha_{j_k}) \ . \qquad (8.1.2)_S$$

In particular,

$$\alpha \wedge \beta = 1[\ \alpha\otimes\beta - \beta\otimes\alpha] \ . \qquad // \text{ no factor of } 1/2 \qquad (4.4.1)_S$$

The affected equations are these:

$$(\lambda^{i_1}\wedge \lambda^{i_2}\wedge \ .....\wedge \lambda^{i_k}) = k!\ \mathrm{Alt}(\lambda^{i_1}\otimes \lambda^{i_2}\otimes \ .....\otimes \lambda^{i_k}) \qquad \text{or} \qquad \lambda^{\wedge I} = k!\mathrm{Alt}(\lambda^I) \qquad (8.3.8)_S$$

$$\mathcal{T}_\wedge \equiv k!\mathrm{Alt}(\mathcal{T}) \qquad \text{and} \qquad \mathcal{S}_\wedge \equiv k'!\mathrm{Alt}(\mathcal{S}) \ . \qquad (8.4.3)_S$$

$$\mathcal{T}_\wedge\wedge \mathcal{S}_\wedge = (k+k')!\ \mathrm{Alt}(\mathcal{T}\otimes\mathcal{S}) \ . \qquad (8.9.a.7)_S$$

$$\mathcal{T}_\wedge\wedge \mathcal{S}_\wedge = \frac{(k+k')!}{k!\ k'!}\ \mathrm{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}_\wedge) \qquad \mathcal{T}_\wedge \in \Lambda^k \ \text{and}\ \mathcal{S}_\wedge \in \Lambda^{k'} \ . \qquad (8.9.g.1)$$

$$\mathcal{T}_\wedge\wedge \mathcal{S}_\wedge\wedge \mathcal{R}_\wedge = \frac{(k+k'+k'')!}{k!\ k'!k''!}\ \mathrm{Alt}(\mathcal{T}_\wedge\otimes\mathcal{S}_\wedge\otimes\mathcal{R}_\wedge) \qquad \mathcal{T}_\wedge \in \Lambda^k, \mathcal{S}_\wedge \in \Lambda^{k'}, \mathcal{R}_\wedge \in \Lambda^{k''} \ . \qquad (8.9.g.2)$$

These now correspond exactly with Spivak's wedge product definition for tensor functions,

$$\omega \wedge \eta = \frac{(k+l)!}{k!\,l!} \mathrm{Alt}(\omega \otimes \eta).$$

page 79

$$(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta)$$
$$= \frac{(k+l+m)!}{k!\,l!\,m!} Alt(\omega \otimes \eta \otimes \theta).$$

page 80

In (8.3.1a) we showed a table comparing our notation to that of Spivak. Here are a few more items:

us	Spivak		
λ^i	φ_i	dual space basis vectors	
Λ^k	--	space of k-multilinear alternating tensor functionals	
$\Lambda^k{}_f$	$\Lambda^k(V)$	space of k-multilinear alternating tensor functions	
$\mathcal{T}^\wedge, \mathcal{R}^\wedge, \mathcal{S}^\wedge$	ω,η,θ	typical elements of Λ^k (and $\Lambda^k{}_f$)	
k,k',k"	k,ℓ,m	ranks (degrees) of the above typical elements	
σ	P	permutation operator	
sgn σ	$(-1)^{S(P)}$	permutation parity, S(P) = swap count	
S_k	G	set (group) of all permutations of [1,2...k], App. A.	(8.9.g.3)

Comments:

1. Spivak refers to a totally antisymmetric tensor function as an alternating function which is the traditional terminology in this realm, hence the operator name Alt.

2. Spivak uses all lower indices, whereas we have used covariant notation where contravariant indices are up and covariant indices are down.

3. The Spivak normalization is compatible with the traditional definition of a "pullback" as described below in Chapter 10.

9. The Wedge Product as a Quotient Space

We present here a wedge product "theory section" which really should be part of Chapter 1, but we wanted to have the reader first immersed in the nuts and bolts approach to the wedge product presented in Chapters 4, 7 and 8. As is the case for Chapter 1, this chapter makes no mention of the components of vectors or tensors.

9.1. Development of L^k as V^k/S

The presentation below is based on the paragraph titled Definition 3.1 on page 5 of Conrad.

Consider the vector space V defined over some field K (the scalars, normally reals). If V used coefficients in a ring R instead of a field K, V would be called an R-module. Since any field K is also a ring, we can regard our usual V as an R-module (any vector space is also an R-module). Statements about R-modules are more general that statements about vector spaces, so for that reason one sees the R-module moniker in discussions of our current topic. We shall use the bare term module.

Thus, the vector space $V^k = V\otimes V...\otimes V$ can be regarded as a module since its vectors are defined over the field K which is also a ring.

The pure elements of V^k have the form $v_1\otimes v_2\otimes v_3 .... \otimes v_k$ (k factors).

Consider the subset S of V^k whose elements have a *repeat* of one of the vectors. That is, suppose we have $v_i = v_j$ for some $i \neq j$ in $v_1\otimes v_2\otimes v_3 .... \otimes v_k$. There could be other vectors which are also equal to v_i, so at least two vectors are the same. For example, if k = 4 one would say $a\otimes x\otimes b\otimes x$ and $x\otimes x\otimes b\otimes x$ were in the subset S. Adding elements of this subset produces another element of the subset, so this subset is itself a module. Thus we are talking about elements of a submodule S of the module V^k. Notice that 0 is an element of S, which can be represented by any element of V^k having one or more vectors being the 0 vector of V, as in (1.1.9).

For k = 4, consider this element of V^k,

$$A' = 3\ a\otimes b\otimes c\otimes d + 5\ a\otimes b\otimes c\otimes a - 2\ a\otimes a\otimes c\otimes d + 3\ a\otimes a\otimes c\otimes a . \tag{9.1.1}$$

If we were to throw out elements of the set S, we would get

$$A = 3\ a\otimes b\otimes c\otimes d . \tag{9.1.2}$$

The set of elements of V^k that is generated by adding all elements of set S to A is called the coset of A, usually written [A]. Thus, the coset of A is A + s where $s \in S$. The elements of V^k can be *partitioned* into an array in this manner, where each row (coset) involves all the $s_i \in S$:

row name	coset →			
[0]	0	$0 + s_1$	$0 + s_2$	.....
[A]	A	$A + s_1$	$A + s_2$	....
[B]	B	$B + s_1$	$B + s_2$	....
...				

(9.1.3)

For example our V^k element A' lies somewhere in the row of this chart labeled on the left by [A].

It turns out that the rows themselves (the cosets) form a module called V^k/S . The elements of this module can be regarded as being those in the first column of the cosets. So A is an element of V^k/S , but A' is not. Strictly speaking, there is an isomorphism between A and [A], but we ignore such details.

Fact: To enumerate the elements of the module V^k/S we write down all the elements of V^k and just set to 0 all terms in which a vector is repeated, such as the last three terms of A' in (9.1.1). We thus filter out such terms, they are "modded out", which is why V^k/S is sometimes called V^k mod S. (9.1.4)

Define L^k to be

$$L^k \equiv V^k/S \quad . \tag{9.1.5}$$

The fact that V^k elements lying in S (those that have repeated vectors) are "thrown out" (modded out, set equal to 0) is reminiscent of the construction (1.1.4) that $F(VxW)/N = V\otimes W$ and certain elements of the full set F(VxW) were similarly modded out (set to 0, such as $(v_2, w_1+w_2) - (v_2,w_1) - (v_2,w_2)$).

Elements of V^k are written $v_1\otimes v_2\otimes v_3 .... \otimes v_k$. This product is "associative" in that parentheses can be placed any way one wants, such as $v_1\otimes(v_2\otimes v_3) .... \otimes v_k$, with no change in value. (9.1.6)

Elements of $L^k \equiv V^k/S$ are written $v_1\hat{}v_2\hat{}v_3 .... \hat{}v_k$. This product is declared to be "associative" in that parentheses can be placed any way one wants, such as $v_1\hat{}(v_2\hat{}v_3) .... \hat{}v_k$, with no change in value. (9.1.7)

Using this definition of the wedge product of k vectors, we can derive some of its properties.

Fact 1: $v_1\hat{}v_2\hat{}v_3 .... \hat{}v_k = 0$ if two (or more) vectors are the same. (9.1.8)

Proof: This follows from the definition of $L^k \equiv V^k/S$ and the Fact (9.1.4) stated above.

Fact 2: $v_1\hat{}v_2 = - v_2\hat{}v_1$ (9.1.9)

Proof: We know that $(v_1+v_2) \hat{} (v_1+v_2) = 0$ since this has the form $v_3 \hat{} v_3$ which is 0 by Fact 1. Expanding,

$$0 = (v_1+v_2) \hat{} (v_1+v_2) = v_1\hat{}v_1 + v_1\hat{}v_2 + v_2\hat{}v_1 + v_2\hat{}v_2 = v_1\hat{}v_2 + v_2\hat{}v_1$$

so

$0 = v_1\hat{}v_2 + v_2\hat{}v_1$ and $v_1\hat{}v_2 = - v_2\hat{}v_1$ QED

Fact 3: Swapping any pair of vectors in v_1^v_2^v_3.... ^v_k creates a minus sign. (9.1.10)

Proof by example: (swap v_1 and v_3 by making use of associativity and Fact 2 three times) :

$$v_3 \wedge v_2 \wedge v_1 .... \wedge v_k = + v_3 \wedge (v_2 \wedge v_1) .... \wedge v_k = - v_3 \wedge (v_1 \wedge v_2) .... \wedge v_k = - (v_3 \wedge v_1) \wedge v_2 .... \wedge v_k$$

$$= + (v_1 \wedge v_3) \wedge v_2 .... \wedge v_k = + v_1 \wedge (v_3 \wedge v_2) .... \wedge v_k = - v_1 \wedge (v_2 \wedge v_3) .... \wedge v_k$$

$$= - v_1 \wedge v_2 \wedge v_3 .... \wedge v_k \qquad \text{QED}$$

Fact 4: $v_{j_1} \wedge v_{j_2} \wedge .... \wedge v_{j_k} = \varepsilon_{j_1 j_2 .... j_k} (v_1 \wedge v_2 \wedge ..... \wedge v_k)$ (9.1.11)

Proof: Fact 4 is the combination of Fact 3 and Fact 1.

Fact 5: $v_{j_1} \wedge v_{j_2} \wedge .... \wedge v_{j_k} = 0$ if the vectors are linearly dependent. (9.1.12)

Proof: See (7.2.6).

In this manner, we can derive all the properties of the wedge product stated in Section 7.2 without having to lean on the construction of the wedge product as a linear combination of tensor products.

However, we know that the elements of L^k are linear combinations of the elements of V^k. We have written in (7.1.3) that

$$v_1 \wedge v_2 \wedge ..... \wedge v_k = (1/k!) \Sigma_{i_1 i_2 .... i_k} \varepsilon_{i_1 i_2 .... i_k} (v_{i_1} \otimes v_{i_2} \otimes ..... \otimes v_{i_k})$$

$$= (1/k!) \Sigma_P (-1)^{S(P)} (v_{P(1)} \otimes v_{P(2)} \otimes ..... \otimes v_{P(k)}) \quad . \qquad (4.6.2)$$

Since in Section 7.2 this linear combination generates all the Facts listed above, and does not contradict any of them, we conclude that this must be *the* linear combination of V^k elements that equals $v_1 \wedge v_2 \wedge ..... \wedge v_k$ (apart from a possible normalization factor).

Alternate Language. Looking at A and A' above, we could say that A and A' are in the same **equivalence class** so that A ~ A'. Two elements of V^k are in the same equivalence class if they differ by an element of S, so we have A' - A = s ∈ S. The elements of the equivalence class of A are then just the coset [A]. The submodule of V^k called V^k/S is called a **quotient module**. Using category diagrams, one can consolidate this notion with that of quotient rings and quotient groups.

9.2. Development of L as T/I

Start with the tensor algebra (vector space) shown in (5.4.1),

$$T(V) = V^0 \oplus V \oplus V^2 \oplus V^3 \oplus .... \tag{9.2.1}$$

The elements of the vector space T(V) form a **ring** with operations ⊕ and ⊗. It is easy to show that T(V) is closed under addition ⊕ and multiplication ⊗ and has the other required ring properties.

Ideal Example 1: Consider the set S of elements of T(V) which are linear combinations of elements of the form A⊗B⊗C (with coefficients in field K) where B is some fixed element of T(V) and A,C ∈ T(V) are allowed to vary. This set is closed under addition. For example, A⊗B⊗C + A'⊗B⊗C' ∈ S. Since coefficients in K can be absorbed into A, one could just say that the elements of S are sums of elements of the form A⊗B⊗C. One could take any 0⊗B⊗C as the "0" element, and -A⊗B⊗C is the additive inverse. The set S is commutative and associative under addition. Therefore S forms an additive subgroup of the ring T(V). Moreover if we left or right multiply (using ⊗) any element of this set by any element of T(V), the result clearly lies in T(V).

$$Q\otimes(A\otimes B\otimes C) \in T(V) \qquad\qquad (A\otimes B\otimes C)\otimes Q \in T(V) \ . \tag{9.2.2}$$

Therefore this set S is a **two-sided ideal** of the ring T(V).

Ideal Example 2: S = sums of elements of the form A⊗B⊗C⊗D⊗E where elements B and D are fixed and A,C are E varied, all letters being ∈ T(V) .

Ideal Example 3: S = sums of elements of the form A⊗x⊗C⊗x⊗E where vector x is fixed and A,C,E ∈ T(V) are varied. This set is the set of all sums of elements of T(V) in which the vector x appears at least twice. Let's call this particular ideal by the name S = I, because this is our ideal of interest.

Now suppose we declare the following equivalence relation

$$A\otimes x\otimes C\otimes x\otimes E \ \sim 0 \qquad\qquad x,A,C,E \in T(V). \tag{9.2.3}$$

Sums of such elements form the ideal I discussed above, and we are in effect setting all elements of this ideal equal to 0.

There then exists a subset of T(V) which we shall call T(V)/I, or T(V) "mod" I. This is a standard algebraic structure where one takes the quotient of a ring R divided by a two-sided ideal I of that ring. The upshot is that the elements of the new quotient set T(V)/I consist of all sums of T(V) elements *except that* any term which matches the form (9.2.3) is filtered out ("modded out") by setting it equal to 0.

Example: $t' = k_1\, a\otimes b\otimes c\otimes d + k_2\, a\otimes b + k_3\, b\otimes c\otimes c + k_4\, a\otimes b\otimes c\otimes a$ = element of T(V)

$$t = k_1\, a\otimes b\otimes c\otimes d + k_2\, a\otimes b \qquad = \text{element of } T(V)/I \tag{9.2.4}$$

In algebra terminology, adding all elements of the form A⊗x⊗C⊗x⊗E to t generates a **coset** associated with t called [t], and T(V)/I is in effect the set of all such cosets. Element t' is one element of the t coset. The elements of T(V)/I themselves form a new ring called the **quotient ring** or **factor ring**. The ring/ideal situation is quite similar to that discussed above for the module/submodule situation V^k/S.

Recall from (7.8.1) that the full wedge (exterior) tensor algebra is given by the direct sum space

$$L(V) = L^0 \oplus L^1 \oplus L^2 \oplus L^3 + ... \quad . \tag{9.2.5}$$

The claim then is that

$$L(V) = T(V)/I \qquad \text{where } I = \text{the ideal of Example 3 above.} \tag{9.2.6}$$

This is then the space of all T(V) elements where all terms in which a vector is repeated are set to 0 and thus are not part of L(V).

Notice that the quotient of Section 9.1 has a finer granularity. It deals with individual $L^k \subset V^k$ spaces, whereas Section 5.2 deals with the entire $L(V) \subset T(V)$.

Many texts refer to L^k as $\Lambda^k(V)$ and L(V) as $\Lambda(V)$. We have reserved the Λ names for the dual spaces.

The category theory approaches to L^k and L(V) are similar to the discussion of Section 1.2 with the main point being that L^k and L(V) are "universal" and therefore uniquely defined up to isomorphism. The role played by k-multilinear functions is played by antisymmetric k-multilinear functions.

10. Differential Forms

In this chapter we consider aspects of the topic of differential forms from the viewpoint of Chapter 2 on the tensor algebra of transformations, and Chapter 8 on the dual exterior algebra of wedge products.

10.1. Differential Forms Defined

A differential form is in fact just an element of the wedge space $\Lambda^k(V)$ described in Chapter 8. Recall that the most general element of $\Lambda^k(V)$ was written in symmetric sum notation as (sums run 1 to n = dimV) ,

$$\mathcal{T}_\wedge = \Sigma_{i_1 i_2 \ldots i_k} T_{i_1 i_2 \ldots i_k} (\lambda^{i_1} \wedge \lambda^{i_2} \ldots\ldots \wedge \lambda^{i_k}) \qquad i_r = 1 \text{ to } n \quad n = \dim(V) \quad (8.4.4)$$

$$\mathcal{T}_\wedge = \Sigma_I T_I \lambda_\wedge{}^I \ . \qquad \text{// the above in multiindex notation} \qquad (10.1.1)$$

This sum is redundant since each basis vector appears k! times. In an ordered sum form, each independent basis vector of Λ^k appears only once,

$$\mathcal{T}_\wedge = \Sigma_{1 \le i_1 < i_2 < \ldots < i_k \le n} A_{i_1 i_2 \ldots i_k} (\lambda^{i_1} \wedge \lambda^{i_2} \ldots\ldots \wedge \lambda^{i_k}) \qquad (8.4.7)$$

$$\mathcal{T}_\wedge = \Sigma'_I A_I \lambda_\wedge{}^I \qquad (10.1.2)$$

where Σ'_I indicates the ordered summation.

If one is given $T_{i_1 i_2 \ldots i_k}$ in the first expansion, a viable expression for $A_{i_1 i_2 \ldots i_k}$ which makes the second expansion valid is this

$$A = k! \text{Alt}(T) \qquad (8.4.16) \qquad (10.1.3)$$

which is a shorthand notation for

$$A_I = k!\ \text{Alt}(T_I) \qquad (10.1.4)$$

which in turn means

$$A_{i_1 i_2 \ldots i_k} = k!\ \text{Alt}(T_{i_1 i_2 \ldots i_k})$$

$$= T_{i_1 i_2 \ldots i_k} - T_{i_2 i_1 \ldots i_k} + \text{all other signed permutations} \ . \qquad (10.1.5)$$

Whereas T is an arbitrary rank-k tensor, A obtained from T in (10.1.4) is a totally antisymmetric rank-k tensor if one allows all values of the i_r.

On the other hand, if one is given $A_{i_1 i_2 \ldots i_k}$ in (10.1.2), it is likely that $A_{i_1 i_2 \ldots i_k}$ is not a totally antisymmetric tensor. One might have, for example, $A_{i_1 i_2 \ldots i_k} = \partial_{i_1} B_{i_2 \ldots i_k}$. The sum in (10.1.2) only "senses" the values of $A_{i_1i_2....i_k} = A_I$ for values of I which are ordered, and A_I for non-ordered I play no role. Given some $A_{i_1i_2....i_k}$ in (10.1.2) a viable expression for $T_{i_1i_2....i_k}$ in (10.1.1) is this,

$$T_{i_1i_2...i_k} = \begin{cases} A_{i_1i_2...i_k} & \text{for } i_1 < i_2 < ...< i_k \\ 0 \text{ for all other values of } i_1,i_2...i_k \end{cases}$$

or (10.1.6)

$$T_I = A_I\,\theta(I\text{=ordered}) \qquad // \ \theta(\text{bool}) = 1 \text{ if bool true else } 0$$

since then

$$\Sigma_I T_I \lambda^{\wedge I} = \Sigma_I[\, A_I\,\theta(I\text{=ordered})\,]\,\lambda^{\wedge I} = \Sigma'_I\, A_I \lambda^{\wedge I}\,. \tag{10.1.7}$$

We shall take the vector space V in $\Lambda^k(V)$ to be $V = R^n$.

Below we shall treat the objects T_I and A_I as rank-k tensor *fields* with an argument in R^n, so we will then have for example,

$$A_{i_1i_2...i_k}(\mathbf{x}) \qquad \mathbf{x} \in R^n \qquad (\text{"x-space"})\,. \tag{10.1.8}$$

In the usual presentation of the theory of differential forms, the dual-space basis vector λ^i is given the purely cosmetic name dx^i,

$$\mathit{dx}^i \equiv \lambda^i = \langle \mathbf{u}^i| = (\mathbf{u}^i)^T\,, \quad \mathbf{u}^i = \text{axis-aligned basis vectors of } R^n \tag{10.1.9}$$

where λ^i was a notation introduced in (2.11.c.2).

This object dx^i is very different from the normal calculus differential dx^i, and for that reason we write dx^i in a red italic font. For example, one can then write,

$$\mathit{dx}^i(\mathbf{v}) = \lambda^i(\mathbf{v}) = \langle \mathbf{u}^i|\,\mathbf{v}\rangle = v^i\,. \tag{10.1.10}$$

In contrast, there is no calculus differential object called $dx^i(\mathbf{v})$.

The differential forms (elements of Λ^k) shown above in (10.1.1) and (10.1.2) are now written in cosmetic notation as

$$\mathcal{T}_\wedge = \Sigma_{i_1i_2....i_k}\, T_{i_1i_2....i_k}\,(\mathit{dx}^{i_1} \wedge \mathit{dx}^{i_2} .....\wedge \mathit{dx}^{i_k}) \qquad \mathcal{T}_\wedge = \Sigma_I T_I\, \mathit{dx}_\wedge^{\,I} \tag{10.1.11}$$

$$\mathcal{T}_\wedge = \Sigma_{1\le i_1<i_2<....<i_k\le n}\, A_{i_1i_2...i_k}\,(\mathit{dx}^{i_1} \wedge \mathit{dx}^{i_2} .....\wedge \mathit{dx}^{i_k}) \qquad \mathcal{T}_\wedge = \Sigma'_I A_I\, \mathit{dx}_\wedge^{\,I}\,. \tag{10.1.12}$$

We have used the hat subscript notation to distinguish dual tensors in V^{*k} from those in $\Lambda^k(V)$,

$\lambda^I = \lambda^{i_1} \otimes \lambda^{i_2} .....\otimes \lambda^{i_k}$ // basis vector in dual space V^{*k}

$\lambda^{\wedge I} = \lambda^{i_1} \wedge \lambda^{i_2} .....\wedge \lambda^{i_k}$ // basis vector in dual space $\Lambda^k(V)$. (10.1.13)

The traditional names for differential forms are α, β, ω and so on, so we take $\mathcal{T}^\wedge \rightarrow \alpha$ and write our arbitrary differential k-form (10.1.12) now as

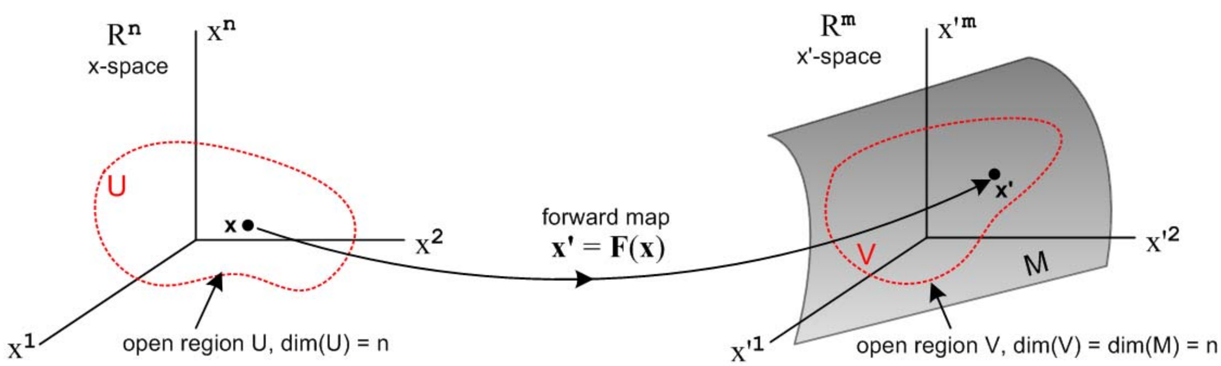
$$\alpha = \Sigma'_I \ f_I(\mathbf{x})\, \lambda^{\wedge I} = \Sigma'_I \ f_I(\mathbf{x})\, dx^{\wedge I} \qquad \alpha \in \Lambda^k(V) \qquad V = \text{x-space} = R^n \qquad (10.1.14)$$

where f_I is the more traditional name for A_I. Once again, $V = R^n$, Euclidean space, where the basis vectors $\mathbf{u}_i = |\mathbf{u}_i\rangle$ are independent of $\mathbf{x}$, and so the $\lambda^i = \langle\mathbf{u}^i|$ are also independent of $\mathbf{x}$.

Comment: Since our monograph deals with both tensor products and wedge products, we feel it is useful to maintain the distinction between $\lambda^{\wedge I}$ and λ^I, or between $dx^{\wedge I}$ and dx^I . Most discussions of differential forms involve only wedge products and the corresponding wedge spaces, so they write $dx^{\wedge I}$ as dx^I. And of course they don't use our red italic notation, so the final result is just dx^I. Furthermore, many presentations don't show the wedge product ^ symbols, so one sees $dx^I = dx_1 dx_2$ which we would write as $dx^{\wedge I} = dx^1 \wedge dx^2$. (Our use of italics is only to maintain the form/calculus distinction for black and white printed copies of this document.)

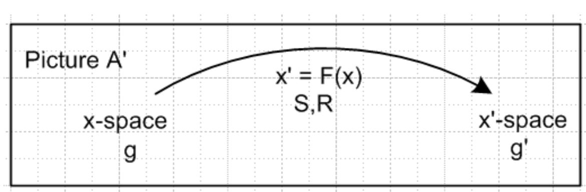
10.2. Differential Forms on Manifolds

Chapter 2 was concerned with the general transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ where x-space and x'-space both had the same dimension N. Here we shall be considering x-space = R^n and x'-space = R^m with $n \le m$. If we allow $\mathbf{x}$ to exhaust some dimension-n region U in x-space, the image $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ will exhaust some region V in x'-space of dimension $n \le m$. If F_i and its derivatives are "smooth" and 1-to-1, the region V in x'-space will lie on a "manifold" M which is embedded in x'-space. Here is a crude graphical representation,



(10.2.1)

Here we have in effect reflected Picture A (2.1.1) left to right to get



(10.2.2)

One can define a differential form $\alpha_{\mathbf{x}'}$ at a point $\mathbf{x}'$ on manifold M in this way,

$$\alpha_{\mathbf{x}'} = \Sigma'_I \, f_I(\mathbf{x}')\, \lambda'^{\wedge I} = \Sigma'_I \, f_I(\mathbf{x}')\, dx'^{\wedge I} = \Sigma'_I \, f_I(\mathbf{x}') <\mathbf{e}'^{\wedge I}| \qquad \in \Lambda'^k \equiv \Lambda^k(R^m) \qquad (10.2.3)$$

where $\lambda'^{\wedge I} = <\mathbf{e}'^{\wedge I}|$ is based on (2.11.c.11). Recall that the $\mathbf{e}'_i$ are axis-aligned basis vectors in x'-space. We think of this differential form $\alpha_{\mathbf{x}'}$ as "being in dual x'-space" to which we give the name Λ'^k .

The manifold M is a "surface" of dimension n within R^m with $n \le m$. The manifold M could be some full chunk of R^m (or all of R^m), in which case it has dimension n = m. If the manifold is a "hypersurface" in R^m it then has dimension n = m-1. In general M is some n-dimensional "surface" embedded within R^m where $1 \le n \le m$.

Recall from (2.5.1) and Fig (2.5.4) (left-right flipped) that the x'-space basis vectors $\mathbf{u}'_i = R\, \mathbf{u}_i$ are the tangent base vectors for the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. For m > n this inverse transformation only exists for points $\mathbf{x}'$ on manifold M, and, for $\mathbf{u}'_i$ with i = 1 to n, the $\mathbf{u}'_i$ continue to be tangent base vectors. The remaining $\mathbf{u}'_i$ for i = n+1 to m can be defined "as needed" to provide a full basis in x'-space R^m .

By the definition of M as the mapping image, we know that the first n $\mathbf{u}'_i$ are "tangent to" the surface M, meaning that tiny arrows $\varepsilon\, \mathbf{u}'_i(\mathbf{x}')$ for $\varepsilon << 1$ lie on M at point $\mathbf{x}'$. Since the full basis is by definition complete in R^m (elements are linearly independent), the remaining $\mathbf{u}'_i(\mathbf{x}')$ for i = n+1 to m are all "normal to" the surface M. This is all specific to some point $\mathbf{x}'$ on M.

For example, for a manifold that is a smooth non-self-intersecting 3D curve embedded in R^3, one would have $\mathbf{u}'_1$ being tangent to the curve at $\mathbf{x}'$, and then $\mathbf{u}'_2$ and $\mathbf{u}'_3$ are both normal to the curve at $\mathbf{x}'$.

On the other hand, if M is a 2D surface in R^3, $\mathbf{u}'_1$ and $\mathbf{u}'_2$ will be tangent to the surface M and $\mathbf{u}'_3$ will be normal to that surface, all at point $\mathbf{x}'$ on M.

The set of n linearly independent basis vectors $\{\mathbf{u}'_1 ... \mathbf{u}'_n\}$ which are tangent to M at $\mathbf{x}'$ are first thought of as having their tails right at the point $\mathbf{x}'$ on M. When these vectors are translated so their tails are all at the origin, the $\{\mathbf{u}'_1 ... \mathbf{u}'_n\}$ then span an n-dimensional vector space. This vector space is usually written $T_{\mathbf{x}'}M$ and is called the **tangent space** to M at point $\mathbf{x}'$ on M, dimension n. As with any vector space, there is a corresponding dual space. The dual space to the tangent space is called the **cotangent space** $T^*_{\mathbf{x}'}M$ and it is the set of all rank-n linear functionals of vectors in $T_{\mathbf{x}'}M$. The name cotangent is like the name covector mentioned below (2.11.a.3) and has nothing to do with the cotangent of any angle.

The conglomeration of all the tangent spaces $T_{\mathbf{x}'}M$ on M has the structure of a fiber bundle and is often called the **tangent bundle**. There is a corresponding dual **cotangent bundle**. See Spivak [1999] Chapter 3, Lang [1999] Chapter III, or wiki on tangent bundles.

As one moves from $\mathbf{x}'$ to a nearby point $\mathbf{x}' + d\mathbf{x}'$ on M, the basis vectors in general will move slightly (M is "smooth"). The dual basis vectors $\mathbf{u}'^i$ of course also move to maintain $\mathbf{u}'^i \bullet \mathbf{u}'_j = \delta^i{}_j$. Thus we have

$\lambda'^i = <\mathbf{u'}^i|$ also depending on $\mathbf{x'}$. We don't want to write this λ'^i as $\lambda'(\mathbf{x'})$ because then we have to write $<\mathbf{u'}^i|\mathbf{v}> = (\lambda'^i(\mathbf{x'}))(\mathbf{v})$ which is rather messy (although Spivak uses this kind of notation with $\mathbf{x'} = \mathbf{p}$ in various places). We hesitate to write the left side $\alpha_{\mathbf{x'}}$ in (10.2.3) as $\alpha(\mathbf{x'})$ because this makes α look like a function, but it is in fact a differential form.

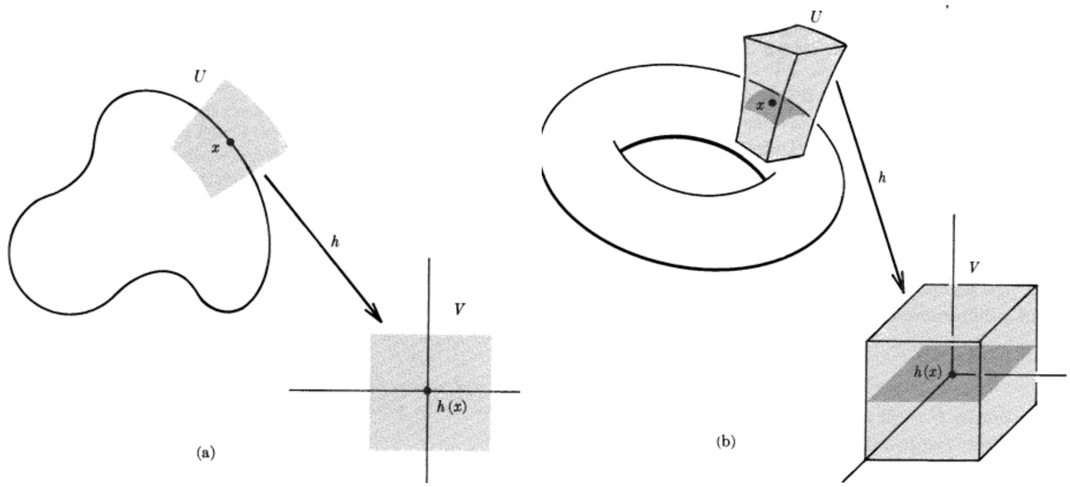
Notice one benefit of the cosmetic notation $\lambda'^{\wedge I} = dx'^{\wedge I}$. The dependence on **x**' can be regarded as being implied by writing *dx'* instead of say *dy*.

It is customary to abbreviate $\alpha_{\mathbf{x'}}$ as just α with the understanding that it is at some point **x'** on M. In proofs below we sometimes call it α' since it is a differential form in dual x'-space.

As already noted, a simple example of a manifold is a non-self-intersecting and "smooth" finite piece of 3D curve hanging in R^3 which is defined by some function $\mathbf{x'} = \mathbf{F}(x)$ where x is a scalar parameter which marks points on the curve. In this case $\alpha_{\mathbf{x'}}$ is a differential 1-form defined at every point **x'** along that curve, and the tangent space at any point **x**' as noted is one dimensional and contains the tangent vector to the curve at that point on the curve.

Our second example of a manifold is a non-self-intersecting and "smooth" finite piece of 2D surface hanging in R^3 which is defined by some function $\mathbf{x'} = \mathbf{F}(\mathbf{x})$ with $\mathbf{x} = (x_1,x_2)$ where every point on the surface is marked by a unique value of **x**. Perhaps this surface is a piece of a toroidal surface or sphere. In this case $\alpha_{\mathbf{x'}}$ is a differential 2-form defined at every point **x'** on that surface. The tangent space at any point **x'** on M is 2 dimensional.

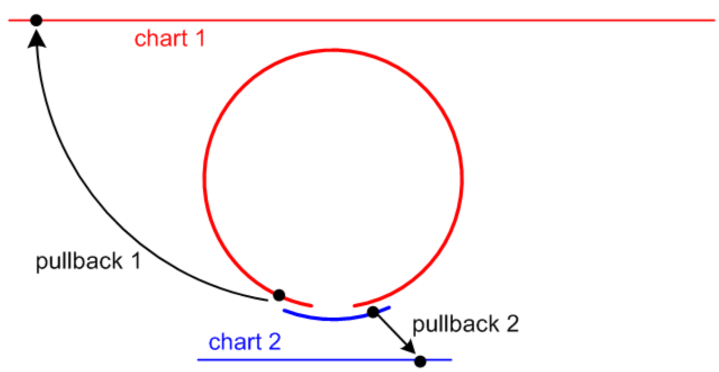
See Sjamaar Chapter 1.1, 6 and 9, Spivak Chapter 5, or elsewhere for a formal definition of a manifold and smoothness. A manifold is roughly a smooth "surface" which can be cobbled together from a set of smooth mappings $\mathbf{x'} = \mathbf{F}_i(\mathbf{x})$ which are said to cover the manifold, the way an atlas of flat maps can cover the entire globe of the Earth. A manifold is a "surface" which is locally smooth in the region of any point **x'** on the manifold. Since $\mathbf{x'} = \mathbf{F}(\mathbf{x})$ must be 1-to-1 between the parameter x-space and x'-space, the manifold cannot be self-intersecting. That is to say, such a point of intersection in x'-space must back-map into at least two different points in x-space hence the map is not 1-to-1. Each mapping has some open domain U_i in R^n and one writes $F_i: U_i \rightarrow M$ and F_i must be 1-to-1 as noted. But $(\partial F_i/\partial x^j) : U_i \rightarrow M$ must also be 1-to-1 to provide clean differentiability at all points on M and in all directions from any such point. This is often stated as (DF_i) must be 1-to-1. What this says is that at any local point in the mapping (on the manifold), there must be differentiable "elbow room" around the point. This is illustrated in Spivak's nice pictures on page 110, where his mapping h is our F^{-1},



(10.2.4)

Spivak would like the mappings h and h^{-1} to both be infinitely differentiable (C^{∞}), in which case he calls h a diffeomorphism. Notice that his U and V are the reverse of ours in (10.2.1) and have one higher dimension. For example, $V_{us} = U_{spivak} \cap M$ which is the gray patch on his toroid above.

Here is a simple example of a manifold (a circle in R^2) being covered by two charts.



(10.2.5)

Most of the circle is "covered" by the larger chart whose t-space is the red line segment at the top. But a small part of the circle shown in blue is covered by a second chart whose x-space is the short lower blue line segment at the bottom. There is some overlap of the charts. So to integrate a 1-form over this circle = manifold, we could do two pullbacks of 1-forms. The circle cannot be covered by just the upper chart extended because then the bottom point on the circle would correspond to both ends of the red segment and then the mapping is not 1-to-1. This is an example of "the seam problem".

Regarding our "cosmetic notation" $dx^i \equiv \lambda^i$ of (10.1.9), the reader can take some support from Lang [1999] page 131, point in $\mathbf{R}^n$, it is customary in the literature to use the notation

$$d\lambda_i(x) = dx_i.$$

This is slightly incorrect, but is useful in formal computations. We shall also use it in this book on occasions. Similarly, we also write (incorrectly)

$$\omega = \sum_{(i)} f_{(i)}\, dx_{i_1} \wedge \cdots \wedge dx_{i_r}$$

instead of the correct

$$\omega(x) = \sum_{(i)} f_{(i)}(x) \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_r}.$$

(10.2.6)

In Lang's second equation, he "incorrectly" writes λ^i as dx^i which is our cosmetic $dx^i \equiv \lambda^i$. We would write his third equation in x-space as $\omega_{\mathbf{x}} = \Sigma'_I f_I(\mathbf{x}) \lambda^{\wedge I}$.

That is our main point from the snippet above, but Lang's first equation also deserves some comment. Recall from (2.11.c.5) that $\lambda^i(\mathbf{x}) = \langle \mathbf{u}^i|\mathbf{x}\rangle = x^i$ where $\mathbf{x}$ is a vector in V. If the $\mathbf{u}^i$ are constant vectors in x-space (as they are for $V = R^n$) then

$$d(\lambda^i(\mathbf{x})) = d(\langle \mathbf{u}^i \mid \mathbf{x}\rangle) = \langle d\mathbf{u}^i|\mathbf{x}\rangle + \langle \mathbf{u}^i| d\mathbf{x}\rangle = \langle \mathbf{u}^i| d\mathbf{x}\rangle = dx^i$$

or

$$d\lambda^i(\mathbf{x}) = dx^i \qquad (10.2.7)$$

and this is the gist of Lang's first equation. This equation is a little confusing for the following reason. Below we show that if we take $\alpha = \Sigma_i f_i(\mathbf{x})\lambda^i$ then $d\alpha = \Sigma_i\Sigma_j (\partial_j f(\mathbf{x}))\lambda^i \wedge \lambda^j$, which is a 2-form. If $f_i(\mathbf{x}) = \delta_{i,n}$ then $(\partial_j f(\mathbf{x})) = 0$. In that case we have $\alpha = \lambda^n$ and $d\alpha = d\lambda^n = 0$ which is then a null 2-form (0 is a valid element of vector space Λ^2). Then we would say

$$(d\lambda^i)(\mathbf{x}) = (0)(\mathbf{x}) = 0 \qquad (10.2.8)$$

which is not the same as (10.2.7) that $d(\lambda^i(\mathbf{x})) = dx^i$.

10.3. The exterior derivative of a differential form

<u>Motivation</u>

The exterior derivative $d\alpha$ plays a key role in the theory of differential forms, as does the notion of the boundary ∂M of a manifold M. Although we shall not derive it, Stokes' Theorem for differential forms says

$$\int_M d\alpha = \int_{\partial M} \alpha \quad .$$

Here α is a k-form, and dα is the exterior derivative of α which we shall see below is a (k+1)-form. The main work involved in proving this theorem involves not so much an understanding of dα as it does dealing with an explicit definition of the boundary ∂M of a manifold M in an arbitrary number of dimensions including issues of orientation. See Sjamaar Chapters 5 and 9.

The single statement above encompasses a large set of integral theorems from analysis only one of which bears the specific Stokes' Theorem moniker. Here we list some of these theorems (red = functionals) ,

$$\int_C \nabla f \bullet dx = \int_{\partial M} f = f(\mathbf{b}) - f(\mathbf{a})$$ "line integral of a gradient theorem" *(H.2.3)*

$$\int_M [\, f(\nabla^2 g) + \nabla f \bullet \nabla g \,]\, dV = \int_{\partial M} f \nabla g \bullet dA$$ "Green's first identity" *(H.3.6)*

$$\int_M (\text{div } \mathbf{F})\, dV = \int_{\partial M} \mathbf{F} \bullet dA$$ "the divergence theorem" *(H.4.6)*

$$\int_M (\partial_1 F_2 - \partial_2 F_1)\, dx^1 \wedge dx^2 = \int_{\partial M} [\, F_1 dx^1 + F_2\, dx^2]$$ "Green's theorem in the plane" *(H.5.6)*

$$\int_M (\text{curl } \mathbf{F}) \bullet dA = \int_{\partial M} \mathbf{F} \bullet dx$$ "traditional Stokes' Theorem" *(H.5.10)*

The sudden appearance of familiar objects like the grad, curl and divergence is part of the Hodge * dual operator "correspondence" we mentioned below (4.3.18). In that correspondence one has

$\alpha = f$	0-form in R^n	$\alpha \leftrightarrow f$	
$d\alpha = \nabla f \bullet dx$	1-form in R^n	$d\alpha \leftrightarrow \nabla f$.	*(H.2.2)*
$\alpha = f$	0-form in R^n	$\alpha \leftrightarrow f$	
$*(d(*d\alpha)) = \nabla^2 f$	0-form on R^n	$*(d(*d\alpha)) \leftrightarrow \nabla^2 f$.	*(H.3.3)*
$\alpha = \mathbf{F} \bullet dx$	1-form on R^n	$\alpha \leftrightarrow \mathbf{F}$	
$*(d(*\alpha)) = \text{div } \mathbf{F}$	0-form on R^n	$*(d(*\alpha)) \leftrightarrow \text{div } \mathbf{F}$	*(H.4.5)*
$\alpha = \mathbf{F} \bullet dx$	1-form in R^3	$\alpha \leftrightarrow \mathbf{F}$	
$*(d\alpha) = [\text{curl } \mathbf{F}] \bullet dx$	1-form in R^3	$*(d\alpha) \leftrightarrow \text{curl } \mathbf{F}$	*(H.5.4)*

where one sees various appearances of the exterior differential operator d on the left side. The action of the Hodge * operator is a follows (see Section H.1),

$dx_\wedge^I$ = some ordered multi-index wedge product of k dx^i in R^n (a basis vector k-form)

$(*dx_\wedge^I) \equiv (\text{sign})_{I,k}\, dx_\wedge^{Ic}$ = ordered wedge product of the *missing* dx^i within R^n (c = complement)

Requirement: $dx_\wedge^I \wedge (*dx_\wedge^I) = dx^1 \wedge dx^2 \ldots dx^n$ which is satisfied by the following sign,

$$(\text{sign})_{I,k} = (-1)^{a+b+..+q}\,(-1)^{k(k+1)/2} \quad \text{where } dx_\wedge^I = dx^a \wedge dx^b \wedge \ldots\wedge dx^q \quad a < b < \ldots < q \qquad (H.1.7)$$

$$\text{Fact: } *(*dx_\wedge^I) = (-1)^{kn+k}\, dx_\wedge^I \qquad (H.1.20)$$

Example: (k=2, n=6) $dx^I = dx^2 \wedge dx^4 \quad \Rightarrow \quad *dx^I = -\, dx^1 \wedge dx^3 \wedge dx^5 \wedge dx^6$

since $(\text{sign})_{I,k} = (-1)^{2+4}\,(-1)^{2(3)/2} = (-1)^6\,(-1)^3 = -1$.

Our intention here is to provide the reader with some *motivation* for slogging through the rest of this section on "d". The above material is treated in Appendix H based on results below.

Definition of the Exterior Derivative

In Section 10.1 we noted that $T_I(\mathbf{x}) = T_{i_1 i_2 \ldots i_k}(\mathbf{x})$ and $A_I(\mathbf{x}) = A_{i_1 i_2 \ldots i_k}(\mathbf{x})$ were rank-k tensor fields with respect to some unspecified Chapter 2 transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ and $d\mathbf{x}' = Rd\mathbf{x}$. We now regard these objects as being just scalar-valued functions which happen to have label I. We refer to either of these functions for the moment as $f(\mathbf{x})$. Such a function by itself is a 0-form because it has no λ^i factors. That is, the object f ,

$$f = f \qquad \in \Lambda^0, \qquad (10.3.1)$$

is a differential 0-form (abbreviated 0-form).

The exterior derivative of such a 0-form is written *df* and is defined as

$$df \equiv \Sigma_{j=1}^{n}\, [\partial f(\mathbf{x})/\partial x^j]\, \lambda^j = \Sigma_{j=1}^{n}\, [\partial_j f(\mathbf{x})]\, \lambda^j \quad . \qquad (10.3.2)$$

Here we put df in red italic so it won't be confused with a calculus differential df of a function $f(\mathbf{x})$. We could have written the 0-form f as *f* , but since then *f* = f there is no reason to do so.

The first thing we notice is that, since f is a 0-form, *df* is a 1-form because the sum is a linear combination of single λ^j dual basis vectors Using the cosmetic notation defined above, we then write (10.3.2) as,

$$df = \Sigma_{j=1}^n [\partial_j f(\mathbf{x})] dx^j \quad . \tag{10.3.3}$$

Now we begin to see the motivation for the cosmetic notation dx^j. The above equation *looks just like* the corresponding calculus equation

$$df = \Sigma_{j=1}^n [\partial_j f(\mathbf{x})] dx^j \ . \tag{10.3.4}$$

In this last equation df(**v**) would make no sense, but in (10.3.3) one can write

$$df(\mathbf{v}) = \Sigma_{j=1}^n [\partial_j f(\mathbf{x})] \lambda^j(\mathbf{v}) = \Sigma_{j=1}^n [\partial_j f(\mathbf{x})] v^j \ . \qquad // (2.11.c.5) \tag{10.3.5}$$

The exterior derivative of a *general* differential form α has an extremely simple definition. Renaming A_I in (10.1.2) to be the more traditional f_I, we write

$$\alpha = \Sigma'_I f_I(\mathbf{x}) \lambda^{\wedge I} \qquad \text{general k-form} \qquad \alpha \in \Lambda^k$$

$$d\alpha \equiv \Sigma'_I (df_I(\mathbf{x})) \wedge \lambda^{\wedge I}$$

$$= \Sigma'_I (\Sigma_{j=1}^n [\partial_j f_I(\mathbf{x})] \lambda^j) \wedge \lambda^{\wedge I} \qquad // \text{ from } (10.3.2)$$

$$= \Sigma_{1 \le i_1 < i_2 < \ldots < i_k \le n} \Sigma_{j=1}^n [\partial_j f_{i_1 i_2 \ldots i_k}(\mathbf{x})] \lambda^j \wedge \lambda^{i_1} \wedge \lambda^{i_2} \ldots \wedge \lambda^{i_k} \ . \tag{10.3.6}$$

Since there are now k+1 wedged dual basis vectors λ^r, this dα must be a (k+1)-form. So,

Fact: If α is a k-form, then dα is a (k+1)-form. (10.3.7)

We pause to take note of a fact that perhaps seems obvious:

Fact: One can compute dα in the same manner for the ordered or the symmetric sum form of α,

$$\alpha = \Sigma'_I A_I(\mathbf{x}) \lambda^{\wedge I} \quad \Rightarrow \quad d\alpha = \Sigma'_I (\Sigma_{j=1}^n [\partial_j A_I(\mathbf{x})] \lambda^j) \wedge \lambda^{\wedge I} \qquad // \text{ ordered sum}$$
$$\alpha = \Sigma_I T_I(\mathbf{x}) \lambda^{\wedge I} \quad \Rightarrow \quad d\alpha = \Sigma_I (\Sigma_{j=1}^n [\partial_j T_I(\mathbf{x})] \lambda^j) \wedge \lambda^{\wedge I} \qquad // \text{ symmetric sum} \tag{10.3.8}$$

where we assume that the $\lambda^{\wedge I}$ are constants in **x**.

<u>Proof</u>: The only question here is whether the dα computed on the second line above is the same as the dα computed on the first line. Assume they are different and call the second line dα". Reorder to get,

$$d\alpha = \Sigma_{j=1}^n \lambda^j \wedge (\Sigma'_I [\partial_j A_I(\mathbf{x})] \lambda^{\wedge I})$$

$$d\alpha'' = \Sigma_{j=1}^n \lambda^j \wedge (\Sigma_I [\partial_j T_I(\mathbf{x})] \lambda^{\wedge I}) \ . \tag{10.3.9}$$

But write (10.1.1) = (10.1.2) and then apply ∂_j to both sides,

$$\Sigma_I\ T_I(\mathbf{x})\ \lambda^{\wedge I} = \Sigma'_I\ A_I\lambda^{\wedge I} \qquad \Rightarrow \qquad \Sigma_I\ \partial_j T_I(\mathbf{x})\ \lambda^{\wedge I} = \Sigma'_I\ \partial_j A_I(\mathbf{x})\ \lambda^{\wedge I}$$

and thus the two right-side expressions in (10.3.9) are the same and so α" = α. QED

So far we have shown that if α is a k-form, then dα is a (k+1)-form.

What can be said about $d^2\alpha \equiv d(d\alpha)$? One might reasonably think this would be a (k+2)-form, but that is not correct. In fact:

Fact: $d^2\alpha = 0$ for any k-form α (differential forms have zero "curvature") . (10.3.10)

Since a differential form involves *linear* functionals, the above Fact seems intuitively reasonable.

Proof: The proof is quite simple if we use the redundant symmetric sum (10.1.1) to express α. Then

$$\alpha = \Sigma_I T_I(\mathbf{x})\ \lambda^{\wedge I}$$

$$d\alpha = \Sigma_I(dT_I(\mathbf{x})) \wedge \lambda^{\wedge I} = \Sigma_I\ (\ \Sigma_{r=1}^n\ [\partial_r T_I(\mathbf{x})]\ \lambda^r) \wedge \lambda^{\wedge I} = \Sigma_I\Sigma_{r=1}^n\ [\partial_r T_I(\mathbf{x})]\ (\lambda^r \wedge \lambda^{\wedge I})$$

$$d(d\alpha) = \Sigma_I\ \Sigma_{r=1}^n\ d[\partial_r T_I(\mathbf{x})]\ (\lambda^r \wedge \lambda^{\wedge I})$$

$$= \Sigma_I\ \Sigma_{r=1}^n\ (\Sigma_{s=1}^n \partial_s[\partial_r T_I(\mathbf{x})]\ \lambda^s\ ) \wedge (\lambda^r \wedge \lambda^{\wedge I})$$

$$= \Sigma_I\ \Sigma_{r=1}^n\ \Sigma_{s=1}^n[\ \partial_s\partial_r T_I(\mathbf{x})]\ (\lambda^s \wedge \lambda^r \wedge \lambda^{\wedge I})$$

$$= 0\ .$$

QED

The result is 0 because in the symmetric sum Σ_{rs} the object $\partial_s\partial_r F_I(\mathbf{x})$ is symmetric under r↔ s while the object $(\lambda^s \wedge \lambda^r \wedge \lambda^{\wedge I})$ is antisymmetric under r↔s. That is to say, if S is symmetric and A antisymmetric,

$$\text{sum} = \Sigma_{rs}\ S_{rs}A_{rs} \overset{\text{swap names } r\leftrightarrow s}{=} \Sigma_{sr}\ S_{sr}A_{sr} \overset{\text{use symmetries}}{=} \Sigma_{rs}\ (+S_{rs})(-A_{rs}) = -\ \Sigma_{rs}\ S_{rs}A_{rs} = -\ \text{sum} = 0 \qquad (10.3.11)$$

Expressing dα in standard form

Recall from above that

$$\alpha = \Sigma'_I\ f_I(\mathbf{x})\ \lambda^{\wedge I} \qquad \text{k-form}$$

$$d\alpha = \Sigma'_I\ (\ \Sigma_{j=1}^n\ [\partial_j f_I(\mathbf{x})]\ \lambda^j) \wedge \lambda^{\wedge I}\ . \qquad \text{(k+1)-form} \qquad (10.3.6)$$

We wish to rewrite dα in a more standardized form. To this end, starting with the k-multiindex I we create a (k+1)-multiindex J as follows

$I = i_1,i_2.....i_k$

$J = j_1,j_2.....j_k,j_{k+1} \equiv i_1,i_2.....i_k, j$ // $j = j_{k+1}$

$J' = j_1,j_2.....j_k = I = i_1,i_2.....i_k$. (10.3.12)

Then one can rewrite the summation appearing in dα above as

$$\Sigma'_I \Sigma_j = \Sigma'_{J'} \Sigma_{j_{k+1}} = \Sigma_{j_1<j_2<...<j_k} \Sigma_{j_{k+1}} . \quad (10.3.13)$$

Then

$$d\alpha = \Sigma_{j_1<j_2<...<j_k} \Sigma_{j_{k+1}} [\partial_{j_{k+1}} f_{j_1j_2...j_k}(\mathbf{x})] \ \lambda^{j_{k+1}} \wedge \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_k} . \quad (10.3.14)$$

Because each swap of vectors in a wedge product of same creates a minus sign,

$$\lambda^{j_{k+1}} \wedge \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_k} = (-1)^k \ \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_k} \wedge \lambda^{j_{k+1}}$$

$$= (-1)^k \ \lambda^{\wedge J} \quad (10.3.15)$$

and then

$$d\alpha = (-1)^k \ \Sigma_{j_1<j_2<...<j_k} \Sigma_{j_{k+1}} [\partial_{j_{k+1}} f_{j_1j_2...j_k}(\mathbf{x})] \ \lambda^{\wedge J} . \quad (10.3.16)$$

The summations appearing above can be written as

$$\Sigma_{j_1<j_2<...<j_k} \Sigma_{j_{k+1}}$$
$$= \Sigma_{j_1<j_2...<j_k<j_{k+1}} + \Sigma_{j_1<j_2...<j_{k+1}<j_k} + ... + \Sigma_{j_{k+1}<j_1<j_2...<j_k} . \quad (10.3.17)$$

Here we are just exhausting all possible locations that j_{k+1} can take relative to the other indices. We don't have to worry about cases like $j_{k+1} = j_2$ because in that case $\lambda^{\wedge J} = 0$ and there is no contribution to the sum (10.3.16). One can then write,

$$(-1)^k d\alpha = \Sigma_{j_1<j_2...<j_{k-1}<j_k<j_{k+1}} [\partial_{j_{k+1}} f_{j_1j_2...j_k}] \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_{k+1}}$$

$$+ \ \Sigma_{j_1<j_2...<j_{k-1}<j_{k+1}<j_k} [\partial_{j_{k+1}} f_{j_1j_2...j_k}] \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_{k+1}}$$

$$+ \ \Sigma_{j_1<j_2...<j_{k+1}<j_{k-1}<j_k} [\partial_{j_{k+1}} f_{j_1j_2...j_k}] \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_{k+1}}$$

.....

$$+ \ \Sigma_{j_{k+1}j_1<j_2...<j_{k-1}<j_k} [\partial_{j_{k+1}} f_{j_1j_2...j_k}] \lambda^{j_1} \wedge \lambda^{j_2} \ .... \wedge \lambda^{j_{k+1}} . \quad (10.3.18)$$

Next, define the following index subscript swap operator S(r,s),

$$S(r,s) \ F_{j_1j_2.\, j_r...j_s...j_k} = F_{j_1j_2.\, j_s...j_r...j_k} . \quad (10.3.19)$$

In each term in (10.3.18) all the summation indices are of course dummy indices and their names can be swapped around at will. Notice that :

$$S(k,k+1)[\Sigma_{j_1<j_2\cdots<j_{k-1}<j_{k+1}<j_k}] = \Sigma_{j_1<j_2\cdots<j_{k-1}<j_k<j_{k+1}}$$

$$S(k,k+1)S(k-1,k+1)[\Sigma_{j_1<j_2\cdots<j_{k+1}<j_{k-1}<j_k}] = S(k,k+1)[\Sigma_{j_1<j_2\cdots<j_{k-1}<j_{k+1}<j_k}]$$

$$= \Sigma_{j_1<j_2\cdots<j_{k-1}<j_k<j_{k+1}}$$

$$S(k,k+1)S(k-1,k+1)S(k-2,k+1)\ [\ \Sigma_{j_1<j_2\cdots<j_{k+1}<j_{k-2}<j_{k-1}<j_k}]$$

$$= S(k,k+1)S(k-1,k+1)\ \ [\ \Sigma_{j_1<j_2\cdots<j_{k-2}<j_{k+1}<j_{k-1}<j_k}]$$

$$= S(k,k+1)\ \ [\ \Sigma_{j_1<j_2\cdots<j_{k-2}<j_{k-1}<j_{k+1}<j_k}] = \Sigma_{j_1<j_2\cdots<j_{k-1}<j_k<j_{k+1}}$$

.....

$$S(k,k+1)S(k-1,k+1)....S(1,k+1)[\Sigma_{j_{k+1}j_1<j_2\cdots<j_{k-1}<j_k}] = \Sigma_{j_1<j_2\cdots<j_{k-1}<j_k<j_{k+1}} \qquad (10.3.20)$$

Thus these swap combinations convert each summation to the standard form shown in the first line of (10.3.18).

So the next step is to apply the swap combinations not just to the summations, but to the entire lines shown in (10.3.18), since one is allowed to do this without changing each line's value since these are just dummy index swaps. The first effect of doing this is that all the summations become that shown on the first line which is just Σ'_J. The second effect is that the λ wedge products can be restored to their first-line ordering by adding a minus sign for each swap. For example,

$$S(k,k+1)S(k-1,k+1)\ \ \lambda^{j_1} \wedge \lambda^{j_2}\ .... \wedge \lambda^{j_{k+1}} = (-1)^2\ \lambda^{j_1} \wedge \lambda^{j_2}\ .... \wedge \lambda^{j_{k+1}}$$

or

$$S(k,k+1)S(k-1,k+1)\ \ \lambda_\wedge{}^J = (-1)^2\, \lambda_\wedge{}^J\ . \qquad (10.3.21)$$

Since on each line going down the number of swaps increases by 1, we pick up alternating signs.

Doing this, one can rewrite (10.3.18) as

$$(-1)^k d\alpha\ =\ \Sigma'_J\,[\partial_{j_{k+1}} f_{j_1 j_2\cdots j_k}]\, \lambda_\wedge{}^J$$

$$-\ \ \Sigma'_J\,[S(k,k+1)\{\partial_{j_{k+1}} f_{j_1 j_2\cdots j_k}\}]\, \lambda_\wedge{}^J$$

$$+\ \ \Sigma'_J\,[S(k,k+1)S(k-1,k+1)\{\partial_{j_{k+1}} f_{j_1 j_2\cdots j_k}\}]\, \lambda_\wedge{}^J$$

.....

$$+\,(-1)^k\ \ \Sigma'_J\,[S(k,k+1)S(k-1,k+1)....S(1,k+1)\{\partial_{j_{k+1}} f_{j_1 j_2\cdots j_k}\}]\, \lambda_\wedge{}^J\quad . \qquad (10.3.22)$$

Here then is the way to write the **derivative of a k-form in standard form**:

$$\alpha = \Sigma'_I \ f_I(\mathbf{x})\, \lambda\hat{}^{I} \qquad \text{k-form}$$

$$d\alpha = (-1)^k \ \Sigma'_J \, Q_J(\mathbf{x}) \ \lambda\hat{}^{J} \qquad \text{(k+1)-form} \qquad (10.3.23)$$

$$Q_J = [\ 1 - S(k,k+1) + S(k,k+1)S(k-1,k+1) - ... +(-1)^k S(k,k+1)S(k-1,k+1)...S(1,k+1)]\ \partial_{j_{k+1}} f_{j_1 j_2 \cdots j_k} \ .$$

The general result is admittedly unwieldy and perhaps has some more pleasant form, but we take it as is and consider some simple examples.

<u>Example 1:</u> Exterior derivative of a 1-form:

$$\alpha = \Sigma'_I \ f_I(\mathbf{x})\, \lambda\hat{}^{I} \ = \Sigma_{i_1} f_{i_1}(\mathbf{x})\, \lambda^{i_1}$$

$$Q_J = \ [\ 1 - S(1,2)\ ]\ \partial_{j_2} f_{j_1} \ = \ (\partial_{j_2} f_{j_1} - \partial_{j_1} f_{j_2})$$

$$d\alpha = (-1)^k \ \Sigma'_J \, Q_J \lambda\hat{}^{J} = - \Sigma_{j_1<j_2} (\partial_{j_2} f_{j_1} - \partial_{j_1} f_{j_2})\, \lambda^{j_1} \wedge \lambda^{j_2}$$

$$= \Sigma_{j_1<j_2} (\partial_{j_1} f_{j_2} - \partial_{j_2} f_{j_1})\, \lambda^{j_1} \wedge \lambda^{j_2} \quad . \qquad (10.3.24a)$$

In cosmetic notation,

$$\alpha = \ \Sigma_{i_1} f_{i_1}(\mathbf{x})\, dx^{i_1}$$

$$d\alpha = \Sigma_{j_1<j_2} \ (\partial_{j_1} f_{j_2} - \partial_{j_2} f_{j_1})\, dx^{j_1} \wedge dx^{j_2} \qquad (10.3.24b)$$

$$d\alpha = \Sigma_{i<j} \ (\partial_i f_j - \partial_j f_i)\, dx^i \wedge dx^j$$

which agrees with Sjamaar p 21 (2-2).

<u>Example 2:</u> Exterior derivative of a 2-form:

$$\alpha = \Sigma'_I \ f_I(\mathbf{x})\, \lambda\hat{}^{I} \ = \Sigma_{i_1<i_2} f_{i_1 i_2}(\mathbf{x}) \ \lambda^{i_1} \wedge \lambda^{i_2}$$

$$Q_J = \ [\ 1 - S(2,3) + S(2,3)S(1,3)\ ]\ \partial_{j_3} f_{j_1 j_2}$$

$$S(2,3)\partial_{j_3} f_{j_1 j_2} = \ \partial_{j_2} f_{j_1 j_3}$$

$$S(2,3)S(1,3)\ \partial_{j_3} f_{j_1 j_2} = S(2,3)\partial_{j_1} f_{j_3 j_2} \ = \partial_{j_1} f_{j_2 j_3}$$

$$Q_J = \ \partial_{j_3} f_{j_1 j_2} - \ \partial_{j_2} f_{j_1 j_3} + \ \partial_{j_1} f_{j_2 j_3}$$

$d\alpha = (-1)^k \Sigma'_J Q_J\lambda^{\wedge J}$

$= + \Sigma_{j_1<j_2<j_3} (\partial_{j_1}f_{j_2j_3} - \partial_{j_2}f_{j_1j_3} + \partial_{j_3}f_{j_1j_2}) \lambda^{i_1} \wedge \lambda^{i_2} \wedge \lambda^{i_3}$. (10.3.25a)

In cosmetic notation,

$\alpha = \Sigma_{i_1<i_2} f_{i_1i_2}(x) \ dx^{i_1} \wedge dx^{i_2}$

$d\alpha = \Sigma_{j_1<j_2<j_3} (\partial_{j_1}f_{j_2j_3} - \partial_{j_2}f_{j_1j_3} + \partial_{j_3}f_{j_1j_2}) \ dx^{j_1} \wedge dx^{j_2} \wedge dx^{j_3}$ (10.3.25b)

$d\alpha = \Sigma_{i<j<k} (\partial_i f_{jk} - \partial_j f_{ik} + \partial_k f_{ij}) \ dx^i \wedge dx^j \wedge dx^k$

which agrees with Sjamaar p 21 (2-4).

Example 3: (last one!) Exterior derivative of a 3-form:

$\alpha = \Sigma'_I f_I(\mathbf{x}) \lambda^{\wedge I} = \Sigma_{i_1<i_2<i_3} f_{i_1i_2i_3}(x) \ \lambda^{i_1} \wedge \lambda^{i_2} \wedge \lambda^{i_3}$

$Q_J = [1 - S(3,4) + S(3,4)S(2,4) - S(3,4)S(2,4)S(1,4)] \partial_{j_4}f_{j_1j_2j_3}$

$S(3,4)\partial_{j_4}f_{j_1j_2j_3} = \partial_{j_3}f_{j_1j_2j_4}$

$S(3,4)S(2,4)\partial_{j_4}f_{j_1j_2j_3} = S(3,4)\partial_{j_2}f_{j_1j_4j_3} = \partial_{j_2}f_{j_1j_3j_4}$

$S(3,4)S(2,4)S(1,4)\partial_{j_4}f_{j_1j_2j_3} = S(3,4)S(2,4)\partial_{j_1}f_{j_4j_2j_3} = S(3,4)\partial_{j_1}f_{j_2j_4j_3} = \partial_{j_1}f_{j_2j_3j_4}$

$Q_J = \partial_{j_4}f_{j_1j_2j_3} - \partial_{j_3}f_{j_1j_2j_4} + \partial_{j_2}f_{j_1j_3j_4} - \partial_{j_1}f_{j_2j_3j_4}$

$d\alpha = (-1)^k \Sigma'_J Q_J\lambda^{\wedge J}$ (10.3.26a)

$= \Sigma_{j_1<j_2<j_3<j_4} (\partial_{j_1}f_{j_2j_3j_4} - \partial_{j_2}f_{j_1j_3j_4} + \partial_{j_3}f_{j_1j_2j_4} - \partial_{j_4}f_{j_1j_2j_3}) \lambda^{i_1} \wedge \lambda^{i_2} \wedge \lambda^{i_3} \wedge \lambda^{i_4}$.

In cosmetic notation,

$\alpha = \Sigma_{i_1<i_2<i_3} f_{i_1i_2i_3}(x) \ dx^{i_1} \wedge dx^{i_2} \wedge dx^{i_3}$ (10.3.26b)

$d\alpha = \Sigma_{j_1<j_2<j_3<j_4} (\partial_{j_1}f_{j_2j_3j_4} - \partial_{j_2}f_{j_1j_3j_4} + \partial_{j_3}f_{j_1j_2j_4} - \partial_{j_4}f_{j_1j_2j_3}) \ dx^{j_1} \wedge dx^{j_2} \wedge dx^{j_3} \wedge dx^{j_4}$

$d\alpha = \Sigma_{i<j<k<\ell} (\partial_i f_{jk\ell} - \partial_j f_{ik\ell} + \partial_k f_{ij\ell} - \partial_\ell f_{ijk}) \ dx^i \wedge dx^j \wedge dx^k \wedge dx^\ell$.

We leave it to the reader to deduce a "general rule by inspection" for the series of terms for any k. This might involve rotations of certain subsets of the subscripts.

Exterior Derivative of products of differential forms

Whereas d(fg) = (df)g+f(dg) in calculus, the result is slightly different if f and g are differential forms :

Fact: If α is a k-form and β is a k'-form, then $d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta)$. (10.3.27)

Proof: This is a "brute force proof". See Sjamaar p 22 for a denser proof which uses (10.4.1) below. The forms are assumed to exist in R^n so **x** has n coordinates (n > k+k') and $\Sigma_s = \Sigma_{s=1}^n$.

Consider, using (10.3.6),

$$\alpha = \Sigma'_I \ f_I(\mathbf{x})\, \lambda\wedge^I \quad // \text{ a k-form} \quad \Rightarrow \quad d\alpha = \Sigma'_I \ \Sigma_s \ [\partial_s f_I(\mathbf{x})]\, \lambda^s \wedge \lambda\wedge^I$$

$$\beta = \Sigma'_J \ g_J(\mathbf{x})\, \lambda\wedge^J \quad // \text{ a k'-form} \quad \Rightarrow \quad d\beta = \Sigma'_J \ \Sigma_s \ [\partial_s g_J(\mathbf{x})]\, \lambda^s \wedge \lambda\wedge^J$$

$$\alpha \wedge \beta = (\Sigma'_I \ f_I(\mathbf{x})\, \lambda\wedge^I) \wedge (\Sigma'_J \ g_J(\mathbf{x})\, \lambda\wedge^J) = \Sigma'_I \Sigma'_J \ f_I(\mathbf{x}) g_J(\mathbf{x})\, \lambda\wedge^I \wedge \lambda\wedge^J$$

$$\Rightarrow \ d(\alpha \wedge \beta) = \Sigma'_I \Sigma'_J \Sigma_s \partial_s [f_I(\mathbf{x}) g_J(\mathbf{x})]\, \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J .$$

Then just evaluate the right side of (10.3.27) :

$$(d\alpha) \wedge \beta = (\Sigma'_I \Sigma_s \ [\partial_s f_I(\mathbf{x})]\, \lambda^s \wedge \lambda\wedge^I) \wedge (\Sigma'_J \ g_J(\mathbf{x})\, \lambda\wedge^J)$$

$$= \Sigma'_I \Sigma'_J \Sigma_s \ [\partial_s f_I(\mathbf{x})] \ g_J(\mathbf{x})\, \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J$$

$$\alpha \wedge (d\beta) = (\Sigma'_I \ f_I(\mathbf{x})\, \lambda\wedge^I) \wedge (\Sigma'_J \ \Sigma_{s=1}^n \ [\partial_s g_J(\mathbf{x})]\, \lambda^s \wedge \lambda\wedge^J)$$

$$= \Sigma'_I \Sigma'_J \Sigma_s \ f_I(\mathbf{x})[\partial_s g_I(\mathbf{x})]\, \lambda^I \wedge \lambda\wedge^s \wedge \lambda\wedge^J$$

$$= \Sigma'_I \Sigma'_J \Sigma_s \ f_I(\mathbf{x})[\partial_s g_I(\mathbf{x})]\, (-1)^k \ \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J .$$

Here $\lambda^I \wedge \lambda\wedge^s = (-1)^k \ \lambda^s \wedge \lambda\wedge^I$ because λ^s has to slide left through k vector wedge products. Then

$$(d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta) = \Sigma'_I \Sigma'_J \Sigma_{s=1}^n \ [\partial_s f_I(\mathbf{x})] \ g_J(\mathbf{x})\, \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J$$
$$+ \ \Sigma'_I \Sigma'_J \Sigma_{s=1}^n \ f_I(\mathbf{x})[\partial_s g_I(\mathbf{x})] \ \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J$$

$$= \Sigma'_I \Sigma'_J \Sigma_{s=1}^n \ \{ [\partial_s f_I(\mathbf{x})] \ g_J(\mathbf{x}) + f_I(\mathbf{x}) \ [\partial_s g_I(\mathbf{x})] \ \} \ \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J$$

$$= \Sigma'_I \Sigma'_J \Sigma_{s=1}^n \ \partial_s [f_I(\mathbf{x}) g_J(\mathbf{x})] \ \lambda^s \wedge \lambda\wedge^I \wedge \lambda\wedge^J$$

$$= d(\alpha \wedge \beta) .$$ QED

Reader Exercises: (10.3.28)
(a) Show that d is a linear operator so $d(s_1\alpha + s_2\beta) = s_1 d\alpha + s_2 d\beta$ for any forms α and β .
(b) Use (10.4.1) below three times in (10.3.27) and show result is consistent with (10.3.27) for $d(\beta \wedge \alpha)$.
(c) Write an expression for $d(\alpha \wedge \beta \wedge \gamma)$ where α,β,γ are forms of rank k, k' and k".
(d) Write an expression for $d(\alpha_1 \wedge \alpha_2 \wedge ... \alpha_M)$ where α_i are k_i-forms.

10.4. Commutation properties of differential forms

Recall these three results from Chapter 8 concerning elements of $\Lambda(V)$,

- $\mathcal{S}^\wedge \wedge \mathcal{T}^\wedge = (-1)^{kk'} \mathcal{T}^\wedge \wedge \mathcal{S}^\wedge$ ranks of the two dual tensors are k and k' .   *(8.9.c.6)*

- In a product of tensors $(\mathcal{T}_1)^\wedge \wedge (\mathcal{T}_2)^\wedge \wedge (\mathcal{T}_3)^\wedge$....of rank k_1, k_2, k_3 ... , if two tensors are swapped $(\mathcal{T}_r)^\wedge \leftrightarrow (\mathcal{T}_s)^\wedge$ (with r < s), the resulting tensor incurs the following sign relative to the starting tensor,

  sign = $(-1)^m$ where $m = (k_{r+1}+k_{r+2} ...+k_{s-1})(k_r+k_s) + k_r k_s$ *(8.9.e.6)*

- $\mathcal{T}^{\wedge N} = 0$ for any $N \geq n+1$ assuming $k \neq 0$.   *(8.9.d.9)*

In the language of differential forms these three results become

- $\alpha \wedge \beta = (-1)^{kk'} \beta \wedge \alpha$ α = k-form, β = k'-form   (10.4.1)

- $\alpha_1 \wedge \alpha_2 \wedge ... \alpha_r ... \alpha_s ... \wedge \alpha_k = (-1)^m \alpha_1 \wedge \alpha_2 \wedge ... \alpha_s ... \alpha_r ... \wedge \alpha_k$

  where $m = (k_{r+1}+k_{r+2} ...+k_{s-1})(k_r+k_s) + k_r k_s$ (10.4.2)

- $\alpha^N = 0$ for $N \geq n+1$ dim(V) = n   α = any k-form with $k \geq 1$

  where $\alpha^N \equiv \alpha \wedge \alpha ... \wedge \alpha$.   (10.4.3)

Equations (10.4.1) and (10.4.3) appear in Sjamaar as "2.1 Proposition" and the preceding equation on page 19 . In Sjamaar, Buck and many other source all ^ symbols are suppressed so (10.4.1) is written $\alpha\beta = (-1)^{kk'}\beta\alpha$ and one must understand that these are wedge products in $\Lambda(V)$.

10.5. Closed and Exact, Poincaré and the Angle Form

Closed: If dα = 0 for a k-form α, α is said to be **closed**. The analogous fact for a function f(**x**) with df = 0 would be that f(**x**) = constant.   (10.5.1)

Exact: Sometimes one finds that a form α can be written α = dβ where β is some other form. If α is a k-form, we know from (10.3.7) that β must be a (k-1)-form. When α = dβ for some form β, α is said to be **exact**. We showed in (10.3.10) that $d^2\beta = 0$ for any form β, so it follows that if α = dβ, then dα = 0 and α is closed. Thus we have shown that :   (10.5.2)

Fact: If α is exact, then α is closed.   (10.5.3)

In 1D calculus if f = dh/dx one says that dh = f dx is an "exact (perfect) differential" and one then writes

$$\int_a^b f(x)\,dx = \int_a^b (\frac{dh}{dx})\,dx = \int_a^b dh = h(a) - h(b) \qquad dh = (\frac{dh}{dx})\,dx . \qquad (10.5.4)$$

In nD calculus if $\mathbf{f} = \nabla h$ one says that $dh = \nabla h \bullet d\mathbf{x}$ is an exact (perfect) differential. The above integral then becomes a line integral over a smooth curve C having endpoints **a** and **b**,

$$\int_a^b \mathbf{f}(\mathbf{x}) \bullet d\mathbf{x} = \int_a^b \nabla h \bullet d\mathbf{x} = \int_C dh = h(\mathbf{b}) - h(\mathbf{a})$$

where

$$dh = \nabla h \bullet d\mathbf{x} = \Sigma_{i=1}^n (\partial_i h(\mathbf{x}))dx^i = \Sigma_{i=1}^n f_i(\mathbf{x})\,dx^i = \mathbf{f}(\mathbf{x}) \bullet d\mathbf{x} . \qquad (10.5.5)$$

The line integral depends only on the line endpoints **a** and **b**, and not on the particular shape of the curve C joining **a** and **b**. For a closed curve **a** = **b** and one finds

$$\int_C dh = h(\mathbf{b}) - h(\mathbf{a})$$

$$\oint dh = h(\mathbf{a}) - h(\mathbf{a}) = 0 . \qquad (10.5.6)$$

In physics if $\mathbf{f}(\mathbf{x})$ is a "conservative force field" (like gravity) then $h(\mathbf{a}) - h(\mathbf{a}) = 0$ is the work done in moving a particle that senses the field (has mass) around a closed path.

A similar theorem exists for $\alpha = dg$ where g is a 0-form (a function) and α is 1-form. Here we provide a preview of things to come. C' is a curve in x'-space running from point **a'** to point **b'**, while C is the pulled-back curve in x-space running from **a** to **b**, where $\mathbf{a}' = \mathbf{F}(\mathbf{a})$ and $\mathbf{b}' = \mathbf{F}(\mathbf{b})$:

$$\int_{C'} \alpha_{\mathbf{x}'} = \int_{C'} dg(\mathbf{x}') \qquad // \alpha_{\mathbf{x}'} = dg \text{ so } \alpha_{\mathbf{x}'} \text{ is an exact 1-form (g is a function)}$$

$$= \int_C F^*(dg) \qquad // \text{ pullback of a 1-form, (10.11.2) with } \beta_{\mathbf{x}} = F^*(dg)$$

$$= \int_C d[F^*(g(\mathbf{x}'))] \qquad // \text{ fact (10.7.22) that d commutes with } F^*$$

$$= \int_C d[g(\mathbf{F}(\mathbf{x})] \qquad // \text{ fact (10.7.19) item 1 (pullback of a function) that } F^*(f(\mathbf{x}')) = f(\mathbf{F}(\mathbf{x}))$$

$$= g(\mathbf{F}(\mathbf{b}))-g(\mathbf{F}(\mathbf{a})) \qquad // \text{ think of } g(\mathbf{F}(\mathbf{x})) \text{ as } h(\mathbf{x}) \text{ so } d[g(\mathbf{F}(\mathbf{x})] = dh$$

$$= g(\mathbf{b}')-g(\mathbf{a}') . \qquad (10.5.7)$$

Then for a closed curve C' the line integral of an exact 1-form vanishes,

$$\oint_{C'} \alpha_{\mathbf{x}'} = g(\mathbf{a}') - g(\mathbf{a}') = 0 \qquad (10.5.8)$$

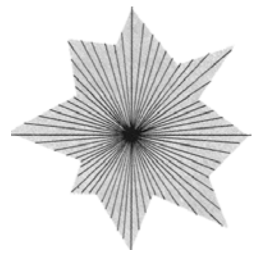
in analogy with (10.5.6). A 1-form α being exact is like dh being an exact differential.

Fact (10.5.3) above says α exact ⇒ α closed. Is it possibly also true that α closed ⇒ α exact and so then the two descriptions are one in the same? The answer is "not quite" as expressed in this claim:

Poincaré Lemma: If any differential form α on R^n is closed for **x** in some open star-shaped domain in R^n which includes the origin, then α is exact. (Poincare for PDF search) (10.5.9)

This Lemma appears on p 94 of Spivak from which we quote,

4-11 Theorem (Poincaré Lemma). *If $A \subset \mathbf{R}^n$ is an open set star-shaped with respect to 0, then every closed form on A is exact.*



and Spivak proceeds to give a detailed proof. In topological language, the star-shaped domain is any domain that is "contractible to a point". Certainly the Lemma is valid for a domain which is an open "cube" or "sphere" (n dimensions) about the origin. The domain need not be convex (as the star shows).

A classic application of this theorem involves the so-called **angle form** defined on R^2 with coordinates (x_1,x_2),

$$\alpha = \Sigma_{i=1}^2 f_i(\mathbf{x})\lambda^i \qquad \text{where} \quad f_1(\mathbf{x}) = -(x_2/r^2) \qquad r^2 = x_1^2 + x_2^2 \ .$$
$$f_2(\mathbf{x}) = \ (x_1/r^2) \qquad (10.5.10)$$

Then

$$d\alpha = \Sigma_i \, df_i(\mathbf{x})\lambda^i = \Sigma_{ij} (\partial_j f_i) \lambda^j \wedge \lambda^i \ .$$

Notice that, using the fact that $\partial_i r = x_i/r$,

$$(\partial_1 f_2) = \partial_1(x_1/r^2) = [r^2 * 1 - x_1 (\partial_1 r^2)] / r^4 = [r^2 - x_1 2r (\partial_1 r)] / r^4 = - [r^2 - x_1 2r(x_1/r)] / r^4$$
$$= [r^2 - 2x_1^2] / r^4 = [x_1^2 + x_2^2 - 2x_1^2] / r^4 = (x_2^2 - x_1^2) / r^4$$

and

$$(\partial_2 f_1) = - \partial_2 (x_2/r^2) = - [r^2 * 1 - x_2 (\partial_2 r^2)] / r^4 = - [r^2 - x_2 2r (\partial_2 r)] / r^4 = - [r^2 - x_2 2r(x_2/r)] / r^4$$
$$= - [r^2 - 2x_2^2] / r^4 = - [x_1^2 + x_2^2 - 2x_2^2] / r^4 = (x_2^2 - x_1^2) / r^4 = (\partial_1 f_2) \ .$$

Thus it turns out that the quantity $(\partial_j f_i)$ is symmetric under i ↔ j. Then by the argument (10.3.11) we get

$$d\alpha = \Sigma_{ij} (\partial_j f_i) \lambda^j \wedge \lambda^i = \Sigma_{ij} (S_{ij})(A^{ji}) = 0 \qquad \Rightarrow \ \alpha = \text{closed}$$

so α is a closed 2-form. As we shall show below in (10.12.21), the line integral of α around a circle centered at the origin gives $\oint \alpha = 2\pi$. Thus the angle form is not exact because if it were one would have $\oint \alpha = 0$ as in (10.5.8). So here is a form α which is closed, but which is not exact. The condition of the

Poincaré Lemma must therefore be violated, and that is indeed the case since the form α is undefined for r = 0 where f_1 and f_2 blow up, so α is then defined only on R^2 punctured at the origin, sometimes written $R^2/\{0\}$ or $R^2 - \{0\}$. Thus we can't have any open star-shaped domain including the origin for α, so Poincaré's Lemma does not apply. Note that $R^2 - \{0\}$ is not "simply connected" due to the puncture hole. The presence of holes ("multiply connected") means that line integrals are no longer path independent. Here a line integral around the hole gives 2π, whereas one not looping the hole gives 0.

Our plan now is first to define the "pullback" of a differential form, and then in later sections to use the pullback to define the meaning of integration of a differential form over a manifold. But we wish to show how the notion of a pullback fits into the general transformation scenario of Chapter 2, and this requires several digressions before we get to the pullback discussion in Section 10.7 through 10.9.

10.6 Transformation Kinematics

Much mathematical hardware accompanies a mapping. In mechanics, the selection of an appropriate set of coordinates and corresponding basis vectors is sometimes referred to as stating the *kinematics* of a problem (as opposed to the dynamics which involves equations of motion). Here we apply this term loosely to the cloud of equations associated with a mapping. Not all these equations will be used in our analysis, but we like being able to see them all in one place just in case something is needed.

In the following Sections we shall move in and out of the Dirac notation of Section 2.11 in a somewhat repetitive fashion intended to make the reader more comfortable with that notation.

The notion of a pullback is often presented as "something new", but the main point of the following sections is to show that the pullback operator is just the R/$\mathcal{R}$ matrix/operator of the underlying transformation.

In Chapter 2 we discussed the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ from x-space to x'-space using Picture A (2.1.1). The vector transformation and "the differential" (the R-matrix) of the transformation were given by

$$V'^a = R^a{}_b V^b \qquad R^a{}_b \equiv (\partial x'^a/\partial x^b) = \partial_b x'^a \equiv (\nabla\mathbf{F})^a{}_b \equiv (\mathbf{DF})^a{}_b \equiv (DF)^a{}_b \qquad \textit{(2.1.2)}$$

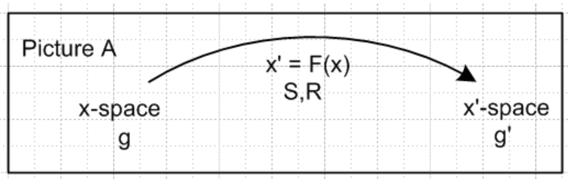
$$dx'^a = R^a{}_b dx^b \qquad d\mathbf{x}' = R d\mathbf{x} \ . \qquad \textit{(2.1.12)} \qquad (10.6.2)$$

Here $V'^a = R^a{}_b V^b$ shows the transformation of a contravariant vector under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. In matrix notation one would write $\mathbf{V}' = R\mathbf{V}$. Repeated indices are always summed unless otherwise stated.

Above we have defined $\nabla\mathbf{F}$ and $\mathbf{DF}$ as alternate names for matrix R because many authors (like Spivak) use this notation. In *Tensor* (E.4.4) we show that this is in fact a "reverse dyadic notation". Often $(\mathbf{DF})^a{}_b$ is written unbolded $(DF)^a{}_b$ so then R = (DF) with the idea that a matrix like R is normally not bolded.

(a) Axis-Aligned Vectors and Tangent Base Vectors : The Kinematics Package

We gather here various facts derived in Chapter 2 which comprise our "kinematics package" for the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. We cosmetically flip Picture A of (2.1.1) left to right.



R^n $\qquad$ R^m $\qquad$ $m \geq n$

(a) $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ transformation $\quad R^i{}_j \equiv (\partial x'^i/\partial x^j) = \partial_j x'^i$ $\quad R = (DF)$
$\mathbf{V}' = R\,\mathbf{V}$ vector $\quad S^i{}_j \equiv (\partial x^i/\partial x'^j) = \partial'_j x^i$ *(2.1.2)*

(b) $\mathbf{e}'_i$ with $(\mathbf{e}'_i)^j = \delta_i{}^j$ — axis-aligned basis vectors in x'-space (i = 1..m) *(2.5.2)*
$\mathbf{e}_i$ $\quad \mathbf{e}_i = S\mathbf{e}'_i$ — tangent base vectors in x-space (i = 1..n) *(2.5.1)*

(c) $\mathbf{u}_i$ with $(\mathbf{u}_i)^j = \delta_i{}^j$ — axis-aligned basis vectors in x-space (i = 1..n) *(2.4.1)*
$\mathbf{u}'_i$ $\quad \mathbf{u}'_i = R\mathbf{u}_i$ — tangent base vectors in x'-space (i = 1..n) *(2.5.1)*
$(\mathbf{u}'_i)^j = R^j{}_k (\mathbf{u}_i)^k$

(2.11.h.8)
(d) $1' = |\mathbf{e}'_i\rangle\langle\mathbf{e}'^i| = |\mathbf{e}'^i\rangle\langle\mathbf{e}'_i| = |\mathbf{u}'_i\rangle\langle\mathbf{u}'^i| = |\mathbf{u}'^i\rangle\langle\mathbf{u}'_i|$ — completeness in x'-space
$1 = |\mathbf{e}_i\rangle\langle\mathbf{e}^i| = |\mathbf{e}^i\rangle\langle\mathbf{e}_i| = |\mathbf{u}_i\rangle\langle\mathbf{u}^i| = |\mathbf{u}^i\rangle\langle\mathbf{u}_i|$ — completeness in x-space

(e) $(\mathbf{u}_j)^i = \mathbf{u}^i \bullet \mathbf{u}_j = \langle\mathbf{u}^i|\mathbf{u}_j\rangle = g^i{}_j = \mathbf{u}'^i \bullet \mathbf{u}'_j = \langle\mathbf{u}'^i|\mathbf{u}'_j\rangle$
$(\mathbf{e}_j)^i = \mathbf{u}^i \bullet \mathbf{e}_j = \langle\mathbf{u}^i|\mathbf{e}_j\rangle = S^i{}_j = R_j{}^i$
$(\mathbf{e}'_j)^i = \mathbf{e}'^i \bullet \mathbf{e}'_j = \langle\mathbf{e}'^i|\mathbf{e}'_j\rangle = g'^i{}_j = \mathbf{e}^i \bullet \mathbf{e}_j = \langle\mathbf{e}^i|\mathbf{e}_j\rangle$
$(\mathbf{u}'_j)^i = \mathbf{e}'^i \bullet \mathbf{u}'_j = \langle\mathbf{e}'^i|\mathbf{u}'_j\rangle = R^i{}_j = S_j{}^i$ *(2.5.8)*

(f) $\mathbf{e}^i = g'^{ij}\,\mathbf{e}_j$ $\quad \mathbf{e}'^i = g'^{ij}\,\mathbf{e}'_j$ $\quad \mathbf{u}^i = g^{ij}\,\mathbf{u}_j$ $\quad \mathbf{u}'^i = g^{ij}\,\mathbf{u}'_j$ *(2.3.2),(2.4.2),(2.5.6)*
$\mathbf{e}_i = g'_{ij}\,\mathbf{e}^j$ $\quad \mathbf{e}'_i = g'_{ij}\,\mathbf{e}'^j$ $\quad \mathbf{u}_i = g_{ij}\,\mathbf{u}^j$ $\quad \mathbf{u}'_i = g_{ij}\,\mathbf{u}'^j$

(g) $\langle\mathbf{e}_j|\mathcal{S}|\mathbf{e}'^i\rangle = \langle\mathbf{e}'^i|\mathcal{R}|\mathbf{e}_j\rangle = g'^i{}_j$ $\quad$ // $\langle\mathbf{e}'^i|R\mathbf{e}_j\rangle = \langle\mathbf{e}'^i|\mathbf{e}'_j\rangle = g'^i{}_j$ from (e)
$\langle\mathbf{e}_j|\mathcal{S}|\mathbf{u}'^i\rangle = \langle\mathbf{u}'^i|\mathcal{R}|\mathbf{e}_j\rangle = S^i{}_j = R_j{}^i$ $\quad$ // $\langle\mathbf{u}'^i|R\mathbf{e}_j\rangle = \langle\mathbf{u}'^i|\mathbf{e}'_j\rangle = R_j{}^i$ from (e)
$\langle\mathbf{u}_j|\mathcal{S}|\mathbf{e}'^i\rangle = \langle\mathbf{e}'^i|\mathcal{R}|\mathbf{u}_j\rangle = R^i{}_j = S_j{}^i$ $\quad$ // $\langle\mathbf{e}'^i|R\mathbf{u}_j\rangle = \langle\mathbf{e}'^i|\mathbf{u}'_j\rangle = R^i{}_j$ from (e)
$\langle\mathbf{u}_j|\mathcal{S}|\mathbf{u}'^i\rangle = \langle\mathbf{u}'^i|\mathcal{R}|\mathbf{u}_j\rangle = g^i{}_j\,.$ $\quad$ // $\langle\mathbf{u}'^i|R\mathbf{u}_j\rangle = \langle\mathbf{u}'^i|\mathbf{u}'_j\rangle = g^i{}_j$ from (e)

(h) $S = R^T$ $\quad \mathcal{S} = \mathcal{R}^T$ $\quad S^i{}_j = (R^T)^i{}_j = R_j{}^i$ *(2.11.f3), (2.11.f1)*
$R = S^T$ $\quad \mathcal{R} = \mathcal{S}^T$ $\quad R^i{}_j = (S^T)^i{}_j = S_j{}^i$

(i) $S = R^{-1}$ $\quad R = S^{-1}$ $\quad RS = SR = 1$ *(2.11.f3)*
$RR^T = R^TR = SS^T = S^TS = 1.$ (10.6.a.1)

In item (g) one has in general $\langle\mathbf{a}|\mathcal{S}|\mathbf{b}\rangle = \langle\mathbf{a}|\mathcal{R}^T|\mathbf{b}\rangle = \langle\mathbf{b}|\mathcal{R}|\mathbf{a}\rangle$ as shown in (2.11.g.11) .

In any equation above, any index or label can be raised or lowered on both sides. The object $g^i{}_j$ is the tensor-correct form of $g^i{}_j = \delta^i{}_j = \delta_{i,j}$, allowing for indices to be raised and lowered, see (2.2.2). Here is a sample Dirac notation manipulation using the above information (implied sum in completeness),

$$|\mathbf{e}^i\rangle \;=\; [\mathit{1}]\,|\mathbf{e}^i\rangle = |\,\mathbf{u}^j\rangle\langle\mathbf{u}_j|\mathbf{e}^i\rangle \;=\; |\,\mathbf{u}^j\rangle R^i{}_j \;=\; R^i{}_j\,|\mathbf{u}^j\rangle \qquad \text{or} \qquad \mathbf{e}^i = R^i{}_j\mathbf{u}^j\ . \qquad (10.6.a.2)$$

The result $\mathbf{e}^i = R^i{}_j\mathbf{u}^j$ appears in (2.4.4) showing that $R^i{}_j$ is the basis change matrix between these two sets of basis vectors. Notice that an equation like $\mathbf{e}^i = \Sigma_{j=1}{}^n\ R^i{}_j\mathbf{u}^j$ is a "vector sum equation" since it has a sum of vectors on the right side. No component indices appear on the vectors in this equation (i and j are labels).

As discussed in Section 2.11 (g), abstract **operators in the Dirac space** will be written in script font. The operator $\mathcal{R}$ for example is completely determined by all its matrix elements $\langle\mathbf{e}'^i|\ \mathcal{R}\ |\ \mathbf{u}_j\rangle\ =\ R^i{}_j$. The identity operator in a Dirac space we then write as *1* for x-space and *1'* for x'-space, as appear in the completeness statements of (10.6.a.1) item (d).

(b) What happens for a non-square tall R matrix?

In Chapter 2 and in *Tensor* it was assumed that $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ was an invertible mapping F: $R^N \rightarrow R^N$. Now however we wish to consider the non-invertible mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ where

$$\begin{aligned} &F:\quad R^n \;\rightarrow\; R^m && m > n \\ &F:\ \text{x-space} \rightarrow \text{x'-space} && \mathbf{x} \in R^n, \qquad \mathbf{x}' \in R^m \qquad \mathbf{F}(\mathbf{x}) = \mathbf{x}'\ . \end{aligned} \qquad (10.6.b.1)$$

In $R^a{}_b = (\partial x'^a/\partial x^b)$ the row index a ranges 1 to m, while column index b ranges 1 to n. Thus the down-tilt R matrix is a "tall" non-square matrix having m rows and n columns with m > n.

As outlined in Section 10.2 and Fig (10.2.1). if we let the variable **x** exhaust some domain U within x-space, the mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ generates a "surface" embedded within x'-space = R^m which has dimension n. We assume that the mapping **F** has appropriate properties so that this surface is a Manifold M.

Thus, the mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is defined in effect for all **x** in R^n (or perhaps for a region U in R^n as in Fig (10.2.1), and produces (as its image) the manifold M within R^m . The inverse mapping $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ is then only defined for points **x'** on the manifold M. For such points, the mapping and its inverse are assumed one-to-one. This inverse mapping is a set of n equations which one can presumably write down. The equations represent $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ only when **x'** lies on M. For other values of **x'**, the set of equations still exists but no longer represents the inverse function $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. This point is hopefully clarified by some Examples.

Example 1: Let U be a square in R^2 x-space with corners (-1,-1) to (1,1). We map this square into R^3 using the following map $\mathbf{x}' = \mathbf{F}(\mathbf{x})$:

$$\begin{aligned} x'^1 &= x^1 \\ x'^2 &= x^2 \\ x'^3 &= \sqrt{2^2 - (x^1)^2 - (x^2)^2} \end{aligned} \qquad \mathbf{x}' = \mathbf{F}(\mathbf{x}) \qquad (10.6.b.2)$$

The image in R^3 x'-space is a partial upper hemispherical surface of radius 2 (see below). What is the inverse mapping $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$? One can take it to be the first two lines above,

$$x^1 = x'^1$$
$$x^2 = x'^2 \qquad\qquad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \qquad\qquad (10.6.b.3)$$

but the inverse mapping only applies to points **x'** on the hemisphere. The above two equations of course *exist* for points **x'** not on the hemisphere, but they only act as the inverse mapping for points on the hemisphere.

Here is Maple code for Example 1. The transformation is first entered and plotted, xp = x' :

```
restart; r := 2;
```

$$r := 2$$

```
xp[1] := x[1];
```

$$xp_1 := x_1$$

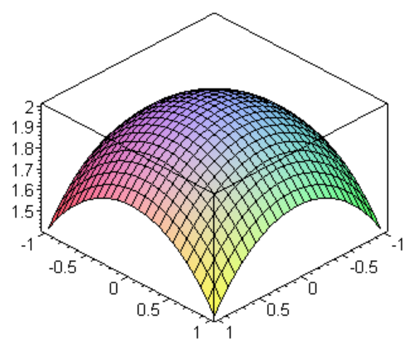
```
xp[2] := x[2];
```

$$xp_2 := x_2$$

```
xp[3] := sqrt(r^2-xp[1]^2 - xp[2]^2);
```

$$xp_3 := \sqrt{4 - x_1^2 - x_2^2}$$

```
plot3d([xp[1],xp[2],xp[3]],x[1]=-1..1,x[2]=-1..1,axes = boxed);
```



Maple then computes the "tall" R matrix, $R^i{}_j \equiv (\partial x'^i/\partial x^j)$,

```
R_ := (i,j) -> diff(xp[i],x[j]);
```

$$R\_ := (i,j) \rightarrow \frac{\partial}{\partial x_j} xp_i$$

```
R := matrix(3,2,R_);
```

$$R := \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{x_1}{\sqrt{4 - x_1^2 - x_2^2}} & -\frac{x_2}{\sqrt{4 - x_1^2 - x_2^2}} \end{bmatrix}.$$

The S matrix $S^i{}_j = (\partial x^i/\partial x'^j)$ is computed by hand from (10.6.b.3) and is then entered into Maple. Maple then computes the matrix products RS and SR,

```
S := matrix(2,3,[1,0,0,0,1,0]);
```

$$S := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

```
evalm(R &* S);
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{x_1}{\sqrt{4-x_1^2-x_2^2}} & -\frac{x_2}{\sqrt{4-x_1^2-x_2^2}} & 0 \end{bmatrix}$$

```
evalm(S &* R);
```

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Notice that RS ≠ 1 while SR = 1.

<u>Example 2</u>: Let U be the same square as in Example 1, but the new mapping is this

$$\begin{aligned} x'^1 &= x^1 + 2x^2 && 1 \\ x'^2 &= 2x^1 + x^2 && 2 \\ x'^3 &= x^1 + 3x^2 && 3 \qquad \mathbf{x}' = \mathbf{F}(\mathbf{x}) \quad . \end{aligned} \tag{10.6.b.4}$$

The image in R^3 x'-space is a tilted plane passing through the origin. We reuse the above Maple code for this example, but don't display the Maple output.

What is the inverse mapping $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$?

If one solves the first two equations for x^1 and x^2 the result is

$$\begin{aligned} x^1 &= -1/3\ x'^1 + 2/3\ x'^2 \\ x^2 &= \ \ 2/3\ x'^1 - 1/3\ x'^2 \qquad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \end{aligned} \tag{10.6.b.5}$$

and this then can be taken to be the inverse mapping $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. Inserting these expressions into the third equation gives

$$5/3\ x'^1 - 1/3\ x'^2 - x'^3 = 0 \tag{10.6.b.6}$$

which is the equation of the tilted image plane passing through the origin whose normal is (5/3,-1/3,-1).

On the other hand, if one instead solves the second two equations in (10.6.b.4) one finds

$$\begin{aligned} x^1 &= 3/5\ x'^2 - 1/5\ x'^3 \\ x^2 &= -1/5\ x'^2 + 2/5\ x'^3 \qquad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \ . \end{aligned} \tag{10.6.b.7}$$

Notice that this inverse mapping is *different* from (10.6.b.5). When these two expressions are inserted into the first equation of (10.6.b.4), one gets

$$x'^1 - 1/5\, x'^2 - 3/5\, x'^3 = 0 \qquad (10.6.b.8)$$

Multiplication by 5/3 gives (10.6.b.6) so this is, of course, the equation for the same tilted plane.

In this Example we find that the inverse equation set $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ is not unique. If we work with the first and third equations in (10.6.b.4) we get a third set of inverse equations which we leave to the reader.

By visual inspection, the R matrix computed from $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ (10.6.b.4) is this:

$$R = R^a{}_b = (\partial x'^a/\partial x^b) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix} \qquad (10.6.b.9)$$

and is the "tall" R matrix for this example. For the two inverse transformations stated in (10.6.b.5) and (10.6.b.7) we compute an S matrix, again by inspection (Maple did the products on the right)

$$S = S^a{}_b = (\partial x^a/\partial x'^b) = \begin{pmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \end{pmatrix} \qquad SR = \begin{pmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(10.6.b.10)

$$S = S^a{}_b = (\partial x^a/\partial x'^b) = \begin{pmatrix} 0 & 3/5 & -1/5 \\ 0 & -1/5 & 2/5 \end{pmatrix} \qquad SR = \begin{pmatrix} 0 & 3/5 & -1/5 \\ 0 & -1/5 & 2/5 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Thus we have found two different "left inverses" S of the tall matrix R. If we try out these S matrices on the right of R, we find

$$RS = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5/3 & -1/3 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$RS = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 0 & 3/5 & -1/5 \\ 0 & -1/5 & 2/5 \end{pmatrix} = \begin{pmatrix} 0 & 1/5 & 3/5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (10.6.b.11)$$

Example 2 serves then to illustrate that a tall R matrix might have multiple left inverses, but those left inverses are not also right inverses. It turns out that there are in fact *no* right inverses for a tall R, as shown in section (c) below.

Before leaving this example, we comment on the "coordinate lines" in x-space using our first inverse solution (10.6.b.5).

$$\begin{aligned} x^1 &= -1/3\, x'^1 + 2/3\, x'^2 \\ x^2 &= 2/3\, x'^1 - 1/3\, x'^2 \end{aligned} \qquad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \qquad \textit{(10.6.b.5)}$$

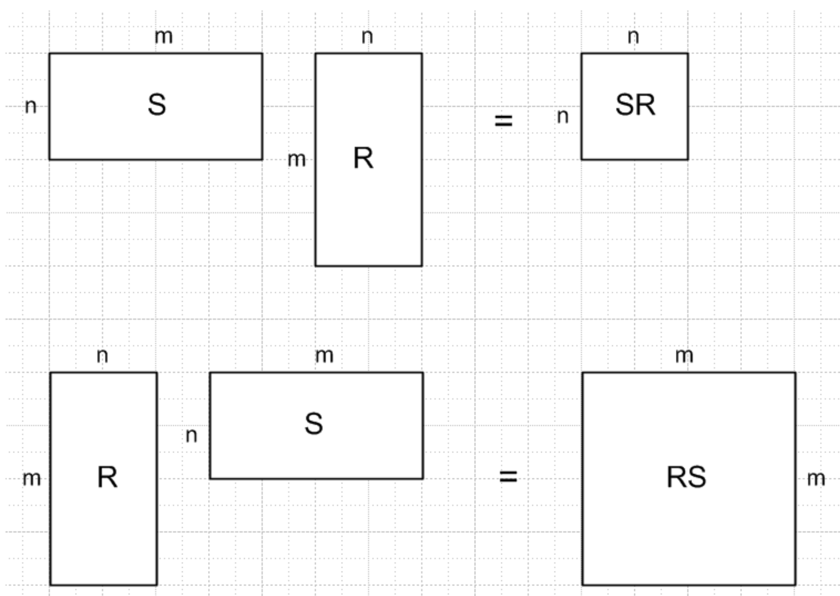
If we vary only x'^1 (keeping the other two coordinates in x'-space fixed) both x^1 and x^2 vary, and not surprisingly they define a certain line in x-space, and this is the coordinate line in x-space for x'^1 . If we instead vary only x'^2, again both x^1 and x^2 vary and they define some other line in x-space, the x'^2 coordinate line. If we vary only x'^3 , then x^1 and x^2 do not vary and this coordinate line is just a point!

Recall that the tangent base vectors $\mathbf{e_n}$ are tangent to the coordinate lines in x-space. As shown in (10.6.a.1) (e) one has $(\mathbf{e_j})^i = S^i{}_j$ so the tangent base vectors are the columns of S, S = $[\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}]$. Looking at S = $\begin{pmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \end{pmatrix}$ for our first inverse solution, we see that the first two tangent base vectors are indeed reasonable tangents to coordinate lines in x-space. Since the third coordinate line is just a point, it can have no tangent base vector, and in fact $\mathbf{e_3}$ = (0,0) which "resolves" this problem.

(c) Some Linear Algebra for non-square matrices

The linear algebra for non-square matrices is a topic often omitted in linear algebra presentations. Here we consider only the special case of two matrices where each has the shape of the transpose of the other, and we cherry-pick a few relevant theorems. As shown below, non-square matrices never have two-sided inverses, so one talks only about the possibility of such a matrix having a "right inverse" or a "left inverse".

Consider then the following matrix products where we assume m > n :



(10.6.c.1)

A nameless matrix rank theorem states the following :

Fact: rank(AB) ≤ min{rank(A), rank(B)} . (10.6.c.2)

Consider first the upper part of Fig (10.6.c.1). S and R each have some rank ≤ n, since n is the smaller matrix dimension. The Fact then says rank(SR) ≤ n. Since SR is an n x n matrix, it *could* therefore have full rank n, and then it is *possible* that one could have SR = 1. This says that it is possible for R to have a left inverse S, and for S to have a right inverse R.

Another nameless theorem states that if R has full rank n then in fact it has at least one left inverse S, and if S is full rank it has at least one right inverse R. The theorem does not say how to compute these inverses, nor does it suggest how many inverses there might be (a non-trivial problem). So,

Fact: tall R has full rank $\Rightarrow$ R has at least one left inverse S
wide S has full rank $\Rightarrow$ S has at least one right inverse R (10.6.c.3)

In our Example 2 above, matrix R in (10.6.b.9) has full rank 2, so we know it has at least one left inverse S. We explicitly found *two* such left inverses S as shown in (10.6.b.10). Since each of these left inverses has R as a right inverse, we know (and confirm) that each S must have full rank 2. Thus, we know (and confirm) that two of the tangent base vectors $\mathbf{e_n}$ are linearly independent (these being columns of S).

Now consider the lower part of Fig. (10.6.c.1). Fact (10.6.c.2) says rank(RS) ≤ n, but the matrix RS is m x m. Thus it cannot possibly have full rank m, so it can never be the m x m identity matrix (which would have rank m). We may then conclude that R has no right inverses and S has no left inverses:

Fact: tall R has no right inverses
wide S has no left inverses (10.6.c.4)

Corollary: A non-square matrix cannot have a two-sided inverse. (10.6.c.5)

If we take $S = R^T$, then the two matrices on the right in the drawing are R^TR and RR^T. Yet another matrix rank theorem says,

Fact: $\text{rank}(RR^T) = \text{rank}(R^TR) = \text{rank}(R)$. (10.6.c.6)

If R has full rank n, then the small matrix R^TR has rank n and so is full rank, $\det(R^TR) \neq 0$, and R^TR is invertible. But the m x m larger matrix RR^T having rank n must have $\det(RR^T) = 0$ and is not invertible.

Fact: If tall R has full rank n, then $(R^TR)^{-1}$ exists.
For any tall R, $(RR^T)^{-1}$ does not exist. (10.6.c.7)

With this in mind, another theorem says that if tall R is full rank, then we know *one* of its left inverses:

Fact: If tall R has full rank n, then one left inverse is given by $S = (R^TR)^{-1}R^T$. (10.6.c.8)

<u>Proof</u>: By the previous fact we know $(R^TR)^{-1}$ exists, so $SR = [(R^TR)^{-1}R^T]R = (R^TR)^{-1} (R^TR) = 1$.

Fact: If wide S has full rank n, then one right inverse is given by $R = S^T (SS^T)^{-1}$. . (10.6.c.8)

<u>Proof</u>: Reader exercise.

We mention in passing two other matrix theorems for arbitrary conforming matrices A,B,C:

Fact: (Sylvester's Inequality)
$\text{rank}(A) + \text{rank}(B) \le \text{rank}(AB) + n$ where n is the conforming dimension (10.6.c.9)

Fact: (Frobenius Inequality)
$\text{rank}(AB) + \text{rank}(BC) \le \text{rank}(ABC) + \text{rank}(B)$ (10.6.c.10)

(d) Implications for the Kinematics Package

The set of relations shown in (10.6.a.1) still stands for F: $R^n \to R^m$ with its tall R matrix, with the exception of the last item (i),

(i) $S = R^{-1}$ $R = S^{-1}$ $RS = SR = 1$ $RR^T = R^TR = SS^T = S^TS = 1$. *(10.6.a.1)*

This must be replaced by

(i) $SR = 1$ $SS^T = R^TR = 1$ (10.6.d.1)

since $RS \neq 1$ and two-sided inverses R^{-1} and S^{-1} do not exist for F: $R^n \to R^m$ with m>n.

A second implication is that certain items in the kinematics package are no longer unique. We have already seen that $S^i{}_j$ is not unique, so anything depending on $S^i{}_j$ is also not unique. Here is a list showing which objects are unique, and which are not:

<u>Metric tensors</u>

g_{ij}, g^{ij} unique
g'^{ij} unique, since $g'^{ij} = R^i{}_a R^j{}_b g^{ab}$
g'_{ij} **not** unique, since $g'_{ij} = R_i{}^a R_j{}^b g_{ab} = S^a{}_i S^b{}_j g_{ab}$ and $S^i{}_j$ not unique

<u>Transformation matrices</u>

$R^i{}_j = S_j{}^i$ unique (tall R matrix from **x**' = **F**(**x**))
$R^{ij} = S^{ji}$ unique since $R^{ij} = g^{ja} R^i{}_a$ and both g^{ja} and $R^i{}_a$ are unique
$R_j{}^i = S^i{}_j$ **not** unique
$R_{ij} = S_{ji}$ **not** unique, since $R_{ij} = g'_{ia} R^a{}_j$ and g'_{ia} not unique

<u>Axis-aligned basis vectors</u>

$(\mathbf{u}_j)^i$	unique since $(\mathbf{u}_j)^i = g_j{}^i$	$(\mathbf{e}'_j)^i$	unique since $(\mathbf{e}'_j)^i = g'^i{}_j$ $(= \delta^i{}_j)$
$(\mathbf{u}_j)_i$	unique since $(\mathbf{u}_j)_i = g_{ji}$	$(\mathbf{e}'_j)_i$	**not** unique since $(\mathbf{e}'_j)_i = g'_{ij}$
$(\mathbf{u}^j)^i$	unique since $(\mathbf{u}^j)^i = g^{ji}$	$(\mathbf{e}'^j)^i$	unique since $(\mathbf{e}'^j)^i = g'^{ij}$
$(\mathbf{u}^j)_i$	unique since $(\mathbf{u}^j)_i = g^j{}_i$	$(\mathbf{e}'^j)_i$	unique since $(\mathbf{e}'^j)_i = g'_i{}^j$ $(= \delta_i{}^j)$

Tangent base vectors

$(\mathbf{e}_j)^i$	**not** unique since $(\mathbf{e}_j)^i = R_j{}^i$	$(\mathbf{u'}_j)^i$	unique since $(\mathbf{u'}_j)^i = R^i{}_j$
$(\mathbf{e}_j)_i$	**not** unique since $(\mathbf{e}_j)_i = R_{ji}$	$(\mathbf{u'}_j)_i$	**not** unique since $(\mathbf{u'}_j)_i = R_{ij}$
$(\mathbf{e}^j)^i$	unique since $(\mathbf{e}^j)^i = R^{ji}$	$(\mathbf{u'}^j)^i$	unique since $(\mathbf{u'}^j)^i = R^{ij}$
$(\mathbf{e}^j)_i$	unique since $(\mathbf{e}^j)_i = R^j{}_i$	$(\mathbf{u'}^j)_i$	**not** unique since $(\mathbf{u'}^j)_i = R_i{}^j$ (10.6.d.2)

(e) Basis vectors for the Tangent Space at point x' on M

From (10.6.a.1) we select as a basis for x-space the set of n axis-aligned basis vectors $\mathbf{u}_i$,

$\{\mathbf{u}_i\}$ $i = 1,2...n$ basis for x-space

$(\mathbf{u}_i)^j = \delta_i{}^j$ components of these basis vectors in x-space .   (10.6.e.1)

These map into a set of n tangent base vectors $\mathbf{u'}_i$ in x'-space,

$\mathbf{u'}_i = R\,\mathbf{u}_i$ $|\mathbf{u'}_i> = \mathcal{R}\,|\mathbf{u}_i>$ *(2.5.1)*

or

$(\mathbf{u'}_i)^j = R^j{}_a\,(\mathbf{u}_i)^a = R^j{}_a\,\delta_i{}^a = R^j{}_i$ $i = 1,2...n$ $j = 1,2..m$.   (10.6.e.2)

We know that $\mathbf{u'}_i = R\,\mathbf{u}_i$ because this is the way *any* vector transforms: $\mathbf{v}' = R\,\mathbf{v}$.

Since there are m basis vectors in x'-space, we define the *rest* of the $\mathbf{u'}_i$ arbitrarily such that the m basis vectors $\{\mathbf{u'}_i\}$ in R^m are linearly independent, so

$\mathbf{u'}_i$ = as needed $i = n+1, n+2 .....m$.   (10.6.e.3)

Note in (10.6.e.2) that $(\mathbf{u'}_i)^j = R^j{}_a(\mathbf{u}_i)^a = \Sigma_{a=1}{}^n\ R^j{}_a(\mathbf{u}_i)^a$ is a "component sum equation", in contrast with the "vector sum equation" $\mathbf{e}^i = \Sigma_{j=1}{}^n R^i{}_j\mathbf{u}^j$ appearing in (10.6.a.2). To summarize for $\mathbf{u'}_i$:

$$\mathbf{u'}_i = \begin{cases} R\,\mathbf{u}_i & i = 1 \text{ through } n \\ \text{as needed} & i = n+1 \text{ through } m \end{cases} \quad . \tag{10.6.e.4}$$

We show just below that the first n $\mathbf{u'}_i$ span the tangent space $T_{x'}M$. Since the remaining $\mathbf{u'}_i$ must be selected so that the full set of m $\mathbf{u'}_i$ is a basis for x'-space, we know that the higher m-n $\mathbf{u'}_i$ must span the **perp space** $(T_{x'}M)^\perp$ of the tangent space, and this space is said to have **codimension** m-n within R^m.

Based on (10.6.e.2) that $R^j{}_i = (\mathbf{u'}_i)^j$, one concludes that the columns of $R^*{}_*$ are the contravariant basis vectors $\mathbf{u'}_i$ which span $T_{x'}M$. Each of these $\mathbf{u'}_i$ has m components and $R^*{}_*$ has m rows.

$R^*{}_* = [\mathbf{u'}_1, \mathbf{u'}_2 ....\mathbf{u'}_n]$.   *(2.5.9)*   (10.6.e.5)

As long as $R^*{}_*$ has full rank n, the columns are linearly independent so the $\mathbf{u'}_i$ form a (complete) basis.

We now show that the first n tangent base vectors $\mathbf{u'_i}$ do in fact span the tangent space $T_{\mathbf{x'}}M$.

Assume that, as **x** ranges over some portion of x-space, the mapping **x'** = **F**(**x**) describes a "smooth surface" M embedded in x'-space, hopefully a manifold or a piece thereof. If we start at some **x** and move to **x** + d**x** in x-space, we move from some point **x'** on M to some nearby point **x'** + d**x'** *on M*. By the definition of M, this d**x'** lies on the surface M and so is tangent to the surface M at **x'** and thus lies in the tangent space $T_{\mathbf{x'}}M$ of M at point **x'**. Applying R to each of the n axis-aligned differentials $d\mathbf{x^i} = dx^i(\mathbf{u_i})$ in x-space (no i sum), we thereby generate a set of n differential vectors $d\mathbf{x'^i} = Rd\mathbf{x^i}$ in x'-space which are in effect a set of short basis vectors which span the tangent space $T_{\mathbf{x'}}M$. Since $d\mathbf{x'^i} = dx'^i(\mathbf{u'_i})$, we may take the basis vectors $\{\mathbf{u'_i}$, i=1,2..n$\}$ as spanning $T_{\mathbf{x'}}M$. The upper $\mathbf{u'_i}$ are orthogonal to M and span the perp space $(T_{\mathbf{x'}}M)^\perp$ as noted.

We know from the fact $\mathbf{u'^i} \bullet \mathbf{u'_j} = \delta^i{}_j$ that the up-label (dual) vectors $\{\mathbf{u'^i}$, i=1,2..n$\}$ also form a basis for the tangent space $T_{\mathbf{x'}}M$. This conclusion can be reached as well by raising all i indices in the previous paragraph. In this case, the set $\{\mathbf{u'^i}$, i=n+1,n+2..m$\}$ are then all orthogonal to the "surface" M.

These last paragraphs and (10.6.e.5) have shown that:

Fact: The first n x'-space tangent base vectors $\mathbf{u'_i}$, which are the columns of full-rank $R^*{}_*$, span the tangent space $T_{\mathbf{x'}}M$ at point **x**' on M, and this is true as well for the $\mathbf{u'^i}$. (10.6.e.6)

10.7 The Pullback Operator $\mathcal{R}$ and properties of the Pullback Function F*

The Pullback Operator $\mathcal{R}$

From (10.6.e.2), or just from the fact that vectors transform as **v**' = R**v,** we know that

$\mathbf{u'_i} = R\mathbf{u_i}$ $\quad |\mathbf{u'_i}\rangle = \mathcal{R}\,|\mathbf{u_i}\rangle$ $\quad$ i = 1,2..n . *(2.5.1)* (10.7.1)

One can say that the n axis-aligned basis vectors $\mathbf{u_i}$ in x-space are "pushed forward" by R to become the tangent-space-spanning vectors $\mathbf{u'_i}$ in x'-space. Applying S to both sides and using (10.6.d.1) that SR = 1, one finds that

$\mathbf{u_i} = S\,\mathbf{u'_i}$ $\quad |\mathbf{u_i}\rangle = \mathcal{S}\,|\mathbf{u'_i}\rangle$ $\quad$ i = 1,2..n . (10.7.2)

From the package (10.6.a.1) item (h) we know that $S = R^T$ and $\mathcal{S} = \mathcal{R}^T$ for the corresponding Dirac operators, so the above may be written,

$\mathbf{u_i} = R^T\,\mathbf{u'_i}$ $\quad |\mathbf{u_i}\rangle = \mathcal{R}^T\,|\mathbf{u'_i}\rangle$ $\quad$ i = 1,2..n . (10.7.3)

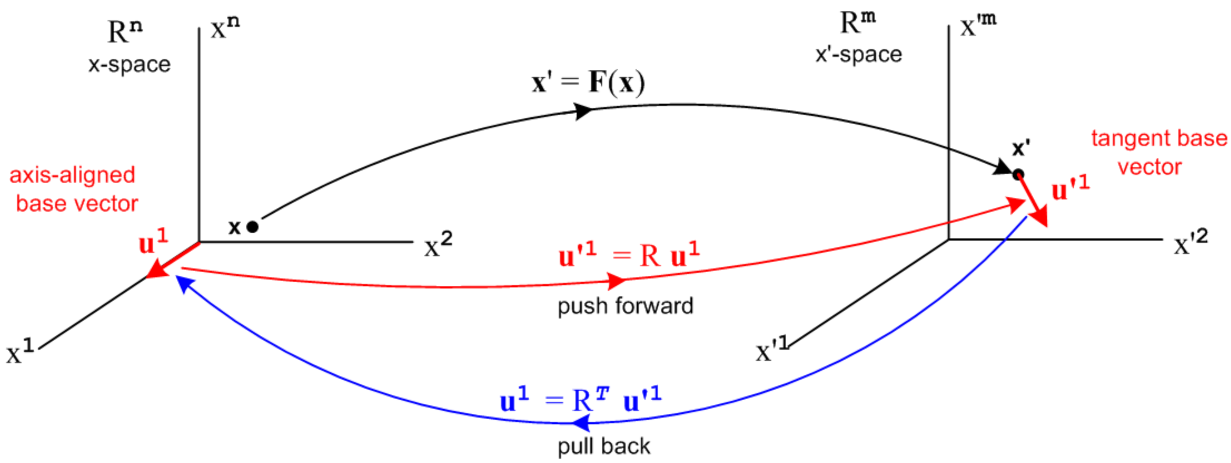
Thus, while operator $\mathcal{R}$ "pushes forward" the $|\mathbf{u_i}\rangle$ to the $|\mathbf{u'_i}\rangle$, the operator $\mathcal{R}^T$ "pulls back" the $|\mathbf{u'_i}\rangle$ from x'-space into the $|\mathbf{u_i}\rangle$ in x-space, just reversing the first process.

For the label-up u and e basis vectors one then has,

$\mathbf{u'^i} = R\mathbf{u^i}$	$\|\mathbf{u'^i}\rangle = \mathcal{R}\,\|\mathbf{u^i}\rangle$	i = 1,2..n	push forward
$\mathbf{u^i} = R^T\,\mathbf{u'^i}$	$\|\mathbf{u^i}\rangle = \mathcal{R}^T\,\|\mathbf{u'^i}\rangle$	i = 1,2..n	pull back

$$\mathbf{e}'^{i} = \mathrm{R}\mathbf{e}^{i} \qquad |\mathbf{e}'^{i}\rangle = \mathcal{R}\,|\mathbf{e}^{i}\rangle \qquad i = 1,2..n \qquad \text{push forward}$$
$$\mathbf{e}^{i} = \mathrm{R}^{T}\,\mathbf{e}'^{i} \qquad |\mathbf{e}^{i}\rangle = \mathcal{R}^{T}\,|\mathbf{e}'^{i}\rangle \qquad i = 1,2..n \qquad \text{pull back .} \qquad (10.7.4)$$

Here is a picture, reminiscent of Fig (2.5.4) (but reversed left to right), showing the above activity just for the $\mathbf{u}^1$ and $\mathbf{u}'^1$ basis vectors,



(10.7.5)

In the **dual space** of bras (linear functionals) (10.7.4) becomes, according to (2.11.g.10),

$$(\mathbf{u}'^{i})^{T} = (\mathbf{u}^{i})^{T}\,\mathrm{R}^{T} \qquad \langle\mathbf{u}'^{i}| = \langle\mathbf{u}^{i}|\mathcal{R}^{T} \qquad i = 1,2..n \qquad \text{push forward}$$
$$(\mathbf{u}^{i})^{T} = (\mathbf{u}'^{i})^{T}\mathrm{R} \qquad \langle\mathbf{u}^{i}| = \langle\mathbf{u}'^{i}|\mathcal{R} \qquad i = 1,2..n \qquad \text{pull back}$$

$$(\mathbf{e}'^{i})^{T} = (\mathbf{e}^{i})^{T}\,\mathrm{R}^{T} \qquad \langle\mathbf{e}'^{i}| = \langle\mathbf{e}^{i}|\mathcal{R}^{T} \qquad i = 1,2..n \qquad \text{push forward}$$
$$(\mathbf{e}^{i})^{T} = (\mathbf{e}'^{i})^{T}\mathrm{R} \qquad \langle\mathbf{e}^{i}| = \langle\mathbf{e}'^{i}|\mathcal{R} \qquad i = 1,2..n \qquad \text{pull back .} \qquad (10.7.6)$$

We refer to the $\mathcal{R}$ operator acting to the left as the **pullback operator**.

A picture similar to (10.7.5), which has dual x-space $(\mathrm{R}^n)^*$ on the left and dual x'-space $(\mathrm{R}^m)'^*$ on the right, would show the push forward $\langle\mathbf{u}'^1| = \langle\mathbf{u}^1|\mathcal{R}^T$ in red and the pullback $\langle\mathbf{u}^1| = \langle\mathbf{u}'^1|\mathcal{R}$ in blue. Below we shall have hybrid pictures showing the non-dual spaces but also showing the mapping of linear functionals between the dual-spaces.

Recall now the notations used in (8.7.1) for basis vectors in the dual wedge product spaces $\Lambda'^k(\mathrm{R}^m)$ and $\Lambda^k(\mathrm{R}^n)$,

$$\lambda'_{\wedge}{}^{I} \equiv \lambda'^{i_1} \wedge \lambda'^{i_2} \wedge .... \wedge \lambda'^{i_k} \qquad = \quad \langle\mathbf{e}'_{\wedge}{}^{I}| \equiv \langle\mathbf{e}'^{i_1}| \wedge \langle\mathbf{e}'^{i_2}| \wedge ..... \wedge \langle\mathbf{e}'^{i_k}|$$

$$\lambda_{\wedge}{}^{I} \equiv \lambda^{i_1} \wedge \lambda^{i_2} \wedge .... \wedge \lambda^{i_k} \qquad = \quad \langle\mathbf{u}_{\wedge}{}^{I}| \equiv \langle\mathbf{u}^{i_1}| \wedge \langle\mathbf{u}^{i_2}| \wedge ..... \wedge \langle\mathbf{u}^{i_k}| \qquad (10.7.7)$$

where the $\mathbf{e}'^{i}$ $(\mathbf{u}^{i})$ are axis-aligned basis vectors in x'-space (x-space). Recall also that,

$<\mathbf{e}^i| = <\mathbf{e}'^i|\mathcal{R}$ *(10.7.6)*

$\mathbf{e}^i = R^i{}_j\, \mathbf{u}^j$ or $|\mathbf{e}^i> = R^i{}_j\, |\mathbf{u}^j> \Rightarrow <\mathbf{e}^i| = R^i{}_j <\mathbf{u}^j| = <\mathbf{e}'^i|\mathcal{R}$. *(2.4.4)*

Then,

$$<\mathbf{e}'{}^{\wedge I}|\mathcal{R} = <\mathbf{e}'^{i_1}|\mathcal{R} \wedge <\mathbf{e}'^{i_2}|\mathcal{R} \wedge \,.....\, <\mathbf{e}'^{i_k}|\,\mathcal{R} \qquad // \ (8.9.d.15)$$

$$= (\,R^{i_1}{}_{j_1}<\mathbf{u}^{j_1}|\,) \wedge (\,R^{i_2}{}_{j_2}<\mathbf{u}^{j_2}|\,) \wedge \,...\, \wedge (\,R^{i_k}{}_{j_k}<\mathbf{u}^{j_k}|\,) \qquad // \ (2.4.4) \text{ above}$$

$$= R^{i_1}{}_{j_1} R^{i_2}{}_{j_2} \,...\, R^{i_k}{}_{j_k} \,(\, <\mathbf{u}^{j_1}| \wedge <\mathbf{u}^{j_2}| \wedge \,...\, \wedge <\mathbf{u}^{j_k}|\,) \qquad // \text{ reorder}$$

$$= \Sigma_J\, R^I{}_J <\mathbf{u}^{\wedge J}| \qquad // \text{ multiindex}$$

or

$$[\lambda'^{\wedge I}\ \mathcal{R}] = <\mathbf{e}'{}^{\wedge I}|\mathcal{R} = \Sigma_J\, R^I{}_J <\mathbf{u}^{\wedge J}| = \Sigma_J\, R^I{}_J\, \lambda^{\wedge J} \ . \qquad (10.7.8)$$

On the last line we write $[\lambda'^{\wedge I}\ \mathcal{R}]$ where $\mathcal{R}$ acts to the left on $\lambda'^{\wedge I}$ as a reminder of what is happening in the Dirac notation. For k=1 one would write $<\mathbf{e}'^i|\mathcal{R} = [\lambda'^i\mathcal{R}]$.

Eq. (10.7.8) shows that the **pullback of a k-form basis vector** $\lambda'^{\wedge I} = <\mathbf{e}'{}^{\wedge I}|$ from dual x'-space to dual x-space is a linear combination of k-form basis vectors $\lambda^{\wedge J} = <\mathbf{u}^{\wedge J}|$ in dual x-space which is then some k-form in dual x-space. The above equations are meaningful for $k \geq 1$.

For k=0, a 0-form in x'-space is just a scalar function f(**x'**). Since there are no basis vectors involved, there is no shuffling with $R^I{}_J$ and the pullback of the scalar f(**x'**) is just itself. That is to say, there is no distinction between the spaces $\Lambda'^0(R^m) = V^0 = K$ and $\Lambda^0(R^n) = V'^0 = K$ where K is the field of scalars . However, in x-space we want any object to be expressed in terms of x-space variables, so we write f(**x'**) as f(**F(x)**) since **x'** = **F**(**x**). Therefore,

Fact: The **pullback of a 0-form** may be written as

$$[f(\mathbf{x}')\ \mathcal{R}] = f(\mathbf{x}') = f(\mathbf{F}(\mathbf{x})) \qquad (10.7.9)$$

so $\mathcal{R}$ is really the unity operator in this $\Lambda^0 = V^0 = K$ (scalars) space. Note that

$$[f(\mathbf{x}')g(\mathbf{x}')\ \mathcal{R}] = f(\mathbf{x}')g(\mathbf{x}') = [f(\mathbf{x}')\ \mathcal{R}]\ [f(\mathbf{x}')\ \mathcal{R}] \ . \qquad (10.7.10)$$

The pullback of a 0-form (a function) times a k-form basis vector is then,

$$[f(\mathbf{x}')\, \lambda'^{\wedge I}\,]\, \mathcal{R} = <f(\mathbf{x}')\, \mathbf{e}'{}^{\wedge I}\,|\,\mathcal{R}$$

$$= (\, f(\mathbf{x}') < \mathbf{e}'{}^{\wedge I}\,|\,)\, \mathcal{R} \qquad // \text{ the space } \Lambda'^k(R^m) \text{ is linear since it is a vector space}$$

$$= f(\mathbf{x}')\, (\, < \mathbf{e}'{}^{\wedge I}\,|\,\mathcal{R}\,) \qquad // \ \mathcal{R} \text{ is a linear operator as in (2.11.g.29)}$$

$$= [f(\mathbf{x}')\ \mathcal{R}]\ [\lambda'^{\wedge I}\ \mathcal{R}\,] \qquad // \text{ using (10.7.9) and (10.7.8)} \quad . \qquad (10.7.11)$$

The **pullback of an arbitrary k-form** is then given by,

$$\alpha_{x'} = \Sigma'_I f_I(\mathbf{x}')\lambda'^{\wedge I} \quad \in \Lambda'^k \qquad \text{// k-form as in (10.2.3)} \qquad (10.7.12)$$

$$[\alpha_{x'}\mathcal{R}] = < \Sigma'_I f_I(\mathbf{x}')\, \mathbf{e}'^{\wedge I} \,|\, \mathcal{R}$$

$$= [\, \Sigma'_I f_I(\mathbf{x}') < \mathbf{e}'^{\wedge I} \,| \,]\, \mathcal{R} \qquad \text{// the space } \Lambda^k(R^m) \text{ is linear since it is a vector space}$$

$$= \Sigma'_I f_I(\mathbf{x}')\, [< \mathbf{e}'^{\wedge I} \,|\, \mathcal{R}\,] \qquad \text{// } \mathcal{R} \text{ is a linear operator as in (2.11.g.29)}$$

$$= \Sigma'_I f_I(\mathbf{x}')\, [\, \lambda'^{\wedge I} \mathcal{R}\,] \qquad \text{// } \lambda'^{\wedge I} = < \mathbf{e}'^{\wedge I} \,|$$

$$= \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\;\; \Sigma_J R^I{}_J\, \lambda^{\wedge J} \qquad \text{// (10.7.8)}$$

$$= \Sigma_J\, [\, \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, R^I{}_J\,]\, \lambda^{\wedge J} \qquad \text{// reorder}$$

$$= \Sigma_J\, G_J(\mathbf{x})\lambda^{\wedge J} \;\in \Lambda^k \qquad \text{where} \quad G_J(\mathbf{x}) \equiv \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, R^I{}_J \;. \qquad (10.7.13)$$

Note that $\alpha_{x'}$ is a k-form in dual x'-space, while $[\alpha_{x'}\mathcal{R}]$ is a linear combination of the $\lambda^{\wedge J}$ and therefore is a k-form in dual x-space. This will be rewritten with the ordered sum Σ'_J in Section 10.8 below. So,

Fact: The pullback of a k-form in $\Lambda'^k(R^m)$ is a k-form in $\Lambda^k(R^n)$. (10.7.14)

Finally, for a general k-form scaled by a function g(**x**'), using the same steps as above,

$$(g(\mathbf{x}')\alpha_{x'})\mathcal{R} \;= < g(\mathbf{x}')\, \alpha_{x'} \,|\, \mathcal{R} \;= g(\mathbf{x}') <\alpha_{x'} \,|\, \mathcal{R} \;=\; [g(\mathbf{x}')\mathcal{R}\,]\, [\alpha_{x'}\mathcal{R}] \;. \qquad (10.7.15)$$

The **pullback of a rank-k tensor function** is obtained by closing $[\alpha_{x'}\mathcal{R}]$ with a vector in the space V^k ,

$$[\alpha_{x'}\mathcal{R}](\mathbf{v}_1,\mathbf{v}_2....\mathbf{v}_k) = \;< \alpha_{x'} \,|\mathcal{R}|\, \mathbf{v}_1,\mathbf{v}_2....\mathbf{v}_k>$$

$$= < \alpha_{x'} \,|\mathcal{R}\, [\; |\, \mathbf{v}_1> \otimes |\, \mathbf{v}_2> \; ..... \otimes |\, \mathbf{v}_k>] \qquad \text{// definition of } |\mathbf{v}_1,\mathbf{v}_2....\mathbf{v}_k>$$

$$= < \alpha_{x'} \,|[\; |\, R\mathbf{v}_1> \otimes |\, R\mathbf{v}_2> \; ..... \otimes |\, R\mathbf{v}_k>] \qquad \text{// (5.6.17)}$$

$$= \; < \alpha_{x'} |\; R\mathbf{v}_1,R\mathbf{v}_2....R\mathbf{v}_k>$$

$$= \alpha_{x'}(R\mathbf{v}_1,R\mathbf{v}_2....R\mathbf{v}_k) \;. \qquad (10.7.16)$$

The object $\alpha_{x'}(R\mathbf{v}_1,R\mathbf{v}_2....R\mathbf{v}_k)$ is a rank-k tensor function in $\Lambda'^k{}_f(R^m)$: the functional $\alpha_{x'}$ lies in $\Lambda'^k(R^m)$ while the k vector arguments $\mathbf{v}'_i = R\mathbf{v}_i$ all lie in R^m . In contrast, the object $[\alpha_{x'}\mathcal{R}]\, (\mathbf{v}_1,\mathbf{v}_2....\mathbf{v}_k)$ is a rank-k tensor function in $\Lambda^k{}_f(R^n)$: the functional $[\alpha_{x'}\mathcal{R}]$ lies in $\Lambda^k(R^n)$ while the k vector arguments $\mathbf{v}_i$ all lie in R^n. The functional $[\alpha_{x'}\mathcal{R}]$ is the pullback of the functional $\alpha_{x'}$. Equation (10.7.16) says that the pulled-back tensor function $[\alpha_{x'}\mathcal{R}]$ in $\Lambda^k{}_f$ when evaluated at arguments $(\mathbf{v}_1,\mathbf{v}_2....\mathbf{v}_k)$ is equal in value to the un-pulled-back tensor function $\alpha_{x'}$ in $\Lambda'^k{}_f$ evaluated at arguments $(R\mathbf{v}_1,R\mathbf{v}_2....R\mathbf{v}_k)$.

These tensor functions are the objects that Spivak [1965] uses and he refers to them as k-tensors. Presentations which use only tensor functions regard (10.7.16) as the *definition* of a pullback $[\alpha_{x'}\mathcal{R}]$ of a differential k-form $\alpha_{x'}$.

The Pullback Function F*

The notation used above with the Dirac operator $\mathcal{R}$ acting to the left on a dual space vector is a bit clumsy, so one defines the following **pullback function** where $<\alpha'|$ is any k-form in $\Lambda'^k(R^m)$,

$$F^*(\alpha') \equiv <\alpha'|\mathcal{R} \qquad // \ \alpha' \equiv \alpha_{x'} \tag{10.7.17}$$

$$F^* : \Lambda'^k(R^m) \rightarrow \Lambda^k(R^n) \ . \tag{10.7.18}$$

Recall that the differential matrix R = (DF) and its associated Dirac operator $\mathcal{R}$ are specific to the underlying general transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, so to be more precise we could have written R_F and $\mathcal{R}_F$. The letter F in the function F* makes this connection explicit.

Various equations above can now be recast in terms of the pullback function F* :

Some Properties of the F* pullback function (10.7.19)

0 $F^*(\alpha') \equiv <\alpha'|\mathcal{R} = <\alpha'|\mathcal{R}_F$ // definition of F*, (10.7.17)

1 $F^*(f(\mathbf{x}')) = f(\mathbf{x}') = f(\mathbf{F}(\mathbf{x}))$ // F* on a 0-form, (10.7.9) for k = 0

2 $F^*(f(\mathbf{x}')\, g(\mathbf{x}')) = F^*(f(\mathbf{x}'))\, F^*(g(\mathbf{x}'))$ // F* on a product of two 0-forms, (10.7.10)

3 $F^*(f(\mathbf{x}')\, \lambda'^{\wedge I}) = F^*(f(\mathbf{x}'))\, F^*(\lambda'^{\wedge I})$ // F* on 0-form and basis-vector k-form, (10.7.11)

4 $F^*(\lambda'^{\wedge I}) = \Sigma_J R^I{}_J\, \lambda^{\wedge J}$ // F* on a basis-vector k-form, (10.7.8) for k ≥ 1

5 $F^*(\lambda'^i) = \Sigma_j R^i{}_j\, \lambda^j$ // F* on a basis-vector 1-form, k=1 of the above

6 $\alpha_{x'} = \Sigma'_I f_I(\mathbf{x}')\lambda'^{\wedge I}$ // general k-form in $\Lambda'^k(R^m)$, (10.7.12)

7 $F^*(\alpha_{x'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\ \Sigma_J R^I{}_J\, \lambda^{\wedge J}$ // F* pulling back a general k-form from Λ'^k, (10.7.13)

8 $F^*(g(\mathbf{x}')\, \alpha_{x'}) = F^*(g(\mathbf{x}'))\, F^*(\alpha_{x'})$ // F* on a 0-form times a general k-form, (10.7.15)

9 $[F^*(\alpha_{x'})](\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_{x'}(R\mathbf{v}_1,R\mathbf{v}_2...R\mathbf{v}_k)$ // F* pulling back a rank-k tensor function, (10.7.16)

Note that Σ_J in items 4 and 7 is the redundant symmetric sum. In (10.8.2) below we restate items 4 and 7 using the ordered sum Σ'_J, and then we restate everything again using cosmetic notation.

Other Properties of the F* pullback function

Fact: F* is **linear**, so $F^*(s_1\alpha' + s_2\beta') = s_1F^*(\alpha') + s_2F^*(\beta')$ where α' and β' are k-forms. (10.7.20)

Proof for k>0 : $F^*(s_1\alpha' + s_2\beta') = \langle s_1\alpha' + s_2\beta'| \mathcal{R}$ // (10.7.19) 0, definition of F*

$= s_1(\langle\alpha' |\mathcal{R}) + s_2(\langle\beta' |\mathcal{R})$ // $\mathcal{R}$ is a linear operator, see (2.11.g.29)

$= s_1F^*(\alpha') + s_2F^*(\beta')$ // (10.7.19) 0, definition of F* twice

Proof for k=0 : $F^*(s_1f(\mathbf{x}') + s_2g(\mathbf{x}')) = s_1f(\mathbf{x}') + s_2g(\mathbf{x}')$ // (10.7.19) 1, definition of F* on a function

$= F^*(s_1f(\mathbf{x}')) + F^*(s_2g(\mathbf{x}'))$ // (10.7.19) 2 definition of F* twice

$= s_1F^*(f(\mathbf{x}')) + s_2F^*(g(\mathbf{x}'))$ // (10.7.19) 2 and 1

Fact: $F^*(\alpha'_1 \wedge \alpha'_2 \wedge ... \wedge \alpha'_N) = F^*(\alpha'_1) \wedge F^*(\alpha'_2) \wedge ... \wedge F^*(\alpha'_N)$ where α'_i is an arbitrary k_i-form. (10.7.21)

Proof: $F^*(\alpha'_1 \wedge \alpha'_2 \wedge .... \wedge \alpha'_N) = [\langle\alpha'_1| \wedge \langle\alpha'_2| \wedge ... \wedge \langle\alpha'_N|] | \mathcal{R}$ // (10.7.19) 0 + Dirac notation

$= [\langle\alpha'_1| \mathcal{R} \wedge \langle\alpha'_2| \mathcal{R} \wedge ... \wedge \langle\alpha'_N| \mathcal{R}]$ // (8.9.d.15)

$= F^*(\alpha'_1) \wedge F^*(\alpha'_2) \wedge ... \wedge F^*(\alpha'_N)$. // (10.7.19) 0 QED

The result is valid if one or more of the forms are 0-forms. In this case, the two ^ operators surrounding a 0-form can be replaced by one ^. For example, $\langle\alpha'_1| \wedge f(x) \wedge \langle\alpha'_3| = f(x) \langle\alpha'_1| \wedge \langle\alpha'_3|$. In vector space notation, one has $\Lambda^n \wedge \Lambda^0 \wedge \Lambda^m = \Lambda^n \wedge \Lambda^m$ where Λ^0 is the space of scalars.

Fact: $F^*(d\alpha') = d(F^*(\alpha'))$ where $\alpha' \in \Lambda'^k$ is a k-form in x'-space (10.7.22)

This Fact says that the pullback function F* commutes with the exterior derivative operator d.

Proof: Show that Left Hand Side = Right Hand Side:

LHS: $\alpha' = \Sigma'_I f_I(\mathbf{x}') \lambda'^{\wedge I} \in \Lambda'^k$ (k-form in x'-space) // (10.7.12)

$d\alpha' = \Sigma'_I \, df_I(\mathbf{x}') \lambda'^{\wedge I} = \Sigma'_I \Sigma_{j=1}^m [\partial'_j f_I(\mathbf{x}')] \lambda'^j \wedge \lambda'^{\wedge I} \in \Lambda'^{k+1}$ // (10.3.6)

$F^*(d\alpha') = \Sigma'_I \Sigma_{j=1}^m [\partial'_j f_I(\mathbf{x}')] F^*(\lambda'^j \wedge \lambda'^{\wedge I}) \in \Lambda^{k+1}$ // (10.7.20) F* linear

$= \Sigma'_I \ \Sigma_{j=1}^m \ [\partial'_j f_I(\mathbf{x}')] \ F^*(\lambda'^j) \ \wedge F^*(\lambda'^{\wedge I})$. // (10.7.21) product

RHS: $\alpha' = \Sigma'_I f_I(\mathbf{x}') \lambda'^{\wedge I} \ \in \ \Lambda'^k$ (k-form in x'-space) // (10.7.12)

$F^*(\alpha') = \Sigma'_I \ F^*(f_I(\mathbf{x}')) F^*(\lambda'^{\wedge I}) \in \ \Lambda^k$ // (10.7.19) 3

$= \Sigma'_I \ f_I(\mathbf{F}(\mathbf{x})) F^*(\lambda'^{\wedge I}) \quad \in \ \Lambda^k$ // (10.7.19) 1

$d(F^*(\alpha')) = \ \Sigma'_I \ df_I(\mathbf{F}(\mathbf{x}))\ F^*(\lambda'^{\wedge I}) \quad \in \ \Lambda^{k+1}$ // (10.3.6)

$= \ \Sigma'_I \ \Sigma_{j=1}^m \ [\partial'_j f_I(\mathbf{x}') \ \Sigma_{r=1}^n \ (\partial x'^j/\partial x^r)] \ \lambda^r \wedge F^*(\lambda'^{\wedge I})$ // (10.3.6) + chain rule

$= \ \Sigma'_I \ \Sigma_{j=1}^m \ [\partial'_j f_I(\mathbf{x}') \ \Sigma_{r=1}^n \ R^j{}_r] \ \lambda^r \wedge F^*(\lambda'^{\wedge I})$ // (2.1.2)

$= \ \Sigma'_I \ \Sigma_{j=1}^m \ [\partial'_j f_I(\mathbf{x}')] \ (\Sigma_{r=1}^n \ R^j{}_r \ \lambda^r) \wedge F^*(\lambda^{\wedge I})$ // regroup

$= \ \Sigma'_I \ \Sigma_{j=1}^m \ [\partial'_j f_I(\mathbf{x}')] \ F^*(\lambda'^j) \wedge F^*(\lambda'^{\wedge I})$. // (10.7.19) 5

The LHS and RHS results are the same, so $F^*(d\alpha') = d(F^*(\alpha'))$. QED

Corollary: $d(F^*(d\alpha')) = 0$. (10.7.23)

Proof: $d(F^*(d\alpha')) \ = d(dF^*(\alpha')) \ = d^2 \ [F^*(\alpha')] = 0$ // (10.7.22) then (10.3.10)

Fact: $F^*(G^*\alpha) = (G \circ F)^* \ \alpha$ where α is a k-form (10.7.24)

Proof: This theorem involves two mappings F and G which are composed to form a third H :

$\mathbf{x}'' = \mathbf{G}(\mathbf{x}')$ $\quad \mathbf{x}' = \mathbf{F}(\mathbf{x})$ $\quad \mathbf{x} \xrightarrow{F} \mathbf{x}' \xrightarrow{G} \mathbf{x}''$ $\quad \mathbf{x} \xrightarrow{H} \mathbf{x}''$

$\mathbf{x}'' = \ \mathbf{G}(\mathbf{F}(\mathbf{x})) \equiv [\mathbf{G} \circ \mathbf{F}\ ](\mathbf{x}) \ = \mathbf{H}(\mathbf{x}) \quad \Rightarrow \quad H^* = (G \circ F)^*$

$d\mathbf{x}'' = R_G \, d\mathbf{x}'$ $\quad d\mathbf{x}' = R_F \, d\mathbf{x} \quad \Rightarrow \quad d\mathbf{x}'' = R_G R_F d\mathbf{x}$

$\mathbf{x}'' = \mathbf{H}(\mathbf{x}) \quad \Rightarrow \quad d\mathbf{x}'' = R_H d\mathbf{x}$ $\quad$ so: $\quad R_H = R_G R_F$ and $\mathcal{R}_H = \mathcal{R}_G \mathcal{R}_F$

Using the definition (10.7.17) that $F^*(\beta) \equiv \langle\beta|\mathcal{R}_F$ and $G^*(\alpha) \equiv \langle\alpha|\mathcal{R}_G$ we find,

$F^*(G^*\alpha) = F^*(\langle\alpha|\mathcal{R}_G\ ) \ = (\langle\alpha|\mathcal{R}_G\ )\ \mathcal{R}_F \ = \langle\alpha|\ \mathcal{R}_G\mathcal{R}_F \ = \ \langle\alpha|\mathcal{R}_H = H^*(\alpha) = (G \circ F)^*\ (\alpha)$ QED

If α is a 0-form (a function) $\alpha = f(\mathbf{x}'')$, then by (10.7.19) item 1,

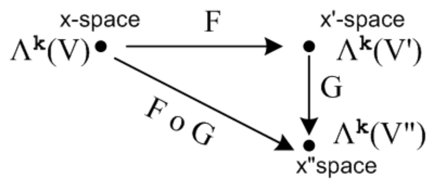
$$G^*(f(\mathbf{x}'')) = f(G(\mathbf{x}'))$$

so

$$F^*(G^*\alpha) = F^*(G^*f(\mathbf{x}'')) = F^*(f(\mathbf{G}(\mathbf{x}'))) = f(\mathbf{G}(\mathbf{F}(\mathbf{x}))) = f((\mathbf{G} \circ \mathbf{F})(\mathbf{x}))$$

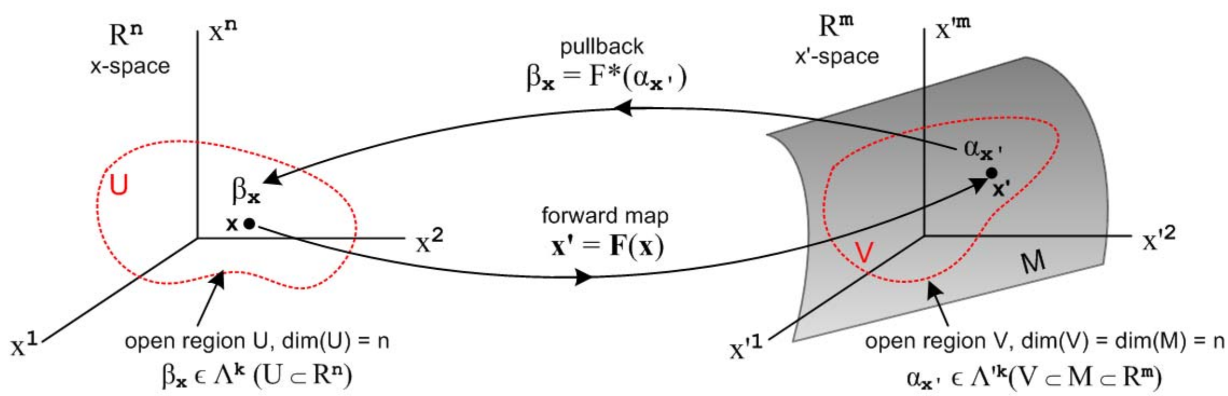
$$= f(\mathbf{H}(\mathbf{x})) = H^*(f(\mathbf{x}'')) = (G \circ F)^*(f(\mathbf{x}'')) = (G \circ F)^*\alpha$$ QED

A Chapter 1 style category diagram for this scenario would be



(10.7.25)

The following hybrid drawing shows the forward mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ between the non-dual spaces, and at the same time the pullback $\beta_{\mathbf{x}} = F^*(\alpha_{\mathbf{x}'})$ from the dual space Λ'^k on the right to dual space Λ^k on the left,



(10.7.26)

Here $\beta_{\mathbf{x}}$ is just a made-up name for the pulled back k-form $\alpha_{\mathbf{x}'}$. Recall that $\mathbf{x}'$ on $\alpha_{\mathbf{x}'}$ means that the k-form $\alpha_{\mathbf{x}'} = \Sigma'_{\mathbf{I}}\, f_{\mathbf{I}}(\mathbf{x}')\, \lambda'^{\wedge\mathbf{I}}$ is specific to the point $\mathbf{x}'$ on manifold M, while the $\mathbf{x}$ on $\beta_{\mathbf{x}}$ means that the k-form $\beta_{\mathbf{x}}$ is specific to the point $\mathbf{x}$ in x-space.

10.8 Alternate ways to write the pullback of a k-form

<u>The ordered sum form of a k-form pullback</u>

Certain expressions above contain the sum $\Sigma_{\mathbf{J}}\, R^{\mathbf{I}}{}_{\mathbf{J}}\, \lambda^{\wedge\mathbf{J}}$. As shown in Appendix A, because the object $R^{\mathbf{I}}{}_{\mathbf{J}}$ has a "factored form" , this sum can be written as on ordered sum $\Sigma'_{\mathbf{J}}$ as follows

$$\Sigma_{\mathbf{J}}\, R^{\mathbf{I}}{}_{\mathbf{J}}\, \lambda^{\wedge\mathbf{J}} = \Sigma'_{\mathbf{J}}\, \det(R^{\mathbf{I}}{}_{\mathbf{J}})\, \lambda^{\wedge\mathbf{J}} \qquad \text{// (A.8.37)} \qquad (10.8.1)$$

where the determinant magically appears. We can then rewrite two items from (10.7.19) :

4 $F^*(\lambda'^{\wedge I}) = \Sigma'_J \det(R^I{}_J)\, \lambda^{\wedge J}$ // F* on a basis-vector k-form

7 $F^*(\alpha_{x'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \det(R^I{}_J)\, \lambda^{\wedge J}$ // F* pulling back a general k-form from Λ'^k

$$= \Sigma'_J\, g_J(\mathbf{x})\lambda^{\wedge J} \quad \text{where} \quad g_J(\mathbf{x}) \equiv \Sigma'_I f_I(\mathbf{F}(\mathbf{x})) \det(R^I{}_J) \ . \tag{10.8.2}$$

It is useful to write out (10.8.2) in full detail. using the cosmetic notation $\lambda^i = dx^i$,

$$F^*(\alpha_{x'}) = \Sigma_{1\le j_1<j_2<....<j_k\le n}\ \Sigma_{1\le i_1<i_2<....<i_k\le m}\ f_{i_1 i_2 ... i_k}(\mathbf{F}(\mathbf{x}))$$

$$*\ \det [R^I{}_J] * (\ dx^{i_1} \wedge dx^{i_2} .....\wedge dx^{i_k})$$

$$= \Sigma_{1\le i_1<i_2<....<i_k\le m}\ f_{i_1 i_2 ... i_k}(\mathbf{F}(\mathbf{x}))$$

$$\Sigma_{1\le j_1<j_2<....<j_k\le n} \det [R^I{}_J] * (\ dx^{i_1} \wedge dx^{i_2} .....\wedge dx^{i_k}) \tag{10.8.3}$$

where $R^I{}_J$ is this kxk matrix,

$$R^I{}_J = \begin{matrix} R^{i_1}{}_{j_1} & R^{i_1}{}_{j_2} \dots & R^{i_1}{}_{j_k} \\ R^{i_2}{}_{j_1} & R^{i_2}{}_{j_2} \dots & R^{i_2}{}_{j_k} \\ ..... & & \\ R^{i_k}{}_{j_1} & R^{i_k}{}_{j_2} \dots & R^{i_k}{}_{j_k} \end{matrix} \tag{10.8.4}$$

where

$$R^i{}_j = (DF)^i{}_j = (\partial F^i/\partial x^j) = \partial_j F^i(\mathbf{x}) \ . \qquad // (10.6.2)$$

The object $\det(R^I{}_J)$ is a k x k minor of the full "tall" m x n matrix R, so $k \le n \le m$ in our application. Remember that, due to the ordered sums, all the i_r are different, and all the j_r are different, so no row or column appears twice in $R^I{}_J$.

<u>The dF^i form of a k-form pullback</u>

It is customary to define the object shown above in (10.8.1) as a certain k-form,

$$dF^{\wedge I} \equiv \Sigma_J R^I{}_J \lambda^{\wedge J} = \Sigma'_J \det(R^I{}_J)\, \lambda^{\wedge J} \ . \tag{10.8.5}$$

The motivation for doing this arises from the k = 1 case where the above becomes

$$dF^i \equiv R^i{}_j\, \lambda^j = R^i{}_j\, dx^j \tag{10.8.6}$$

where we replace λ^j by its cosmetic notation dx^j. The above equation then "looks just like" the normal calculus differential one obtains from transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ so $x'^i = F^i(\mathbf{x})$,

$$dF^i = (\partial F^i/\partial x^j)\, dx^j = (\partial x'^i/\partial x^j)\, dx^j = R^i{}_j dx^j \,. \tag{10.8.7}$$

That is to say,

$$\begin{aligned} dF^i &= R^i{}_j dx^j && \textit{(10.8.7)} \\ dF^i &= R^i{}_j\, dx^j && \textit{(10.8.6)} \end{aligned} \tag{10.8.8}$$

and in this same cosmetic notation we can rewrite (10.8.5) as

$$dF_\wedge{}^I = \Sigma_J R^I{}_J\, dx_\wedge{}^J = \Sigma'_J \det(R^I{}_J)\, dx_\wedge{}^J \,. \tag{10.8.9}$$

Using this new object, we write (10.8.2) as

4 $\quad F^*(\lambda'_\wedge{}^I) = dF_\wedge{}^I$ $\qquad$ // F* on a basis-vector k-form

7 $\quad F^*(\alpha_{\mathbf{x}'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, dF_\wedge{}^I$ $\qquad$ // F* pulling back a general k-form from Λ'^k $\qquad$ (10.8.10)

Therefore,

$$\begin{aligned} dF_\wedge{}^I = F^*(\lambda'_\wedge{}^I) &= F^*(\lambda'^{i_1}) \wedge F^*(\lambda'^{i_2}) \wedge \dots \wedge F^*(\lambda'^{i_k}) && \text{// (10.7.21) product rule} \\ &= [\lambda'^{i_1}\mathcal{R}] \wedge [\lambda'^{i_2}\mathcal{R}] \wedge \dots \wedge [\lambda'^{i_k}\mathcal{R}] && \text{// (10.7.17) def F*} \\ &= \langle \mathbf{e}'^{i_1}|\mathcal{R} \wedge \langle \mathbf{e}'^{i_2}|\mathcal{R} \wedge \dots \wedge \langle \mathbf{e}'^{i_k}|\mathcal{R} && \text{// (2.11.c.11) def } \lambda'^{i_1} \\ &= \langle \mathbf{e}^{i_1}| \wedge \langle \mathbf{e}^{i_2}| \wedge \dots \wedge \langle \mathbf{e}^{i_k}| && \text{// (10.7.6) pullbacks} \\ &= \langle \mathbf{e}_\wedge{}^I| \in \Lambda^k \,. && \text{// recall that } \lambda_\wedge{}^I = \langle \mathbf{u}_\wedge{}^I| \end{aligned} \tag{10.8.11}$$

so $dF_\wedge{}^I$ is just our old friend $\langle \mathbf{e}_\wedge{}^I|$.

As an example of (10.8.9) we write for k = 2,

$$dF^{i_1} \wedge dF^{i_2} = \Sigma_{1\le j_1<j_2\le n} \det \begin{pmatrix} \dfrac{\partial x'^{i_1}}{\partial x^{j_1}} & \dfrac{\partial x'^{i_1}}{\partial x^{j_2}} \\ \dfrac{\partial x'^{i_2}}{\partial x^{j_1}} & \dfrac{\partial x'^{i_2}}{\partial x^{j_2}} \end{pmatrix} dx^{j_1} \wedge dx^{j_2} \,. \tag{10.8.12}$$

Note that both sides of the above equation are 2-forms in Λ^2, whereas $dx'^{i_1} \wedge dx'^{i_2}$ is a 2-form in Λ'^2, so we cannot identify $dF^{i_1} \wedge dF^{i_2}$ with $dx'^{i_1} \wedge dx'^{i_2}$ even though $x'^i = F^i(\mathbf{x})$. But from (10.8.10) 4,

$$dF^{i_1} \wedge dF^{i_2} = F^*(\lambda'^{i_1} \wedge \lambda'^{i_2}) = F^*(dx'^{i_1} \wedge dx'^{i_2}) \qquad (10.8.13)$$

so in fact $dF^{i_1} \wedge dF^{i_2}$ is just the pullback of $dx'^{i_1} \wedge dx'^{i_2}$. This is just restating the basic fact of (10.7.6) that $\langle e^i| = \langle e'^i|\mathcal{R}$ so $\langle e^i|$ is the pullback of $\langle e'^i|$.

The reader will hopefully appreciate our use of red italic font to distinguish differential form objects from calculus objects of the same name. Otherwise things can be very confusing, especially in presentations where all ^ symbols are suppressed and where symmetric and ordered sums are both written as Σ_I.

Summary all in cosmetic notation for x' = F(x)

$$dx'^{\wedge I} \equiv dx'^{i_1} \wedge dx'^{i_2} \wedge .... \wedge dx'^{i_k} \quad \in \Lambda'^k(R^m)$$
$$dx^{\wedge I} \equiv dx^{i_1} \wedge dx^{i_2} \wedge .... \wedge dx^{i_k} \quad \in \Lambda^k(R^n) \qquad (10.7.7) \qquad (10.8.14)$$

Some Properties of the F* pullback function (10.7.19) (10.8.15)

0 $F^*(\alpha') \equiv \langle\alpha' | \mathcal{R} = \langle\alpha' | \mathcal{R}_F$ // definition of F*, (10.7.17)

1 $F^*(f(\mathbf{x}')) = f(\mathbf{x}') = f(\mathbf{F}(\mathbf{x}))$ // F* on a 0-form, (10.7.9) for k = 0

2 $F^*(f(\mathbf{x}')\, g(\mathbf{x}')) = F^*(f(\mathbf{x}'))\, F^*(g(\mathbf{x}'))$ // F* on a product of two 0-forms, (10.7.10)

3 $F^*(f(\mathbf{x}')\, dx'^{\wedge I}) = F^*(f(\mathbf{x}'))\, F^*(dx'^{\wedge I})$ // F* on 0-form and basis-vector k-form, (10.7.11)

4 $F^*(dx'^{\wedge I}) = \Sigma_J R^I{}_J\, dx^{\wedge J}$ // F* on a basis-vector k-form, (10.7.8) for k ≥ 1

5 $F^*(dx'^i) = \Sigma_j R^i{}_j\, dx^j$ // F* on a basis-vector 1-form, k=1 of the above

6 $\alpha_{x'} = \Sigma'_I f_I(\mathbf{x}')\, dx'^{\wedge I}$ // general k-form in Λ'^k, (10.7.12)

7 $F^*(\alpha_{x'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\ \Sigma_J R^I{}_J\, dx^{\wedge J}$ // F* pulling back a general k-form from Λ'^k, (10.7.13)

8 $F^*(g(\mathbf{x}')\, \alpha_{x'}) = F^*(g(\mathbf{x}'))\, F^*(\alpha_{x'})$ // F* on a 0-form and general k-form, (10.7.15)

9 $[F^*(\alpha_{x'})](\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_{x'}(R\mathbf{v}_1,R\mathbf{v}_2...R\mathbf{v}_k)$ // F* pulling back a rank-k tensor function, (10.7.16)

Other Properties of the F* pullback function (10.8.16)

Fact: F* is **linear**, so $F^*(s_1\alpha' + s_2\beta') = s_1F^*(\alpha') + s_2F^*(\beta')$ where α' and β' are k-forms. *(10.7.20)*

Fact: $F^*(\alpha'_1 \wedge \alpha'_2 \wedge ... \wedge \alpha'_N) = F^*(\alpha'_1) \wedge F^*(\alpha'_2) \wedge ... \wedge F^*(\alpha'_N)$ where α'_i is an arbitrary k_i-form. *(10.7.21)*

Fact: $F^*(d\alpha') = d(F^*(\alpha'))$ where $\alpha' \in \Lambda'^k$ is a k-form in x'-space *(10.7.22)*

Corollary: $d(F^*(d\alpha')) = 0$. *(10.7.23)*

Fact: $F^*(G^*\alpha) = (G \circ F)^* \alpha$ where α is a k-form . *(10.7.24)*

<u>The ordered sum form of a k-form pullback</u>

$\Sigma_J R^I{}_J\, dx^{\wedge J} = \Sigma'_J \det(R^I{}_J)\, dx^{\wedge J}$ // (A.8.37) *(10.8.1)* (10.8.17)

4 $F^*(dx'^{\wedge I}) = \Sigma'_J \det(R^I{}_J)\, dx^{\wedge J}$ // F* on a basis-vector k-form (10.8.18)

7 $F^*(\alpha_{x'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \det(R^I{}_J)\, dx^{\wedge J}$ // F* pulling back a general k-form from Λ'^k

$= \Sigma'_J g_J(\mathbf{x})\, dx^{\wedge J}$ where $g_J(\mathbf{x}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x})) \det(R^I{}_J)$ *(10.8.2)* (10.8.19)

<u>The dF^i form of a k-form pullback</u>

$dF^{\wedge I} \equiv \Sigma_J R^I{}_J\, dx^{\wedge J} = \Sigma'_J \det(R^I{}_J)\, dx^{\wedge J} = F^*(dx'^{\wedge I}) = \langle \mathbf{e}^{\wedge I}|$ *(10.8.9,10,11)* (10.8.20)

$$F^*(dx'^{i_1} \wedge dx'^{i_2}) = dF^{i_1} \wedge dF^{i_2} = \Sigma_{1\le j_1<j_2\le n} \det\begin{pmatrix} \dfrac{\partial x'^{i_1}}{\partial x^{j_1}} & \dfrac{\partial x'^{i_1}}{\partial x^{j_2}} \\ \dfrac{\partial x'^{i_2}}{\partial x^{j_1}} & \dfrac{\partial x'^{i_2}}{\partial x^{j_2}} \end{pmatrix} dx^{j_1} \wedge dx^{j_2} .$$

$$\equiv \Sigma_{1\le j_1<j_2\le n} \frac{\partial(x'^{i_1}, x'^{i_2})}{\partial(x^{j_1}, x^{j_2})}\, dx^{j_1} \wedge dx^{j_2} \qquad (10.8.12) \qquad (10.8.21)$$

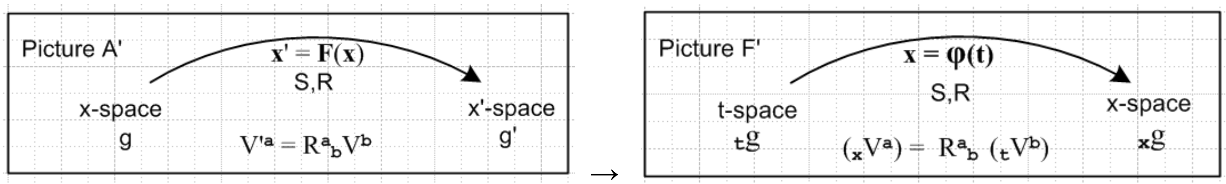
$dF^i = R^i{}_j\, dx^j$ *(10.8.8)* (10.8.22)

$F^*(dx'^{\wedge I}) = F^*(dx'^{i_1} \wedge dx'^{i_2} \ldots \wedge dx'^{i_k}) = dF^{i_1} \wedge dF^{i_2} \ldots \wedge dF^{i_k} = dF^{\wedge I}$ *(10.8.10) 4* (10.8.23)

$F^*(\alpha_{x'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, dF^{\wedge I}$ *(10.8.10) 7* (10.8.24)

10.9 A Change of Notation and Comparison with Sjamaar and Spivak

To this point we have maintained the notation of Chapter 2 (and *Tensor*) for transformations $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. To compare our results with other sources, we shall now make the following change of notation :



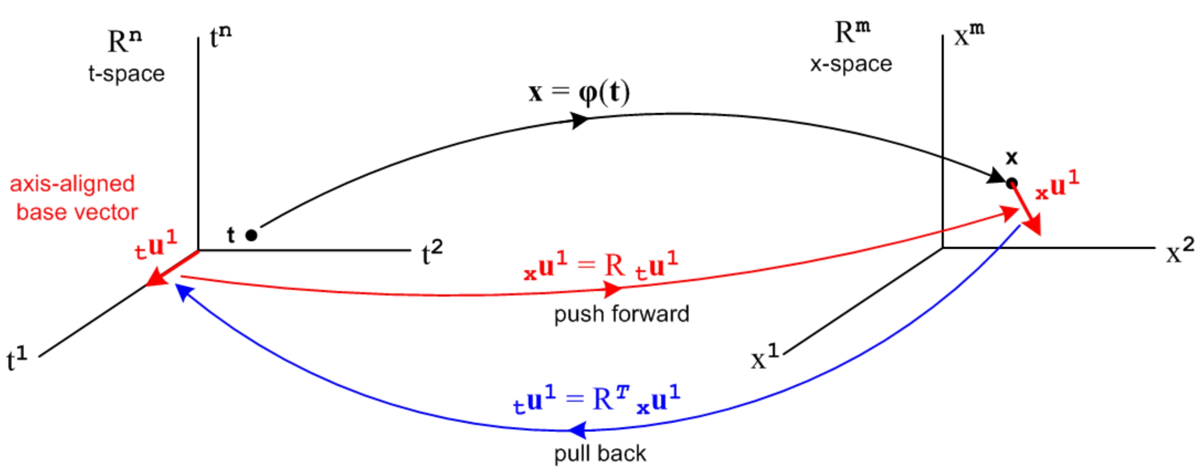
x-space	→	t-space		
x'-space	→	x-space		
F	→	**φ**	general transformation name	
x' = **F**(**x**)	→	**x** = **φ**(**t**)	general transformation equation	
R = (DF)	→	R = (Dφ)	differential matrix	
F*	→	φ*	pullback function	(10.9.1)

Confusingly, in Picture A' the left-side space has the name x-space, while in Picture F' this happens to be the name of the right-side space. This is just a coincidental new definition of x-space.

This notation is convenient for presenting results, but it is somewhat clumsy for developing those results as we have done above. Having primes and no primes is very efficient compared to the other changes one must make to develop in this new notation. For example one must write $\lambda'^{i} \rightarrow {}_{x}\lambda^{i}$ and $\lambda^{i} \rightarrow {}_{t}\lambda_{i}$. Similarly, one has $\mathbf{u}'^{i} \rightarrow {}_{x}\mathbf{u}^{i}$ for the tangent base vectors which span T_xM and $\mathbf{u}^{i} \rightarrow {}_{t}\mathbf{u}^{i}$ are the axis-aligned basis vectors in t-space ("parameter space"). Just for the record, these changes are all shown in Appendix E.

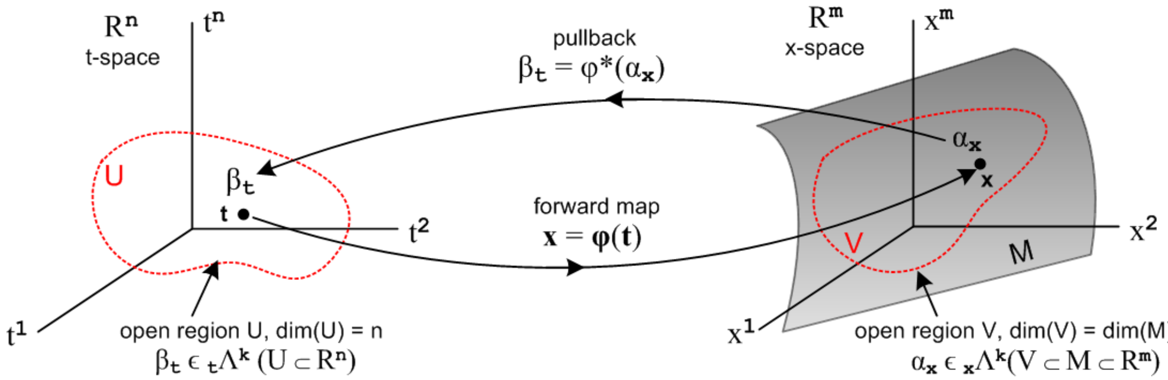
Drawings for the new spaces

In terms of the new t-space and x-space, this drawing (a translation of Fig (10.7.5)) shows the push forward and pull back of the first basis vectors in R^n and R^m,



(10.9.2)

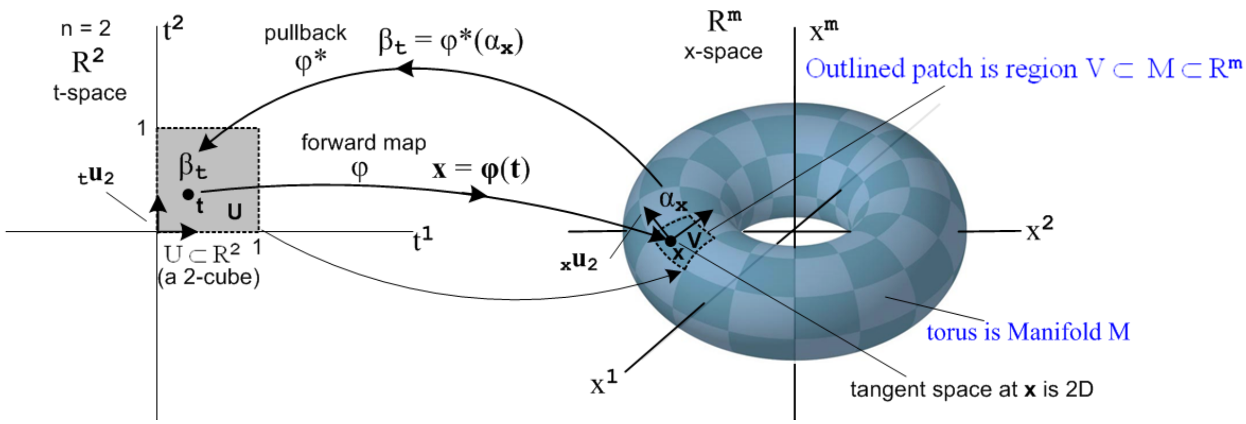
The next drawing (translation of Fig (10.7.26)) shows the pullback of a differential k-form α_x from ${}_x\Lambda^k(R^m)$ to k-form β_t in ${}_t\Lambda^k(R^n)$,



(10.9.3)

Here β_t is just a made-up name for the pulled back k-form α_x. Recall that **x** on α_x means that the k-form $\alpha_x = \Sigma'_I f_I(\mathbf{x})\ {}_x\lambda^{\wedge I}$ is specific to the point **x** on manifold M, while the **t** on β_t means that the k-form β_t is specific to the point **t** in t-space.

Here is a more practical picture for the special case n = 2 and k = 2:



(10.9.3a)

Here the open region U is a unit square $[0,1]^2$ which maps into a patch on a torus. That is, if m = 3 the object on the right is a torus in R^3, but we can imagine it to be a torus embedded in R^m for any $m \geq 3$.

The space of functionals defined on $U \subset R^2$ is a 2-dimensional dual space $(R^{*2})(U)$. On this space we can define either 1-forms or 2-forms. The above picture suggests a 2-form since the region U is an area, and since we will later *associate* $dt^1 \wedge dt^2$ with the calculus differential dt^1dt^2 which represents an area (we are not there yet).

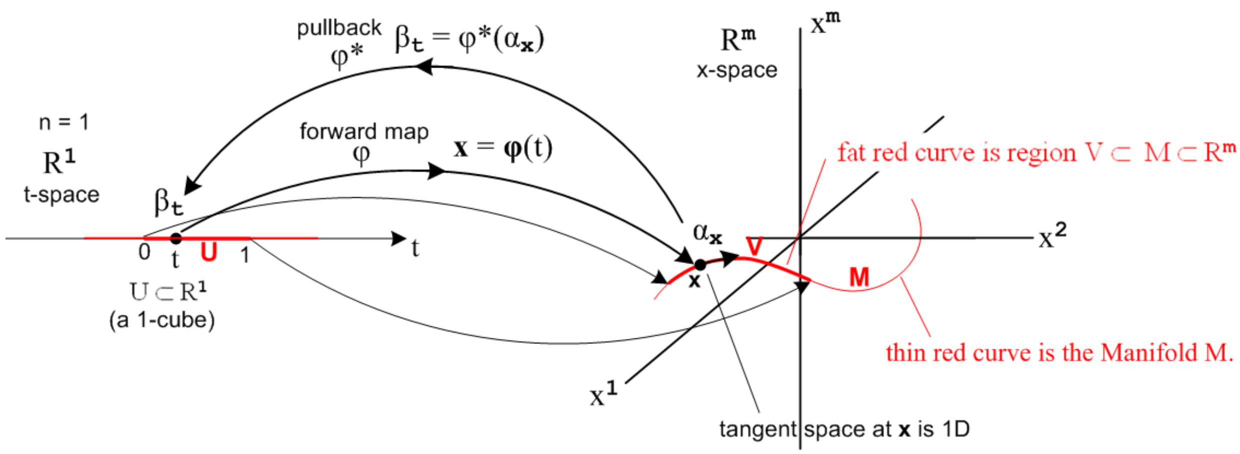
The picture shows the "forward map" $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$, suggesting that forward means left to right in the picture. Then α_x is "pulled back" right to left from dual x-space to dual t-space where it becomes β_t.

One could imagine a set of 16x6 = 96 mappings like the one shown above which would "cover the torus", using one little patch for each mapping (with some small overlap between patches). One would then have an atlas of 96 square maps like that on the left which would serve to cover the surface of Planet Toroid. This is the basic idea of a manifold. In the torus example, one could do the job with only 2 maps. Doing it with a single map does not fly since then some seam curve on the torus would map back to two boundaries of the square and the mapping is then not one-to-one and smooth. Manifold mappings have to be continuous in both mapping directions at every point, and a seam is a place without continuity.

The aspect ratio of the 2-cube on the left is not significant. One could change it to be an arbitrary rectangle in t-space and select a **φ** to make it map to the same small image patch in x-space. Or one could construct a mapping **φ** which maps the unit 2-cube $[0,1]^2$ to the entire left half of the torus. See Sjamaar.

The black arrows on the left are the t-space basis vectors $_t\mathbf{u}_i$ (only $_t\mathbf{u}_2$ is labeled). These map according to $_x\mathbf{u}^i = R\ _t\mathbf{u}^i$ (formerly $\mathbf{u}'^i = R\ \mathbf{u}^i$) into basis vectors which are tangent to M, and these vectors then span the tangent space T_xM at point **x** on M. It is clear that the two $_x\mathbf{u}^i$ vary as the point **x** on M is varied.

As another example consider this situation with n = 1 and k = 1,



(10.9.3b)

Now the domain in t-space is U = 1-cube [0,1] which maps to a (generally non-planar) red curve which is embedded in R^m . Here α_x and β_t are 1-forms. The red curve segment V lies on the manifold curve M as shown, just as the patch of the previous example lay on the torus. There is only one basis vector $_t\mathbf{u}$ in t-space (not shown) and it maps to the unlabeled black arrow on the right which is $_x\mathbf{u}$ and is of course tangent to the curve at **x**.

We now reproduce the "Summary in all cosmetic notation" given above at the end of Section 10.8 but in terms of this new notation:

Summary all in cosmetic notation for x = φ(t)

$dx_\wedge^I \equiv dx^{i_1} \wedge dx^{i_2} \wedge .... \wedge dx^{i_k}$ basis vector $\in\ _x\Lambda^k(R^m)$

$dt_\wedge^I \equiv dt^{i_1} \wedge dt^{i_2} \wedge .... \wedge dt^{i_k}$ basis vector $\in\ _t\Lambda^k(R^n)$ *(10.8.14)* (10.9.4)

<u>Some Properties of the φ* pullback function</u> $R = (D\varphi)$ *(10.8.15)*   (10.9.5)

0    $\varphi^*(\alpha_\mathbf{x}) \equiv \langle\alpha_\mathbf{x} | \mathcal{R} = \langle\alpha_\mathbf{x} | \mathcal{R}_\mathbf{F}$ // definition of φ*, (10.7.17)

1    $\varphi^*(f(\mathbf{x})) = f(\mathbf{x}) = f(\boldsymbol{\varphi}(\mathbf{t}))$ // φ* on a 0-form, (10.7.9) for k = 0

2    $\varphi^*(f(\mathbf{x})\, g(\mathbf{x})) = \varphi^*(f(\mathbf{x}))\, \varphi^*(g(\mathbf{x}))$ // φ* on a product of two 0-forms, (10.7.10)

3    $\varphi^*(f(\mathbf{x})\, dx^{\wedge\mathbf{I}}) = \varphi^*(f(\mathbf{x}))\, \varphi^*(dx^{\wedge\mathbf{I}})$ // φ* on 0-form and basis-vector k-form, (10.7.11)

4    $\varphi^*(dx^{\wedge\mathbf{I}}) = \Sigma_\mathbf{J} R^\mathbf{I}{}_\mathbf{J}\, dt^{\wedge\mathbf{J}}$ // φ* on a basis-vector k-form, (10.7.8) for k ≥ 1

5    $\varphi^*(dx^i) = \Sigma_j R^i{}_j\, dt^j$ // φ* on a basis-vector 1-form, k=1 of item 4

6    $\alpha_\mathbf{x} = \Sigma'_\mathbf{I} f_\mathbf{I}(\mathbf{x})\, dx^{\wedge\mathbf{I}}$ // general k-form in ${}_\mathbf{x}\Lambda^k$, (10.7.12)

7    $\varphi^*(\alpha_\mathbf{x}) = \Sigma'_\mathbf{I} f_\mathbf{I}(\boldsymbol{\varphi}(\mathbf{t}))\, \Sigma_\mathbf{J} R^\mathbf{I}{}_\mathbf{J}\, dt^{\wedge\mathbf{J}}$ // φ* pulling back a general k-form from ${}_\mathbf{x}\Lambda^k$, (10.7.13)

8    $\varphi^*(g(\mathbf{x})\, \alpha_\mathbf{x}) = \varphi^*(g(\mathbf{x}))\, \varphi^*(\alpha_\mathbf{x})$ // φ* on a 0-form and general k-form, (10.7.15)

9    $[\varphi^*(\alpha_\mathbf{x})](\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_\mathbf{x}(R\mathbf{v}_1,R\mathbf{v}_2...R\mathbf{v}_k)$ // φ* pulling back a rank-k tensor function, (10.7.16)

<u>Other Properties of the φ* pullback function</u>   *(10.8.16)*   (10.9.6)

These five items are translations of (10.7.20) through (10.7.24) :

Fact: φ* is linear, so $\varphi^*(s_1\alpha + s_2\beta) = s_1\varphi^*(\alpha) + s_2\varphi^*(\beta)$ where α and β are k-forms.   (10.9.7)

Fact: $\varphi^*(\alpha_1 \wedge \alpha_2 \wedge ... \wedge \alpha_N) = \varphi^*(\alpha_1) \wedge \varphi^*(\alpha_2) \wedge ... \wedge \varphi^*(\alpha_N)$ where α_i is an arbitrary k_i-form.   (10.9.8)

Fact: $\varphi^*(d\alpha) = d(\varphi^*(\alpha))$ where $\alpha \in {}_\mathbf{x}\Lambda^k$ is a k-form in x-space   (10.9.9)

Corollary: $d(\varphi^*(d\alpha)) = 0$.   (10.9.10)

Fact: $\varphi^*(\psi^*\alpha) = (\psi \circ \varphi)^*\, \alpha$ where α is a k-form .   (10.9.11)

<u>The ordered sum form of a k-form pullback</u> $R = (D\varphi)$

$\Sigma_\mathbf{J} R^\mathbf{I}{}_\mathbf{J}\, dx^{\wedge\mathbf{J}} = \Sigma'_\mathbf{J} \det(R^\mathbf{I}{}_\mathbf{J})\, dt^{\wedge\mathbf{J}}$ // (A.8.37)   *(10.8.17)*   (10.9.12)

4    $\varphi^*(dx^{\wedge\mathbf{I}}) = \Sigma'_\mathbf{J} \det(R^\mathbf{I}{}_\mathbf{J})\, dt^{\wedge\mathbf{J}}$ // φ* on a basis-vector k-form   *(10.8.18)*   (10.9.13)

7    $\varphi^*(\alpha_\mathbf{x}) = \Sigma'_\mathbf{I} f_\mathbf{I}(\boldsymbol{\varphi}(\mathbf{t}))\, \Sigma'_\mathbf{J} \det(R^\mathbf{I}{}_\mathbf{J})\, dt^{\wedge\mathbf{J}}$ // φ* pulling back a general k-form from ${}_\mathbf{x}\Lambda^k$

$= \Sigma'_\mathbf{J}\, g_\mathbf{J}(\mathbf{t})\, dt^{\wedge\mathbf{J}}$ where $g_\mathbf{J}(\mathbf{t}) = \Sigma'_\mathbf{I} f_\mathbf{I}(\boldsymbol{\varphi}(\mathbf{t}))\, \det(R^\mathbf{I}{}_\mathbf{J})$ *(10.8.19)*   (10.9.14)

The $d\varphi^i$ form of a k-form pullback

$$d\varphi^{\wedge I} \equiv \Sigma_J R^I{}_J\, dt^{\wedge J} = \Sigma'_J \det(R^I{}_J)\, dt^{\wedge J} = \varphi^*(dx^{\wedge I}) = \langle {}_t e^{\wedge I}| \qquad (10.8.20) \qquad (10.9.15)$$

$$\varphi^*(dx^{i_1} \wedge dx^{i_2}) = d\varphi^{i_1} \wedge d\varphi^{i_2} = \Sigma_{1\le j_1<j_2\le n} \det\begin{pmatrix} \frac{\partial\varphi^{i_1}}{\partial t^{j_1}} & \frac{\partial\varphi^{i_1}}{\partial t^{j_2}} \\ \frac{\partial\varphi^{i_2}}{\partial t^{j_1}} & \frac{\partial\varphi^{i_2}}{\partial t^{j_2}} \end{pmatrix} dt^{j_1} \wedge dt^{j_2} .$$

$$\equiv \Sigma_{1\le j_1<j_2\le n} \frac{\partial(\varphi^{i_1}, \varphi^{i_2})}{\partial(t^{j_1}, t^{j_2})} dt^{j_1} \wedge dt^{j_2} \qquad (10.8.21) \qquad (10.9.16)$$

$$d\varphi^i = R^i{}_j\, dt^j \qquad (10.8.22) \qquad (10.9.17)$$

$$\varphi^*(dx^{\wedge I}) = \varphi^*(dx^{i_1} \wedge dx^{i_2} \ldots \wedge dx^{i_k}) = d\varphi^{i_1} \wedge d\varphi^{i_2} \ldots \wedge d\varphi^{i_k} = d\varphi^{\wedge I} \qquad (10.8.10) \qquad (10.9.18)$$

$$\varphi^*(\alpha_{\mathbf{x}}) = \Sigma'_I f_I(\boldsymbol{\varphi}(\mathbf{t}))\, d\varphi^{\wedge I} \qquad (10.8.24) \qquad (10.9.19)$$

Comparison with Sjamaar

Our document was strongly motivated by Sjamaar's excellent notes, so it seems useful to make some connection to those notes. In most of our document we used the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ but in Section 10.9 we changed this to be $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ to bring things closer to Sjamaar and other authors.

Sjamaar uses $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ in his Ch 3 on pullbacks, $\mathbf{x} = \mathbf{c}(t)$ in Ch 4 on 1-forms, and $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ in Ch 5 on integration and Ch 6 on manifolds. He does not stress the notion of an underlying transformation as we have done because he has many more important details to attend to, but he does show $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ in his figure on page 39. All wedge product symbols ^ are suppressed with the idea that almost all products are wedge products, so one sees equations like $dx_1 dx_2 = - dx_2 dx_1$. His sum Σ_I is almost always an ordered sum which we write as Σ'_I.

Here then is a sampling of our equations above and how they appear in Sjamaar's 2015 notes :

$\alpha \wedge \beta = (-1)^{kk'} \beta \wedge \alpha$ $\qquad$ α = k-form, β = k'-form $\qquad$ (10.4.1)

$\boxed{\beta\alpha = (-1)^{kl}\alpha\beta}$ $\qquad$ Sja p 19, "graded commutivity"

Fact: $d^2\alpha = 0$ for any k-form α (differential forms have zero "curvature") . $\qquad$ (10.3.10)

2.6. Proposition. *$d(d\alpha) = 0$ for any form α. In short,*

$\boxed{d^2 = 0.}$ $\qquad$ Sja p 22

$\alpha = \Sigma'_{\mathbf{I}}\ f_{\mathbf{I}}(\mathbf{x})\ {}_{\mathbf{x}}\lambda^{\wedge \mathbf{I}} = \Sigma'_{\mathbf{I}}\ f_{\mathbf{I}}(\mathbf{x})\ dx^{\wedge \mathbf{I}}$ // a k-form *(10.1.14)*

$$\alpha = \sum_I f_I\, dx_I, \qquad \alpha = \sum_I f_I\, dy_I.$$ Sja p 19,39

$\varphi^*(f(\mathbf{x})\ dx^{\wedge I}) = \varphi^*(f(\mathbf{x}))\ \varphi^*(dx^{\wedge I})$ *(10.9.5) 2*

$$\phi^*(\alpha) = \sum_I \phi^*(f_I)\phi^*(dy_I).$$ Sja p 39, related to the above

$\varphi^*(dx^{\wedge I}) = \varphi^*(dx^{i1} \wedge dx^{i2} ...\wedge dx^{ik}) = d\varphi^{i1} \wedge d\varphi^{i2} ...\wedge d\varphi^{ik} = d\varphi^{\wedge I}$ *(10.9.18)*

$$\phi^*(dy_I) = \phi^*(dy_{i_1}\, dy_{i_2} \cdots dy_{i_k}) = d\phi_{i_1}\, d\phi_{i_2} \cdots d\phi_{i_k}$$ Sja p 39

Fact: φ^* is linear, so $\varphi^*(s_1\alpha + s_2\beta) = s_1\varphi^*(\alpha) + s_2\varphi^*(\beta)$ where α and β are k-forms. *(10.9.7)*
Fact: $\varphi^*(\alpha_1 \wedge \alpha_2 \wedge ...\wedge \alpha_N) = \varphi^*(\alpha_1) \wedge \varphi^*(\alpha_2) \wedge ...\wedge \varphi^*(\alpha_N)$ where α_i is an arbitrary k_i-form. *(10.9.8)*
Fact: $\varphi^*(\psi^*\alpha) = (\psi \circ \varphi)^*\ \alpha$ where α is a k-form *(10.9.11)*

3.10. Proposition. *Let $\phi\colon U \to V$ be a smooth map, where U is open in $\mathbf{R}^n$ and V is open in $\mathbf{R}^m$. The pullback operation is*
(i) *linear: $\phi^*(a\alpha + b\beta) = a\phi^*(\alpha) + b\phi^*(\beta)$ for all scalars a and b and all k-forms α and β on V;*
(ii) *multiplicative: $\phi^*(\alpha\beta) = \phi^*(\alpha)\phi^*(\beta)$ for all k-forms α and l-forms β on V;*
(iii) *natural: $\phi^*(\psi^*(\alpha)) = (\psi \circ \phi)^*(\alpha)$, where $\psi\colon V \to W$ is a second smooth map with W open in $\mathbf{R}^l$, and α is a k-form on W.*

Sja p 40

Fact: $\varphi^*(d\alpha) = d(\varphi^*(\alpha))$ where $\alpha \in {}_{\mathbf{x}}\Lambda^{\mathbf{k}}$ is a k-form in x-space *(10.9.9)*

$$\boxed{\phi^* d = d\phi^*.}$$ Sja p 41

$\varphi^*(dx^{\wedge I}) = d\varphi^{\wedge I} = \Sigma'_{\mathbf{J}}\ \det((D\varphi)^{\mathbf{I}}{}_{\mathbf{J}})\ dt^{\mathbf{J}}$ for $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ *(10.9.18), (10.9.15)*

$$d\phi_{i_1}\, d\phi_{i_1} \cdots d\phi_{i_k} = \sum_J \det(D\phi_{I,J})\, dx_J$$ Sja p 44 for $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$

$\varphi^*(\alpha_{\mathbf{x}}) = \Sigma'_{\mathbf{J}}\ g_{\mathbf{J}}(\mathbf{t})\ dt^{\wedge \mathbf{J}}$ where $g_{\mathbf{J}}(\mathbf{t}) = \Sigma'_{\mathbf{I}}\ f_{\mathbf{I}}(\boldsymbol{\varphi}(\mathbf{t}))\ \det(R^{\mathbf{I}}{}_{\mathbf{J}})$ for $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ *(10.9.14)*

$\phi^*(\alpha) = \sum_J g_J\, dx_J$ *with*

$$\boxed{g_J = \sum_I \phi^*(f_I)\det(D\phi_{I,J}).}$$ Sja p 44 for $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$

$dx^i \equiv \lambda^i = \langle \mathbf{u}^i| = (\mathbf{u}^i)^T$ $\mathbf{u}^i$ = axis-aligned basis vectors of R^n *(10.1.1)*

$$dx_i = \mathbf{e}_i^T = (0 \ \cdots\ 0\ \ 1\ \ 0\ \cdots\ 0),$$ Sja p 92

$(\alpha_{j_1}\wedge \alpha_{j_2}\wedge \;.....\wedge \alpha_{j_k})(v_{i_1},v_{i_2}.....v_{i_k}) = (1/k!) \det [\, \alpha_{j*}(v_{i*})\,]$ *(8.2.8a)*

$(\alpha_1 \wedge \alpha_2 \wedge \;.....\wedge \alpha_k)\,(v_1, v_2.........v_k) = (1/k!) \det [\, \alpha_*(v_*)\,]$

$$\boxed{\mu_1\mu_2\cdots\mu_k(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k) = \det(\mu_i(\mathbf{v}_j))_{1\le i,j\le k}.}$$

// Sja p 94, Spivak normalization so no (1/k!)

The tensor function pullback

For the pullback of a tensor *function* we have stated

$$[\varphi^*(\alpha_\mathbf{x})](\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_\mathbf{x}(R\mathbf{v}_1,R\mathbf{v}_2...R\mathbf{v}_k) \quad \text{for } \mathbf{x} = \boldsymbol{\varphi}(\mathbf{t}) \qquad \textit{(10.9.5) item 9} \qquad (10.9.20)$$

where recall $[\varphi^*(\alpha_\mathbf{x})](\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \langle\alpha_\mathbf{x} \mid \mathcal{R} \mid \mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k\rangle$. If $\mathcal{R}$ acts to the left, one gets the left side of (10.9.20), while if $\mathcal{R}$ acts to the right one gets the right side. Here the function $\varphi^*(\alpha_\mathbf{x})$ is the pullback of the function $\alpha_\mathbf{x}$ and the pulled-back function is associated with t-space, so we wish to write the right side expression entirely in t-space variables. To this end we replace $\alpha_\mathbf{x}$ by $\alpha_{\varphi(t)}$ on the right of (10.9.20). The tensor functional $[\varphi^*(\alpha_\mathbf{x})]$ is in dual t-space, so we can write it as $[\varphi^*(\alpha_\mathbf{x})]_t$ similar to the β_t appearing in Fig (10.9.3) above. The vectors $\mathbf{v}_i$ are in t-space R^n. We write $R = (D\varphi) = (D^{(t)}\varphi)$ to show that the derivatives are with respect to t. Then in more detail we can write the above tensor function pullback equation as

$$[\varphi^*(\alpha_\mathbf{x})]_t(\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_{\varphi(t)}(\,[D^{(t)}\varphi(\mathbf{t})]\mathbf{v}_1, [D^{(t)}\varphi(\mathbf{t})]\mathbf{v}_2... [D^{(t)}\varphi(\mathbf{t})]\mathbf{v}_k) \quad // \; \mathbf{x} = \boldsymbol{\varphi}(\mathbf{t}) \qquad (10.9.21)$$

where the expression on the right contains only t variables (no x variables), as appropriate for expressing the t-space tensor function $[\varphi^*(\alpha_\mathbf{x})]_t(\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k)$.

Translating (10.9.21) according to $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t}) \;\rightarrow\; \mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ gives

$$[\varphi^*(\alpha_\mathbf{y})]_\mathbf{x}(\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_{\varphi(x)}(\,[D^{(x)}\varphi(\mathbf{x})]\mathbf{v}_1, [D^{(x)}\varphi(\mathbf{x})]\mathbf{v}_2... [D^{(x)}\varphi(\mathbf{x})]\mathbf{v}_k) \quad // \; \mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})\,. \qquad (10.9.22)$$

It is this equation we then compare to Sjamaar's page 96 equation,

$$\boxed{\phi^*(\alpha)_\mathbf{x}(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k) = \alpha_{\phi(\mathbf{x})}(D\phi(\mathbf{x})\mathbf{v}_1, D\phi(\mathbf{x})\mathbf{v}_2,\ldots,D\phi(\mathbf{x})\mathbf{v}_k).} \qquad (10.9.23)$$

He writes $[D^{(x)}\varphi(\mathbf{x})]$ as $D\varphi(\mathbf{x})$ and $[\varphi^*(\alpha_\mathbf{y})]_\mathbf{x}$ as $\varphi^*(\alpha)_\mathbf{x}$.

The tensor function pullback equation also appears in **Spivak** but not quite as we have written it. Spivak says on the top of page 90 and the bottom of page 89,

$$f^*\omega(p)(v_1,\;.\;.\;.\;,v_k) \;=\; \omega(f(p))(f_*(v_1),\;.\;.\;.\;,f_*(v_k)). \qquad f_*(v_p) \doteq (Df(p)(v))_{f(p)}.$$

which we interpret to mean

$$[f^*(\omega)](\mathbf{p})(\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \omega(\mathbf{f}(\mathbf{p}))\;(\,(D^{(p)}f)\mathbf{v}_1,\;(D^{(p)}f)\mathbf{v}_2,\;...\;(D^{(p)}f)\mathbf{v}_k\,)\,.$$

Replacing $\omega\to\alpha$, $f\to\varphi$ and $\mathbf{p}\to\mathbf{x}$ gives

$$[\varphi^*(\alpha)](\mathbf{x})(\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha(\varphi(\mathbf{x}))(\ (D^{(\mathbf{x})}\varphi)\mathbf{v}_1,\ (D^{(\mathbf{x})}\varphi)\mathbf{v}_2,\ ...\ (D^{(\mathbf{x})}\varphi)\mathbf{v}_k\ ) .$$

We then interpret $(\mathbf{x})$ on the left and $(\boldsymbol{\varphi}(\mathbf{x}))$ on the right as spatial locations in the respective non-dual spaces, so the above becomes

$$[\varphi^*(\alpha)]_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) = \alpha_{\varphi(\mathbf{x})}(\ (D^{(\mathbf{x})}\varphi)\mathbf{v}_1,\ (D^{(\mathbf{x})}\varphi)\mathbf{v}_2,\ ...\ (D^{(\mathbf{x})}\varphi)\mathbf{v}_k\ ) \qquad (10.9.24)$$

in agreement with our (10.9.22) and with Sjamaar's form (10.9.23). Sjamaar 2015 refers to $\varphi^*\alpha$ as the pullback of α, but Spivak writing in 1965 does not use the term pullback in his book. Having a name for something is always helpful.

Spivak's entire presentation is in terms of tensor *functions*, there are no functionals *per se*. He uses our tensor function pullback equation as the definition of a pullback (not calling it by that name). We have tried to define the pullback more generally in terms of the general transformation $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ so that it has a meaning for vectors, dual vectors (functionals), and tensor functions.

10.10 Integration of functions over surfaces and curves

In Section 10.11 we are going to make this claim concerning the integration of an arbitrary differential k-form over a manifold "surface" $\mathbf{x'} = \mathbf{F}(\mathbf{x})$ embedded in R^m :

$$\alpha' = \Sigma'_I\ f_I(\mathbf{x'})\,\lambda'^{\wedge I} = \Sigma'_I\ f_I(\mathbf{x'})\ dx'^{i_1} \wedge dx'^{i_2}\ ....\wedge\ dx'^{i_k} \qquad \text{// the k-form in x'-space}$$

$$\int_F \alpha' = \Sigma'_I\ \Sigma'_J \int_{[0,1]^k} f_I(\mathbf{F}(\mathbf{x}))\ \det(R^I{}_J(\mathbf{x}))\ \ dx^{j_1} \wedge dx^{j_2} \wedge ...\wedge\ dx^{j_k} \qquad R= (DF)$$

$$= \Sigma'_I\ \Sigma'_J\ \left(\int_0^1 \int_0^1 ..... \int_0^1 \right) f_I(\mathbf{F}(\mathbf{x}))\ \det(R^I{}_J(\mathbf{x}))\ dx^{j_1}dx^{j_2}....dx^{j_k} . \qquad (10.10.1)$$

The general idea is that the k-form α' in dual x'-space is first pulled back to a different k-form in dual x-space, and then the wedge product $dx^{j_1} \wedge dx^{j_2} \wedge ...\wedge dx^{j_k}$ of basis functionals appearing in this pulled back k-form is mysteriously replaced by a product of ordinary calculus differentials $dx^{j_1}dx^{j_2}....dx^{j_k}$.

The end result is that $\int_F \alpha'$ is some calculus-computable real number. In this notation, the form α' is integrated over a surface F in x'-space determined by $F = \mathbf{F}([0,1]^k)$. We assume that the manifold M can be covered by a single mapping $\mathbf{x'} = \mathbf{F}(\mathbf{x})$, otherwise we create at atlas of mappings as described at the end of Section 10.2.

Before delving into this subject, it seems useful to have some discussion of integrals of functions over surfaces and curves in R^3 without any mention of differential forms. This discussion takes the form of a set of seemingly simple examples.

INTEGRATION OF FUNCTIONS OVER SURFACES

Example 1: Compute the average temperature on a flat plate S' in the z' = 0 plane in R^3 (x'-space). (10.10.2)

What is the meaning of "average temperature"? We partition the plate into a large array of N x N tiny squares of equal area $dA'_i = \Delta x'\Delta y'$, and measure the temperatures T_i simultaneously in all N^2 locations. The average temperature is then $\langle T\rangle = (1/A') \lim_{N\to\infty} [\Sigma_i T_i dA'_i]$ where A' = ab is the area of the plate. So this is an area-weighted average temperature which is cast into a standard-issue 2D Riemann integral,

$$\langle T\rangle = (1/A') \int_{S'} dA'\, T(\mathbf{x'}) = (1/ab) \int_0^a dx' \int_0^b dy'\, T(x',y',0) . \qquad dA' = dx'dy' \qquad (10.10.3)$$

This same kind of integral would be used to compute the average mass density of a flat plate which has areal mass density ρ(x',y'),

$$\langle\rho\rangle = (1/A') \int_{S'} dA'\, \rho(x',y') . \qquad (10.10.4)$$

One could compute the center of mass location of a plate of mass m with similar integrals,

$$\langle x'\rangle = (1/m) \int_{S'} dA'\, x'\, \rho(x',y')$$

$$\langle y'\rangle = (1/m) \int_{S'} dA'\, y'\, \rho(x',y')$$

$$m = \int_{S'} dA'\, \rho(x',y') . \qquad (10.10.5)$$

In this Example we put primes on the variables because they exist in x'-space. There is no need to do any "pulling back" of the area element dA' = dx'dy' to some x-space. The integral is done directly in x'-space.

Example 2: Compute the average normal component of a magnetic field **B** on the same flat plate. (10.10.6)

We are still in x'-space. This problem is similar to the temperature problem with $T \to B_{n'}$ and the result is

$$\langle B_{n'}\rangle = (1/A') \int_{S'} dA'\, B_{n'}(\mathbf{x'}) = (1/A') \int_{S'} dA'\, \mathbf{B}(\mathbf{x'}) \bullet \hat{\mathbf{n}}'$$

$$= (1/A') \int_{S'} d\mathbf{A'} \bullet \mathbf{B}(\mathbf{r'}), \qquad d\mathbf{A'} = dA'\, \hat{\mathbf{n}}' \qquad (10.10.7)$$

where in this problem it happens that

$$\hat{\mathbf{n}}' = \text{unit normal vector, normal to the surface of the plate at } (x',y',0) = \hat{\mathbf{z}}' = \text{constant} .$$

Then

$$<B_{n'}> = <B_{z'}> = (1/ab) \int_0^a dx' \int_0^b dy'\, B_z(x',y',0) \ . \qquad (10.10.8)$$

For both these examples, one could consider a round plate instead of a square plate, and then one would use dA' = dx'dy' → (r)drdθ where the Jacobian J = r appears. One could show that such differential area patches dA' cover the plate surface perfectly with no overlaps and no missed regions.

Example 1a: Compute the average temperature on a spherical shell of radius R in R^3.
Example 2a: Compute the average normal component of a magnetic field **B** on this shell. (10.10.9)

Treating this smooth surface as behaving locally like a flat plate, we use the general expressions (10.10.3) and (10.10.7) obtained for Examples 1 and 2 above,

$$<T> = (1/A') \int_{S'} dA'\, T(\mathbf{x'})$$

$$<B_{n'}> = (1/A') \int_{S'} dA'\, \mathbf{B}(\mathbf{x'}) \bullet \hat{\mathbf{n}}' \ . \qquad (10.10.10)$$

In spherical coordinates, $A' = 4\pi R^2$, $\hat{\mathbf{n}}' = \hat{\mathbf{r}}$ and $dA' = R^2 \sin\theta d\theta d\varphi$. This area measure can be deduced by looking at a picture of spherical coordinates where $dA' = (Rd\theta)(R\sin\theta d\varphi)$ is a surface patch. Then,

$$<T> = (1/4\pi) \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta\, T(R\sin\theta\cos\varphi, R\sin\theta\sin\varphi, R\cos\theta)$$

$$<B_{n'}> = <B_r> = (1/4\pi) \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta\, B_r(R\sin\theta\cos\varphi, R\sin\theta\sin\varphi, R\cos\theta) \ . \qquad (10.10.11)$$

In the language of our earlier sections, we can think of this surface being defined by an underlying transformation

$$\mathbf{x'} = \mathbf{F}(\theta,\varphi) : \qquad \begin{aligned} x' &= R\sin\theta\cos\varphi \\ y' &= R\sin\theta\sin\varphi \\ z' &= R\cos\theta \end{aligned} \qquad (10.10.12)$$

where we would draw "parameter space" = $R^n = R^2$ on the left (with coordinates θ and φ) and $R^m = R^3$ on the right. So one has (θ,φ)-space on the left, and x'-space on the right. In writing $dA' = R^2 \sin\theta d\theta d\varphi$, we are "pulling back" an area patch on the sphere in x'-space to a rectangular area dθdφ in (θ,φ)-space, and we pick up an area conversion factor $R^2\sin\theta$. Similarly, the functions T and B_r are "pulled back" so they are written in the form T(**F**(θ,φ) and B_r(**F**(θ,φ). Although nothing has been said about "differential forms", one suspects that this example can somehow be cast into a 2-form scenario.

Comment: We hope the reader will overlook the fact that if **B** really is a magnetic field, then $<B_{n'}> = 0$ when integrated over any *closed* surface S' (like a spherical shell) due to the divergence theorem and the non-existence of magnetic monopoles, div **B** = 0. The concerned reader can think of **B** as some other vector field.

Example 1b: Compute the average temperature on an *arbitrary smooth surface* S' in x'-space.
Example 2b: Compute the average normal component of a magnetic field **B** on such a surface. (10.10.13)

Start again with (10.10.10),

$$<T> = (1/A')\int_{S'} dA'\, T(\mathbf{x'})$$

$$<B_{n'}> = (1/A')\int_{S'} dA'\; \mathbf{B}(\mathbf{x'}) \bullet \hat{\mathbf{n}}' = (1/A')\int_{S} d\mathbf{A'} \bullet \mathbf{B}(\mathbf{x'}) \;,\quad d\mathbf{A'} = dA'\,\hat{\mathbf{n}}' \;. \qquad \textit{(10.10.10)}$$

The *meaning* of these integrals is clear: dA' is a local area element at point **x'** on the surface, $\hat{\mathbf{n}}'$ is a local unit normal at a point **x'** on the surface, and A' is the total area of the surface. One just has to figure out what these quantities are for a given surface. Notice that $\int_{S'} d\mathbf{A'} \bullet \mathbf{B}(\mathbf{x'})$ is the classic "surface integral of a vector field" as one might encounter in an electrostatic flux calculation (**B** = **E**) or in a fluid flow situation (**B** = **v**).

At point **x**' on the surface there is a tangent space $T_{x'}M$ (Section 10.2) which is spanned by the tangent base vectors $\mathbf{u'}_1$ and $\mathbf{u'}_2$ which appear in the kinematics package (10.6.a.1). Recall that these vectors are generally not orthogonal. The magnitude of the area of the 2-piped subtended by these vectors is $|\mathbf{u'}_1 \times \mathbf{u'}_2|$. But we want a differential 2-piped at point **x**' with some small extents $d\xi_1$ and $d\xi_2$ in these two directions, so then $dA' = |\,(d\xi^1\mathbf{u'}_1) \times (d\xi^2\mathbf{u'}_2)\,| = d\xi^1 d\xi^2\; |\,\mathbf{u'}_1 \times \mathbf{u'}_2|$.

Meanwhile, we know from (10.6.e.2) that $\mathbf{u'}_1 = R\mathbf{u}_1$ and $\mathbf{u'}_2 = R\mathbf{u}_2$, these being vector transformations under **x**' = **F**(**x**). Therefore,

$$R(dx^1\mathbf{u}_1) = dx^1\mathbf{u'}_1$$
$$R(dx^2\mathbf{u}_2) = dx^2\mathbf{u'}_2 \;. \qquad (10.10.14)$$

Thus, the small rectangle spanned by $(dx^1\mathbf{u}_1, dx^2\mathbf{u}_2)$ in x-space is mapped into a small 2-piped spanned by $(dx^1\mathbf{u'}_1, dx'^2\mathbf{u}_2)$ in x'-space. We can take this 2-piped to be the 2-piped discussed above by setting $d\xi_1 = dx_1$ and $d\xi_2 = dx_2$ and then we have $dA' = |\,\mathbf{u'}_1 \times \mathbf{u'}_2\,|\, dx^1 dx^2$.

Recall from (10.6.e.3) that $\mathbf{u'}_3$ is constructed "as needed" so as to form a complete basis for R^3 at point **x**' on the surface S'. We can take $\mathbf{u'}_3 = \mathbf{u'}_1 \times \mathbf{u'}_2$ and then $\mathbf{u'}_3$ can be identified with **n**', a normal vector at point **x**' on the surface. We then need to know that magnitude of this vector to know dA'. Since $R^m = R^3$ is a Cartesian space, up and down vector component indices are the same, so (implied sums)

$$|\,\mathbf{u'}_3|^2 = |\,\mathbf{n'}\,|^2 = |\,\mathbf{u'}_1 \times \mathbf{u'}_2\,|^2 = (\,\mathbf{u'}_1 \times \mathbf{u'}_2) \bullet (\,\mathbf{u'}_1 \times \mathbf{u'}_2) = (\,\mathbf{u'}_1 \times \mathbf{u'}_2)_i(\,\mathbf{u'}_1 \times \mathbf{u'}_2)_i$$

$$= [\varepsilon_{iab}\,(\mathbf{u'}_1)^a(\mathbf{u'}_2)^b]\,[\varepsilon_{icd}\,(\mathbf{u'}_1)^c(\mathbf{u'}_2)^d]$$

$$= \varepsilon_{iab}\varepsilon_{icd}(\mathbf{u'}_1)^a(\mathbf{u'}_2)^b(\mathbf{u'}_1)^c(\mathbf{u'}_2)^d$$

$= (\delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc})\, (\mathbf{u'}_1)^a(\mathbf{u'}_2)^b(\mathbf{u'}_1)^c(\mathbf{u'}_2)^d$ // see e.g. *Tensor* (D.10.22)

$= (\mathbf{u'}_1)^a(\mathbf{u'}_2)^b(\mathbf{u'}_1)^a(\mathbf{u'}_2)^b - (\mathbf{u'}_1)^a(\mathbf{u'}_2)^b(\mathbf{u'}_1)^b(\mathbf{u'}_2)^a$

$= R^a{}_1R^a{}_2\, R^a{}_1R^b{}_2 - R^a{}_1R^b{}_2\, R^b{}_1R^a{}_2$ // kin. package (10.6.a.1) item (e)

$= \Sigma_a (R^a{}_1)^2\ \Sigma_b (R^b{}_2)^2 - (\Sigma_a R^a{}_1R^a{}_2)\ (\Sigma_b R^b{}_1R^b{}_2)$

$= [\Sigma_a (R^a{}_1)^2]\ [\Sigma_a (R^a{}_2)^2] - [\Sigma_a R^a{}_1R^a{}_2]^2$

$\equiv [\ K(\mathbf{x})\ ]^2$ // since $R^i{}_j = R^i{}_j(\mathbf{x})$ in general

or

$$|\mathbf{u'}_3| = |\mathbf{n'}| = K(\mathbf{x}) = \sqrt{[\Sigma_a (R^a{}_1)^2]\ [\Sigma_a (R^a{}_2)^2] - [\Sigma_a R^a{}_1R^a{}_2]^2}\ . \tag{10.10.15}$$

We could have used the vector identity $(\mathbf{A} \times \mathbf{B}) \bullet (\mathbf{A} \times \mathbf{B}) = A^2B^2 - (\mathbf{A}\bullet\mathbf{B})^2$ in place of the εε product method, but εε products are good to know about and we give a reasonable source above for the reader interested in their generalizations. Finally then we have an expression for differential area dA',

$$dA' = |\ \mathbf{u'}_1 \times \mathbf{u'}_2\ |\ dx^1dx^2 = |\ \mathbf{u'}_3|\ dx^1dx^2 = |\ \mathbf{n'}\ |\ dx^1dx^2 = K(\mathbf{x})\ dx^1dx^2\ . \tag{10.10.16}$$

The vector $\mathbf{n'} = \mathbf{u'}_3$ has the following components.

$(\mathbf{n'})_i = (\mathbf{u'}_1 \times \mathbf{u'}_2)_i = \varepsilon_{iab}\, (\mathbf{u'}_1)^a(\mathbf{u'}_2)^b = \varepsilon_{iab}\, R^a{}_1R^b{}_2$ // $R^i{}_j \equiv (\partial x'^i/\partial x^j)$

so

$$(\mathbf{n'})_1 = \varepsilon_{1ab}\, R^a{}_1R^b{}_2 = R^2{}_1R^3{}_2 - R^3{}_1R^2{}_2 = \det\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} = \frac{\partial(x'^2, x'^3)}{\partial(x^1, x^2)}$$

$$(\mathbf{n'})_2 = \varepsilon_{2ab}\, R^a{}_1R^b{}_2 = R^3{}_1R^1{}_2 - R^1{}_1R^3{}_2 = \det\begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix} = \frac{\partial(x'^3, x'^1)}{\partial(x^1, x^2)}$$

$$(\mathbf{n'})_3 = \varepsilon_{3ab}\, R^a{}_1R^b{}_2 = R^1{}_1R^2{}_2 - R^2{}_1R^1{}_2 = \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} = \frac{\partial(x'^1, x'^2)}{\partial(x^1, x^2)}\ . \tag{10.10.17}$$

On the far right we use a common Jacobian-like notation for the 2x2 determinants, where recall from (2.1.2) that $R^i{}_j \equiv (\partial x'^i/\partial x^j)$. We thus obtain this alternate expression for K^2,

$$K^2 = |\ \mathbf{n'}|^2 = \det{}^2\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} + \det{}^2\begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix} + \det{}^2\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix}$$

$$= \left[\frac{\partial(x'^2, x'^3)}{\partial(x^1, x^2)}\right]^2 + \left[\frac{\partial(x'^3, x'^1)}{\partial(x^1, x^2)}\right]^2 + \left[\frac{\partial(x'^1, x'^2)}{\partial(x^1, x^2)}\right]^2 . \tag{10.10.18}$$

From (10.10.17) the vector $\mathbf{n'}$ and the unit normal vector $\hat{\mathbf{n}}' = \mathbf{n'} / |\mathbf{n'}| = \mathbf{n'}/ K$ may then be written,

$$\mathbf{n'} = (\frac{\partial(x'^2, x'^3)}{\partial(x^1, x^2)}, \frac{\partial(x'^3, x'^1)}{\partial(x^1, x^2)}, \frac{\partial(x'^1, x'^2)}{\partial(x^1, x^2)}) = (\det \begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}, \det \begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix}, \det \begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix})$$

$$\mathbf{\hat{n}'} = \frac{1}{K(x)} (\frac{\partial(x'^2, x'^3)}{\partial(x^1, x^2)}, \frac{\partial(x'^3, x'^1)}{\partial(x^1, x^2)}, \frac{\partial(x'^1, x'^2)}{\partial(x^1, x^2)}) . \tag{10.10.19}$$

Using (10.10.18) for dA' the solutions to our problems are,

$$\langle T\rangle = (1/A') \int_{S'} dA'\, T(\mathbf{x'}) = (1/A') \int_S T(\mathbf{F(x)})\, K(\mathbf{x})\, dx^1dx^2$$

$$\langle B_{n'}\rangle = (1/A')\int_{S'} dA'\ \mathbf{B(x')} \bullet \mathbf{\hat{n}'} = (1/A')\int_S \mathbf{B(F(x))} \bullet \mathbf{\hat{n}'}\, K(\mathbf{x})\, dx^1dx^2$$

$$= (1/A')\int_S \mathbf{B(F(x))} \bullet \mathbf{n'}\, dx^1dx^2 \tag{10.10.20}$$

where $\mathbf{\hat{n}'}$ and K(**x**) are as shown above. Notice that the resulting integral is over the region S in x-space which maps into the surface S' in x'-space under **x'** = **F**(**x**). The area A' is given by

$$A' = \int_{S'} dA' = \int_S K(\mathbf{x})\, dx^1dx^2 . \tag{10.10.21}$$

The scalar integral shown in (10.10.20) appears on Buck page 368 (7-3) where T = f, S' = Σ, S = D, and where x^1,x^2 = u,v. The vector integral appears on p 403 where **B** = **F** .

The reader will no doubt notice that in writing dA' = K(**x**) dx^1dx^2 we are in fact "pulling back" some tilted non-rectangular 2-piped patch dA' on the surface S' in x'-space to a rectangular patch dx^1dx^2 in x-space and in doing so we pick up a Jacobian-like factor K(**x**). We are also "pulling back" the integrand functions T(**x'**) and **B**(**x'**) by writing them as T(**F(x)**) and **B**(**F(x)**). Again we arrive at this "pulling back" concept without ever mentioning "differential forms". The pullback integrals shown above are completely well-defined and it is then just a matter of doing the integrals analytically or numerically.

Recall that the square of the area transformation factor K is given by either

$$K^2 = [\Sigma_{a=1}{}^3 (R^a{}_1)^2]\, [\Sigma_{a=1}{}^3 (R^a{}_2)^2] - [\Sigma_{a=1}{}^3 R^a{}_1 R^a{}_2]^2 . \tag*{(10.10.15)}$$

or

$$K^2 = \det^2 \begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} + \det^2 \begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} + \det^2 \begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} . \tag*{(10.10.18)}$$

In the second form K^2 is the sum of the squares of the three 2 x 2 minors of the 3 x 2 "tall" R matrix. See for example Buck page 299 where K = k and $R^i{}_j = a_{ij}$. The two expressions above for K^2 look totally unrelated and it seems strange that they are equal. It turns out that K^2 can be written in yet another way,

$$K^2 = \det(R^T R) \tag{10.10.22}$$

where R^T is the "matrix transpose" of R and not the "covariant transpose" discussed in Section 2.11 (f). Recall from Fig (10.6.c.1) that R^TR is a square n x n matrix and therefore *has* a determinant.

Lest one have doubts, we have Maple compute K^2 in all three ways and show that the three results are the same:

Create a general 3 x 3 R matrix:

```
R := matrix(3,2): print(R);
```

$$\begin{bmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \\ R_{3,1} & R_{3,2} \end{bmatrix}$$

Compute K^2 using (10.10.15) and call it K2a:

```
K2a := sum(R[a,1]^2,a=1..3)*sum(R[a,2]^2,a=1..3) - (sum(R[a,1]*R[a,2],a=1..3))^2;
```

$$K2a := (R_{1,1}^2 + R_{2,1}^2 + R_{3,1}^2)(R_{1,2}^2 + R_{2,2}^2 + R_{3,2}^2) - (R_{1,1} R_{1,2} + R_{2,1} R_{2,2} + R_{3,1} R_{3,2})^2$$

Extract the three 2x2 submatrices from R and call them A,B,C:

```
A := submatrix(R,[1,2],[1,2]);
```

$$A := \begin{bmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \end{bmatrix}$$

```
B := submatrix(R,[1,3],[1,2]);
```

$$B := \begin{bmatrix} R_{1,1} & R_{1,2} \\ R_{3,1} & R_{3,2} \end{bmatrix}$$

```
C := submatrix(R,[2,3],[1,2]);
```

$$C := \begin{bmatrix} R_{2,1} & R_{2,2} \\ R_{3,1} & R_{3,2} \end{bmatrix}$$

Compute K^2 using (10.10.18), call it K2b:

```
K2b := (det(A))^2 + (det(B))^2  + (det(C))^2;
```

$$K2b := (R_{1,1} R_{2,2} - R_{1,2} R_{2,1})^2 + (R_{1,1} R_{3,2} - R_{1,2} R_{3,1})^2 + (R_{2,1} R_{3,2} - R_{2,2} R_{3,1})^2$$

Compute the matrix $R^T R$:

```
RTR := evalm(transpose(R) &* R);
```

$$RTR := \begin{bmatrix} R_{1,1}^2 + R_{2,1}^2 + R_{3,1}^2 & R_{1,1} R_{1,2} + R_{2,1} R_{2,2} + R_{3,1} R_{3,2} \\ R_{1,1} R_{1,2} + R_{2,1} R_{2,2} + R_{3,1} R_{3,2} & R_{1,2}^2 + R_{2,2}^2 + R_{3,2}^2 \end{bmatrix}$$

Compute K^2 using (10.10.22), call it K2c:

```
> K2c := det(RTR);
```

$$K2c := R_{1,1}^2 R_{2,2}^2 + R_{1,1}^2 R_{3,2}^2 + R_{2,1}^2 R_{1,2}^2 + R_{2,1}^2 R_{3,2}^2 + R_{3,1}^2 R_{1,2}^2 + R_{3,1}^2 R_{2,2}^2 - 2 R_{1,1} R_{1,2} R_{2,1} R_{2,2} - 2 R_{1,1} R_{1,2} R_{3,1} R_{3,2} - 2 R_{2,1} R_{2,2} R_{3,1} R_{3,2}$$

Show that all three K^2 expressions are the same:

```
simplify(K2a-K2b);
```

$$0$$

```
simplify(K2a-K2c);
```

$$0$$

Appendix F shows why $K^2 = \det(R^T R)$ in more generality and then Appendix G shows why K^2 may always be written as the sum of the squares of the full-width minors of R as in (10.10.18).

INTEGRATION OF FUNCTIONS OVER CURVES

As much as possible, this section mimics the previous section on integration of surfaces.

Example 3: Compute the average temperature on a piece of straight wire C' of length a in R^3 in x'-space. (10.10.23)

Let $\hat{\mathbf{t}}'$ be a unit vector which is *tangent* to the wire at some point $\mathbf{x}'$ on the wire. Let $d\mathbf{x}'$ be an arbitrary differential distance vector whose tail is located at position $\mathbf{x}'$ on the wire. Then $ds' = d\mathbf{x}' \bullet \hat{\mathbf{t}}'$ is a small distance along the wire. In analogy with the flat plate of Example 1, the length-weighted average temperature of a straight wire is

$$\langle T\rangle = (1/L')\int_{C'} ds'\, T(\mathbf{x}') \qquad ds' = d\mathbf{x}' \bullet \hat{\mathbf{t}}' \ . \tag{10.10.24}$$

In this particular example, the wire is placed on the x' axis so L' = a, $\hat{\mathbf{t}}' = \hat{\mathbf{x}}'$, $ds' = d\mathbf{x}' \bullet \hat{\mathbf{t}}' = dx'$. Then,

$$\langle T\rangle = (1/a) \int_0^a dx'\, T(x',0,0) \ . \tag{10.10.25}$$

Example 4: Compute the average tangential magnetic field **B** on this same straight wire C'. (10.10.26)

This problem is similar to Example 2 (but $\hat{\mathbf{n}}' \to \hat{\mathbf{t}}'$) and the solution is

$$\langle B_{t'}\rangle = (1/L') \int_{C'} ds'\, B_{t'}(\mathbf{x}') = (1/L') \int_{C'} ds'\, \mathbf{B}(\mathbf{x}') \bullet \hat{\mathbf{t}}' , \quad ds' = d\mathbf{x}' \bullet \hat{\mathbf{t}}' = dx'$$

$$= (1/L') \int_{C'} d\mathbf{x}' \bullet \mathbf{B}(\mathbf{x}') , \qquad ds'\, \hat{\mathbf{t}}' = d\mathbf{x}' \ . \tag{10.10.27}$$

so

$$\langle B_{t'}\rangle = (1/a) \int_0^a dx'\, B_{x'}(\mathbf{x}') \ . \tag{10.10.28}$$

Example 3a: Compute the average temperature on a ring of wire C' of radius R in the x',y' plane of R^3.
Example 4a: Compute the average normal component of a magnetic field **B** on this ring. (10.10.29)

Treating this smooth curve as behaving locally like a straight wire, we use the general expressions (10.10.24) and (10.10.27) obtained for Example 3 and 4 above,

$$\langle T\rangle = (1/L')\int_{C'} ds'\, T(\mathbf{x}')$$

$$\langle B_{t'}\rangle = (1/L') \int_{C'} d\mathbf{x}' \bullet \mathbf{B}(\mathbf{x}') \ . \tag{10.10.30}$$

The ring is assumed centered at the origin of the x',y' plane so we use cylindrical coordinates with z' = 0, which then is just polar coordinates, so $\hat{\mathbf{t}}' = \hat{\boldsymbol{\theta}}$, $d\mathbf{x}' = Rd\theta\, \hat{\boldsymbol{\theta}}$, $ds' = d\mathbf{x}' \bullet \hat{\mathbf{t}}' = Rd\theta$, and $L' = 2\pi R$. Then,

$$<T> \ = \ (1/2\pi) \int_0^{2\pi} d\theta\ T(R\cos\theta, R\sin\theta, 0)$$

$$<B_{t'}> \ = (1/2\pi) \int_0^{2\pi} d\theta\ B_\theta(R\cos\theta, R\sin\theta, 0) \qquad (10.10.31)$$

where the last argument of the integrand functions indicates z' = 0 for our placement of the ring in R^3.

Again, the differential distance element $ds' = Rd\theta$ is being "pulled back" from x'-space = R^3 to θ-space = R^1, and the integrand functions are pulled back according to $T(\mathbf{F}(\theta))$ and $B_\theta(\mathbf{F}(\theta))$ where

$$\mathbf{x}' = \mathbf{F}(\theta) : \qquad \begin{aligned} x' &= R\cos\theta \\ y' &= R\sin\theta \\ z' &= 0 \ . \end{aligned} \qquad (10.10.32)$$

Example 3b: Compute the average temperature on an *arbitrary smooth wire* C' in R^3.
Example 4b: Compute the average normal component of a magnetic field **B** on this wire. (10.10.33)

Start again with (10.10.30),

$$<T> = \ (1/L')\int_{C'} ds'\ T(\mathbf{x}') \ = \ (1/L')\int_{C'} d\mathbf{x}' \bullet \hat{\mathbf{t}}'\ T(\mathbf{x}')$$

$$<B_t> = \ (1/L') \int_{C'} ds'\ \mathbf{B}(\mathbf{x}') \bullet \hat{\mathbf{t}}' \ = (1/L') \ \int_{C'} d\mathbf{x}' \bullet \mathbf{B}(\mathbf{x}'), \qquad ds' = d\mathbf{x}' \bullet \hat{\mathbf{t}}' \ . \qquad (10.10.34)$$

The curve C' exists in x'-space $R^m = R^3$ and we take x-space to be $R^n = R^1$. Then curve C in x-space is just the line segment there from $x^1 = 0$ to a and this maps into curve C' under $\mathbf{x}' = \mathbf{F}(x)$. In other words, the curve C' in x'-space is being pulled back to a straight line segment C of x-space. To be consistent, we should be calling the x'-space curve F instead of C', and in Example 2b we should call the surface F instead of S', since in both cases the curve and surface are generated by $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, but we shall sacrifice consistency for clarity.

The most pressing issue now is how to compute the unit tangent vector $\hat{\mathbf{t}}'$. Reaching into our kinematics package (10.6.a.1) and nearby discussion, we realize that

$$\mathbf{t}' \ = \ \mathbf{u}'_1\ , \quad \hat{\mathbf{t}}' \ = \mathbf{u}'_1 \,/\, |\,\mathbf{u}'_1| \ . \qquad (10.10.35)$$

This is because the tangent space $T_{x'}M$ is spanned by the single tangent base vector $\mathbf{u}'_1$, while the other two vectors $\mathbf{u}'_2$ and $\mathbf{u}'_3$ are selected "as needed" to span the perp space to $T_{x'}M$ in R^3. We invent some differential distance dξ so that $d\mathbf{x}' = d\xi\, \mathbf{u}'_1 = \ d\xi\, \mathbf{t}'$ points along the curve C' at point **x'**.

We know from (10.6.e.2) that $\mathbf{u}'_1 = R\mathbf{u}_1$, this being a vector transformation under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Therefore

$$R(dx^1\mathbf{u}_1) = dx^1\mathbf{u'}_1 . \qquad (10.10.36)$$

Thus, the small differential vector $d\mathbf{x} = dx^1\mathbf{u}_1$ in x-space (tangent to C) is mapped into a small differential vector $d\mathbf{x}' = dx^1\mathbf{u'}_1$ in x'-space, tangent to C' at point $\mathbf{x}'$ on C'. Thus we select $d\xi = dx^1$ and conclude that

$$d\mathbf{x'} = dx^1\,\mathbf{u'}_1 \qquad \text{so} \qquad ds' = |d\mathbf{x}'| = |\,dx^1\,\mathbf{u'}_1| = |\,\mathbf{u'}_1|\,dx^1 = |\,\mathbf{t'}\,|\,dx^1 . \qquad (10.10.37)$$

The distance ds' in x'-space is thus being pulled back to distance dx^1 in x-space with scaling factor $|\,\mathbf{t'}\,|$.

The components of the vector $\mathbf{t'} = \mathbf{u'}_1$ are, from (10.6.a.1) item (e),

$$(\mathbf{t'})^i = R^i{}_1(x)$$

$$\mathbf{t'} = (R^1{}_1,\ R^2{}_1,\ R^3{}_1\ ) = (\ (\partial x'^1/\partial x^1), (\partial x'^2/\partial x^1), (\partial x'^3/\partial x^1)\ ) \qquad (10.10.38)$$

and then

$$|\,\mathbf{t'}|^2 = (R^1{}_1)^2 + (R^2{}_1)^2 + (R^3{}_1)^2 = (\partial x'^1/\partial x^1)^2 + (\partial x'^2/\partial x^1)^2 + (\partial x'^3/\partial x^1)^2$$

$$\equiv K^2(x) \qquad \text{// a new and different K from that of Example 2b}$$

or

$$|\,\mathbf{u'}_1| = |\,\mathbf{t'}\,| = K(x) = \sqrt{(R^1{}_1)^2 + (R^2{}_1)^2 + (R^3{}_1)^2} \qquad (10.10.39)$$

and then

$$ds' = |\,\mathbf{t'}|\ dx^1 = K(x)\,dx^1. \qquad \text{// } x = x^1 \qquad (10.10.40)$$

Just as in (10.10.15), (10.10.18) and (10.10.22), the factor $K^2(x)$ can be written three ways,

$$K^2(x) = (R^1{}_1)^2 + (R^2{}_1)^2 + (R^3{}_1)^2 = \Sigma_{a=1}{}^3(R^a{}_1)^2$$

$$K^2(x) = \det{}^2(R^1{}_1) + \det{}^2(R^2{}_1) + \det{}^2(R^3{}_1) \qquad \text{// minors of R are all 1 x 1}$$

$$K^2(x) = \det[R^TR] = R^TR = (R^1{}_1,\ R^2{}_1,\ R^3{}_1\ )\begin{pmatrix} R^1{}_1 \\ R^2{}_1 \\ R^3{}_1 \end{pmatrix} = (R^1{}_1)^2 + (R^2{}_1)^2 + (R^3{}_1)^2$$

or

$$K^2(x) = \det[R^TR] = (R^TR)^1{}_1 = \Sigma_{a=1}{}^3\ (R^T)^1{}_aR^a{}_1 = \Sigma_{a=1}{}^3\ R^a{}_1R^a{}_1 = \Sigma_{a=1}{}^3(R^a{}_1)^2 . \qquad (10.10.41)$$

Here we don't need a Maple program to verify that all three forms give the same result. Note that the "tall" R matrix is the 3x1 matrix shown on the second last line above.

The solutions to our two exercise problems are then (we write $\langle B_{t'}\rangle$ in many equivalent ways),

$$<T> \ = (1/L') \int_{C'} ds'\ T(\mathbf{x'}) \ = (1/L') \int_0^a dx^1\ K(x)\ T(\mathbf{F}(x))$$

$$<B_{t'}> = (1/L') \int_{C'} ds'\ \mathbf{B}(\mathbf{x'}) \bullet \hat{\mathbf{t}}' \ = (1/L') \int_{C'} d\mathbf{x}' \bullet \mathbf{B}(\mathbf{x'})$$

$$= (1/L') \int_C d\mathbf{x} \bullet \mathbf{B}(\mathbf{F}(x)) \ = (1/L') \int_0^a dx^1\ \mathbf{B}(\mathbf{F}(x)) \bullet \mathbf{u}_1 \ = (1/L') \int_0^a dx^1\ \mathbf{B}(\mathbf{F}(x)) \bullet \mathbf{t}'$$

$$= (1/L') \int_0^a dx^1\ B_i(\mathbf{F}(x))\ (\mathbf{t}')^i \ = (1/L') \int_0^a dx^1\ B_i(\mathbf{F}(x))\ R^i{}_1(x) \qquad (10.10.42)$$

where

$d\mathbf{x} = dx^1\mathbf{u}_1$ // below (10.10.36)

$d\mathbf{x}' = dx^1\ \mathbf{u}'_1 = dx^1\ \mathbf{t}'$ // (10.10.37) and (10.10.35)

$ds' = |\ \mathbf{t}'\ |\ dx^1 \ = K(x)\ dx^1$ // (10.10.40)

$L' \ = \int_{C'} ds' \ = \int_0^a dx^1\ K(x)$ = arc length of the curve C' in x'-space . (10.10.43)

Notice that $\int_{C'} d\mathbf{x}' \bullet \mathbf{B}(\mathbf{x'})$ is the classic "line integral of a vector field".

If the variable $x^1 = x$ were time t, then the above $K^2(t) = (\partial x'^1/\partial t)^2 + (\partial x'^2/\partial t)^2 + (\partial x'^3/\partial t)^2$ could be interpreted as the square of the velocity of a particle moving along the curve C',

$$K^2(t) = v'^2_1 + v'^2_2 + v'^2_3 \ = (\mathbf{v}')^2 \qquad \Rightarrow \qquad K(t) = |\ \mathbf{v}'\ | = |\ \partial_t\mathbf{x}'\ | \ = |\ \partial_t\mathbf{F}(t)\ | . \qquad (10.10.44)$$

The scalar integral in (10.10.42) appears on Buck page 367 (7-1) where

$T = f$, $C' = \boldsymbol{\gamma}$, $x = t$, $[0,a] \rightarrow [a,b]$ and $K = |\partial_t\boldsymbol{\gamma}|$

so

$$\int_0^a dx^1\ K(x)\ T(\mathbf{F}(x)) \ \rightarrow \ \int_a^b dt\ \ |\partial_t\boldsymbol{\gamma}|\ f(\boldsymbol{\gamma}(\mathbf{t})) \ . \qquad \text{// Buck 367 (7-1)}$$

The vector integral in (10.10.42) appears for R^2 on Buck page 376 (7-7) where

$\mathbf{B} = (A,B)$, $(x'^1, x'^2) = (\varphi,\psi)$, $\mathbf{t}' = (\ (\partial x'^1/\partial t), (\partial x'^2/\partial t)\ ) = (\partial_t\varphi,\partial_t\psi)$, and $[0,a] \rightarrow [a,b]$

so

$$<B_{t'}> = \int_0^a dx^1\ \mathbf{B}(\mathbf{F}(x)) \bullet \mathbf{t}' \ \rightarrow \ \int_a^b dt\ [\ A(\boldsymbol{\gamma}(t))\ (\partial_t\varphi) \ + B(\boldsymbol{\gamma}(t))\ \partial_t\psi\ ] . \qquad \text{// Buck 376 (7-7)}$$

Comments on the above examples

As will be seen formally in the next section, the surface and curve integrations discussed above fall into the realm of 2-form and 1-form integrations. In the above examples, there was no mention of "functionals" or "dual spaces" or "wedge products" or "cosmetic notation" or even of "differential forms". No mention was made of "surface orientation". The calculations were performed on an *ad hoc* basis as any journeyman might approach these problems. There was, however, some discussion of "pulling back" integrand functions and differential areas and differential lengths from R^m to R^n, but there was no mention of pulling back functionals between the corresponding dual spaces.

The method of differential forms provides a systematic method for doing integrations over "surfaces" (manifolds) of any dimension embedded in a space of any same or larger dimension, where the spaces can have arbitrary metric tensors, and where orientation is tracked.

10.11 Integration of differential k-forms over Surfaces

Using our cosmetic notation for k-form functionals, we write

$$\alpha_{\mathbf{x}'} = \Sigma'_I f_I(\mathbf{x}')\, dx'^{\wedge I} \qquad \text{// original k-form in } \Lambda'^k(R^m) \qquad (10.8.15)\ 6$$

$$F^*(\alpha_{\mathbf{x}'}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \det(R^I{}_J)\, dx^{\wedge J} \qquad \text{// k-form pulled back into } \Lambda^k(R^n)\ . \qquad (10.8.19)$$

The pulled back k-form can be rewritten compactly as

$$\beta_{\mathbf{x}} \equiv F^*(\alpha_{\mathbf{x}'}) = \Sigma'_J g_J(\mathbf{x})\, dx^{\wedge J} \qquad \text{where} \qquad g_J(\mathbf{x}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\det(R^I{}_J)\ . \qquad (10.11.1)$$

Besides grouping terms into $g_J(\mathbf{x})$ we have assigned the name $\beta_{\mathbf{x}}$ to the pulled-back k-form. This pulled-back k-form is written out in detail in (10.8.3) with a display of the $R^I{}_J$ matrix in (10.8.4)

The point $\mathbf{x}'$ lies on a "surface" in x'-space which is generated by a defining transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. A region we shall call S in x-space maps into a region S' on the manifold as shown in Fig (10.2.1) with S=U and S'=V. The letter S suggests the word "Surface".

The integral of the original k-form $\alpha_{\mathbf{x}'}$ over surface S' is then set equal to the integral of the pulled-back k-form $F^*(\alpha_{\mathbf{x}'})$ over the pulled-back surface S,

$$\int_{S'} \alpha_{\mathbf{x}'} \equiv \int_S \beta_{\mathbf{x}}\ . \qquad (10.11.2)$$

Below we shall refer to this as our **first definition**, as if the right side defines the meaning of the left side. That is to say, the integral of a k-form $\alpha_{\mathbf{x}'}$ over some complicated surface (manifold) in x'-space is reduced to an integral of a different k-form $\beta_{\mathbf{x}}$ over a relatively simple surface in x-space. This is reminiscent of our examples in Section 10.10 where we had, for example,

$$\int_{S'} dA'\ \mathbf{B}(\mathbf{x}') \bullet \hat{\mathbf{n}}' = \int_S \mathbf{B}(\mathbf{F}(\mathbf{x})) \bullet \hat{\mathbf{n}}'\, K(\mathbf{x})\, dx^1 dx^2 \qquad (10.10.20)$$

where $\int_{S'}$ is over an arbitrary surface in x'-space while $\int_S$ is a straightforward integral in Cartesian x-space. The big difference however is that in Section 10.10 we were dealing with calculus integrals, whereas here we are dealing with k-forms which are functionals in certain dual spaces. In the case of the calculus integral examples, one can regard the shift from x'-space to x-space as nothing more than a "change of variables" and there is no "first definition" of anything.

We then come to our **second definition** which is this:

$$\int_S \beta_{\mathbf{x}} = \int_S [\, \Sigma'_J\, g_J(\mathbf{x})\, dx_{\wedge}^{J} \,] \equiv \int_S \Sigma'_J\, g_J(\mathbf{x})\, dx^{j_1} dx^{j_2} \ldots dx^{j_k} \qquad (10.11.3)$$

and one ends up the a well-defined multivariable calculus integral.

One must ask: how is it that the functional

$$dx_{\wedge}^{J} = \lambda_{\wedge}^{J} = \lambda^{j_1} \wedge \lambda^{j_2} \wedge \ldots \wedge \lambda^{j_k} = dx^{j_1} \wedge dx^{j_2} \wedge \ldots \wedge dx^{j_k}$$

disappears and is replaced by $dx^{j_1}dx^{j_2} \ldots dx^{j_k}$? This transition is not so easy to detect in some sources because the wedge product hats are suppressed and dx^j and dx^j are typeset identically.

One answer to this question is the following. One writes

$$\beta_{\mathbf{x}}(S) \equiv \int_S \beta_{\mathbf{x}} \equiv \int_S \Sigma'_J\, g_J(\mathbf{x})\, dx^{j_1} dx^{j_2} \ldots dx^{j_k} \ . \qquad (10.11.4)$$

In this point of view, one regards the k-form $\beta_{\mathbf{x}}$ as a functional which acts on regions of R^n to produce a real number so there is a mapping (a "functional" is any mapping to the reals),

$$\beta_{\mathbf{x}} : \ S \subset R^n \ \rightarrow \ R \ . \qquad (10.11.5)$$

This is a different kind of functional from the functional $dx_{\wedge}^{J} = \lambda_{\wedge}^{J} \in \Lambda^k(R^n)$ which is a vector in the dual space shown. Whereas $dx_{\wedge}^{J}$ is a linear functional, $\beta_{\mathbf{x}}(S)$ is not a linear functional, for example, since doubling the region S is not likely to double the resulting real number $\beta_{\mathbf{x}}(S)$.

This seems to be the approach taken by Loring Tu, where we quote from his p 263,

Definition 23.8. Let $\omega = f(x)\,dx^1 \wedge \cdots \wedge dx^n$ be a C^∞ n-form on an open subset $U \subset \mathbb{R}^n$, with standard coordinates $x^1, \ldots, x^n$. Its *integral* over a subset $A \subset U$ is defined to be the Riemann integral of $f(x)$:

$$\int_A \omega = \int_A f(x)\,dx^1 \wedge \cdots \wedge dx^n := \int_A f(x)\,dx^1 \cdots dx^n, \qquad (10.11.6)$$

Here he is stating our "second definition". Buck also takes this approach, referring in his Definition on page 381 to a k-form ω as a "region-functional". He writes as a 3-form example,

$\omega = A(x,y,z)\, dxdydz$ // meaning $\omega = A(x,y,z)\, dx\wedge dy\wedge dz$

$\omega(\Omega) = \int\int\int_{\Omega} A(x,y,z)\, dxdydz$ Ω = a region in the definition domain of ω

Arm-waving comment: We know that an exterior derivative increases the rank of a k-form by one unit. It is not unreasonable then to say that a k-fold integration of a k-form reduces the rank of that k-form by k units down to rank 0, which is a scalar function and in the above situation just a number, the value of the integral.

In any event, the end result for the integral of a k-form over a manifold region S' in x'-space is this:

$$\int_{S'} \alpha_{\mathbf{x}'} = \int_{S'} [\Sigma'_I f_I(\mathbf{x}')\, dx'^{i_1} \wedge dx'^{i_2} \wedge \ldots \wedge dx'^{i_k}] \qquad // \alpha_{\mathbf{x}'} = \Sigma'_I f_I(\mathbf{x}')\, dx'^{\wedge I}$$

$$\equiv \int_{S} [\Sigma'_J g_J(\mathbf{x})\, dx^{j_1} \wedge dx^{j_2} \wedge \ldots \wedge dx^{j_k}] \qquad // \text{first definition (pullback)}$$

$$\equiv \int_{S} [\Sigma'_J g_J(\mathbf{x})\, dx^{j_1} dx^{j_2} \ldots dx^{j_k}] \qquad // \text{second definition}$$

where

$$g_J(\mathbf{x}) = \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\det(R^I{}_J) \quad \text{and} \quad \mathbf{x}' = \mathbf{F}(\mathbf{x}),\ R = (DF)\,. \qquad (10.11.7)$$

As we shall see in Section 10.12, this specification reproduces the "journeyman" integration results shown in the examples of Section 10.10. Recall that Σ'_I and Σ'_J are ordered sums.

An Alternate Approach

In this thread we take a narrower view of the functional sense of the k-form integral. We treat $\int_{S'} \alpha_{\mathbf{x}'}$ as if it were a discrete sum over the points **x**' on the surface S'. Since $\alpha_{\mathbf{x}'}$ is a certain functional, the integral $\int_{S'} \alpha_{\mathbf{x}'}$ is then also a functional, being a sum of functionals. In some sense the analysis below is a microscale interpretation of the region-functional approach noted above. The development below is done in Dirac notation, but it could be restated using the pullback function F* .

In a first step, we write the pullback of the original k-form "sum" as,

$$[\int_{S'} \alpha_{\mathbf{x}'}]\, \mathcal{R} = \int_{S} \beta_{\mathbf{x}} \qquad (10.11.8)$$

where $\mathcal{R}$ is the Dirac pullback operator used in Section 10.7. Recall that the pulled-back k-form is,

$$\beta_{\mathbf{x}} = \Sigma'_J [g_J(\mathbf{x})]\, dx^{\wedge J} = \Sigma'_J [g_J(\mathbf{x})]\, \lambda^{\wedge J} \qquad // \text{(10.11.1) and definition (10.1.9) that } dx^{\wedge J} \equiv \lambda^{\wedge J}$$

$$= \Sigma'_J [\Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\det(R^I{}_J)]\, \lambda^{\wedge J} \qquad // \text{insert } g_J(\mathbf{x}) \text{ from (10.11.1)}$$

$= \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\ [\Sigma'_J \det(R^I{}_J)\ \lambda^{\wedge J}]$ // reorder

$= \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\ [\Sigma_J R^I{}_J \lambda^{\wedge J}]$ // (10.8.1) to get symmetric J sum and no det

$= \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\ [\Sigma_M R^I{}_M \lambda^{\wedge M}]$. // rename dummy multiindex J→M (10.11.9)

Since $\lambda^{\wedge M} = \langle \mathbf{u}^{\wedge M}|$ from (2.11.c.2), we rewrite (10.11.8) in Dirac notation,

$$\langle \int_{S'} \alpha_{\mathbf{x}'} |\ \mathcal{R} = \int_S \langle \beta_{\mathbf{x}}| = \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\ \Sigma_M R^I{}_M \langle \mathbf{u}^{\wedge M}| \tag{10.11.10}$$

which is interpreted as a functional in $\Lambda^k(R^n)$ (see fiber comment below).

In a second step we close this functional with a certain "measure ket" $|\mu\rangle$ defined as

$$|\mu\rangle \equiv \Sigma'_J |\mathbf{dx}^J\rangle = \Sigma'_J |\mathbf{dx}^{j_1}\rangle \otimes |\mathbf{dx}^{j_2}\rangle.\otimes ... \otimes |\mathbf{dx}^{j_k}\rangle$$

$$= \Sigma'_J |\ \mathbf{dx}^{j_1}, \mathbf{dx}^{j_2} ... \mathbf{dx}^{j_k}\rangle \tag{10.11.11}$$

where the differential vectors are aligned with the axes of x-space R^n,

$$\mathbf{dx}^j \equiv dx^j\, \mathbf{u}_j \quad \text{// no sum on j} \qquad \text{or} \qquad |\mathbf{dx}^j\rangle = dx^j\,|\mathbf{u}_j\rangle\ . \tag{10.11.12}$$

Here $\mathbf{dx}^j$ is a vector in R^n and $|\mathbf{dx}^J\rangle$ is a vector in $(R^n)^k$ called V^k in Chapter 5. Thus,

$$|\mu\rangle = \Sigma'_J dx^{j_1}dx^{j_2} ... dx^{j_k}|\ \mathbf{u}_{j_1}, \mathbf{u}_{j_2} ... \mathbf{u}_{j_k}\rangle$$

$$= \Sigma'_J dx^J\,|\mathbf{u}_J\rangle \quad \text{// multiindex notation,} \quad dx^J \equiv dx^{j_1}dx^{j_2} ... dx^{j_k}\,. \tag{10.11.13}$$

For example, for k = 1,2,3 in $R^n = R^3$ the vector $|\mu\rangle$ would be

$|\mu\rangle = |\mathbf{dx}^1\rangle + |\mathbf{dx}^2\rangle + |\mathbf{dx}^3\rangle$ // k = 1

$|\mu\rangle = |\mathbf{dx}^1, \mathbf{dx}^2\rangle + |\mathbf{dx}^1, \mathbf{dx}^3\rangle + |\mathbf{dx}^2, \mathbf{dx}^3\rangle$ // k = 2

$|\mu\rangle = |\mathbf{dx}^1, \mathbf{dx}^2, \mathbf{dx}^3\rangle$. // k = 3 (10.11.14)

We then define "the integral of the k-form $\alpha_{\mathbf{x}'}$ in x'-space" as follows,

$$\text{"}\int_{S'} \alpha_{\mathbf{x}'}\text{"} \equiv \langle \int_{S'} \alpha_{\mathbf{x}'} |\ \mathcal{R}\,|\mu\rangle = \langle \int_S \beta_{\mathbf{x}}|\ \mu\rangle = \int_S \langle\beta_{\mathbf{x}}|\ \mu\rangle \tag{10.11.15}$$

where we end up with the integral of a certain tensor function over S. Next, write

$$\int_S <\beta_{\mathbf{x}} \,|\, \mu> \;=\; \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma_M R^I{}_M \;<\mathbf{u}^{\wedge M} \,|\, \mu> \qquad // (10.11.10)$$

$$= \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma_M R^I{}_M \;\; \Sigma'_J <\mathbf{u}^{\wedge M} \,|\, \mathbf{dx}^J> \qquad // (10.11.11)$$

$$= \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma_M R^I{}_M \;\; \Sigma'_J \, dx^J \;<\mathbf{u}^{\wedge M} \,|\, \mathbf{u}_J> \;. \qquad // (10.11.13) \qquad (10.11.16)$$

In our Chapter 8 normalization for wedge products, we write

$$<\mathbf{u}^{\wedge M} \,|\, \mathbf{u}_J> \;=\; (\lambda^{m_1} \wedge \lambda^{m_2} \wedge ... \wedge \lambda^{m_k})\, (\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, ... \mathbf{u}_{j_k}) \qquad // (2.11.c.2)\ \lambda^i \equiv <\mathbf{u}^i|$$

$$= (1/k!)\, \det(\delta^M{}_J) \;. \qquad // (8.3.9.b) \qquad (10.11.17)$$

In the Spivak normalization of the wedge product (see below (8.1.3) the (1/k!) is replaced by 1, and we shall now continue in the Spivak normalization, so

$$<\mathbf{u}^{\wedge M} \,|\, \mathbf{u}_J> \;=\; \det(\delta^M{}_J) \;. \qquad // \text{ using Spivak wedge product normalization} \qquad (10.11.18)$$

Inserting this into (10.11.16) gives

$$\int_S <\beta_{\mathbf{x}} \,|\, \mu> \;=\; \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \, dx^J \;\; [\Sigma_M R^I{}_M \, \det(\delta^M{}_J)] \qquad (10.11.19)$$

where we have shifted the M sum to the right. Now consider,

$$\Sigma_M \; R^I{}_M \,[\det(\delta^M{}_J\,)]$$

$$= \Sigma_M \; R^I{}_M \,[\Sigma_P\, (-1)^{S(P)} \; \delta^M{}_{P(J)} \;] \qquad // (A.1.21)$$

$$= \Sigma_P\, (-1)^{S(P)} \; \Sigma_M \; R^I{}_M \, \delta^M{}_{P(J)} \qquad // \text{ reorder}$$

$$= \Sigma_P (-1)^{S(P)} \; R^I{}_{P(J)} \qquad // \text{ k matrix multiplications}$$

$$= \det(R^I{}_J) \;. \qquad // (A.1.21) \qquad (10.11.20)$$

Inserting this result into (10.11.19) gives

$$\int_S <\beta_{\mathbf{x}} \,|\, \mu> \;=\; \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \, dx^J \; \det(R^I{}_J)]$$

$$= \int_S \; \Sigma'_I f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \; \det(R^I{}_J)\, dx^J \;. \qquad (10.11.21)$$

The final result then is

$$" \int_{S'} \alpha_{\mathbf{x}'} " \equiv < \int_{S'} \alpha_{\mathbf{x}'} \mid \mathcal{R} \mid \mu > = < \int_S \beta_{\mathbf{x}} \mid \mu > = \int_S < \beta_{\mathbf{x}} \mid \mu >$$

$$= \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x})) \Sigma'_J \det(R^I{}_J) dx^{j_1} dx^{j_2} \ldots dx^{j_k}$$

$$= \int_S g_J(\mathbf{x}) dx^{j_1} dx^{j_2} \ldots dx^{j_k} \quad . \tag{10.11.22}$$

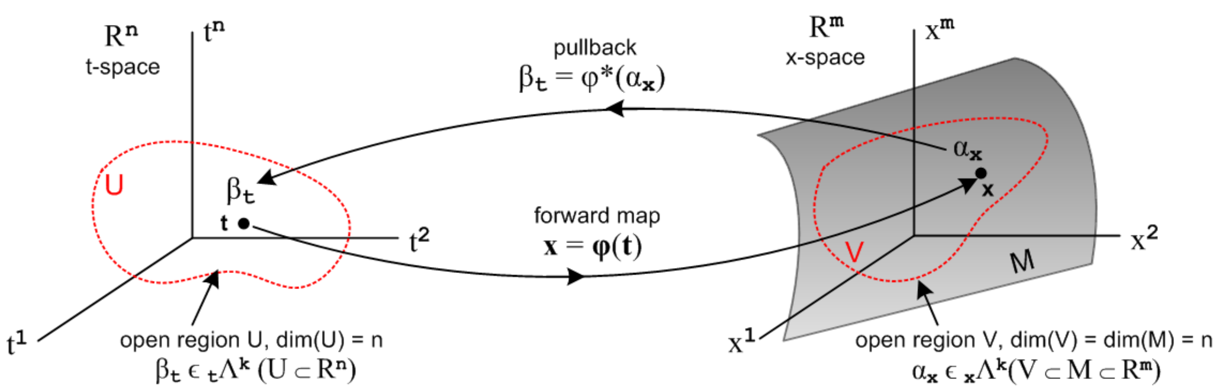
This result then is the same as (10.11.7) obtained by making the "two definitions". Our resulting tensor function turns out to be just a constant function which is a real number which is the result of doing the above regular calculus multivariable integration.

Our alternate approach lacks rigor since the integration is treated as a sum over points **x**' on a manifold and this really means that the Dirac space used above is some kind of fiber bundle space (the tangent bundle of Section 10.2). Moreover, the measure ket $|\mu> = \Sigma'_J \mid \mathbf{dx}^J>$ seems arbitrary, but it does manage to "sweep up" all contributions to the integration and we do get the correct result. The method does at least provide an alternative explanation of how the functional $dx\wedge^J$ wedge product is replaced by the calculus product dx^J.

10.12 Integration of 1-forms

General Review of k-form integration

This section is presented in the $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ notation introduced in Section 10.9 and illustrated in hybrid Fig (10.9.3) which we replicate here,



(10.9.3)

The main result of Section 10.11 is this description of the integration of a k-form over a surface,

$$\int_{S'} \alpha_{x'} \ = \int_{S'}[\Sigma'_I f_I(x')\, dx^{i1} \wedge dx^{i2} \wedge ... \wedge dx^{ik}\,] \qquad // \alpha_{x'} = \Sigma'_I f_I(x')\, dx'^{\wedge I}$$

$$\equiv \int_S [\,\Sigma'_J g_J(x)\, dx^{j1} \wedge dx^{j2} \wedge ... \wedge dx^{jk}\,] \qquad // \text{ first definition (pull back)}$$

$$\equiv \int_S [\,\Sigma'_J g_J(x)\, dx^{j1}dx^{j2} ... dx^{jk}] \qquad // \text{ second definition}$$

where

$$g_J(x) = \Sigma'_I f_I(F(x)) \det(R^I{}_J) \quad \text{and} \quad x' = F(x)\ ,\ R = (DF)\ . \qquad (10.11.7)$$

Using the $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ notation we rewrite the above (with some specialization) as,

$$\int_\varphi \alpha_x \ = \int_\varphi \Sigma'_I f_I(x)\, dx^{i1} \wedge dx^{i2} \wedge ... \wedge dx^{ik} \qquad // \alpha_x = \Sigma'_I f_I(x)\, dx^{\wedge I}$$

$$\equiv \int_{[0,1]^k} \Sigma'_J g_J(t)\, dt^{j1} \wedge dt^{j2} \wedge ... \wedge dt^{jk} \qquad // \text{ first definition (pull back)}$$

$$\equiv (\int_0^1 \int_0^1 ... \int_0^1)\, \Sigma'_J g_J(t)\, dt^{j1}dt^{j2} ... dt^{jk} \qquad // \text{ second definition}$$

where

$$g_J(t) = \Sigma'_I f_I(\varphi(t)) \det(R^I{}_J) \quad \text{and} \quad x = \varphi(t)\ ,\ R = (D\varphi)\ . \qquad (10.12.1)$$

Here the pulled-back integration region formerly called S is taken to be the unit cube in k dimensions, written above as $[0,1]^k$ and referred to as a **k-cube**. The pre-pullback integration region formerly called S' is here called φ, with the idea that this region is $\boldsymbol{\varphi}([0,1]^k)$.

Note: A **k-chain** is a linear combination of k-cubes and is used by both Sjamaar (p 65) and Spivak (p 97) in their derivations of Stokes' Theorem. In fact, Spivak's entire Chapter 4 which includes his discussion of tensor products, wedge products and tensor functions is entitled *Integration on Chains*.

Integration of 1-forms

We wish now to look in more detail at the integration of 1-forms. There is much repetition of statements below because the meaning of objects tends to quietly diffuse away as one proceeds.

Consider this general 1-form in x-space R^m ,

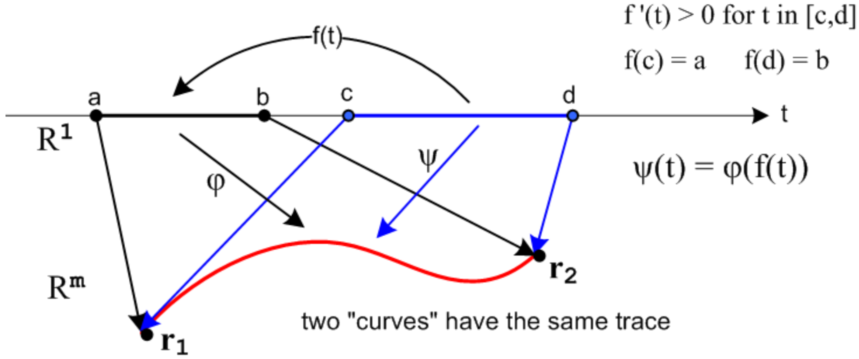
$$\alpha_x = \Sigma_i f_i(x)\, {}_x\lambda^i = \Sigma_i f_i(x)\, dx^i \quad . \qquad (10.12.2)$$

We wish to define a meaning for the integration of this 1-form α_x over a piece of the curve $\mathbf{x} = \boldsymbol{\varphi}(t)$,

$$\int_\varphi \alpha_x = \int_\varphi \Sigma_i f_i(x)\, dx^i \quad = \text{integral of a 1-form over a piece of the curve } \boldsymbol{\varphi} \text{ in } R^m \ . \qquad (10.12.3)$$

The transformation $\mathbf{x} = \boldsymbol{\varphi}(t)$ is a mapping $\boldsymbol{\varphi}$: $t \rightarrow R^m$. Variable t is often called "the parameter".

Comment: Officially it is the mapping $\boldsymbol{\varphi}$ which is "the curve", but one loosely refers to the image (trace) of this mapping in R^m as "the curve". The distinction is necessary because many mappings can have the same image curve, such as $\boldsymbol{\varphi}(t)$ and $\boldsymbol{\varphi}(t^2)$, where the parameter is "re-speeded" (reparametrized). This picture shows the general respeeding idea :



(10.12.4)

Here the same red curve is the image of two different transformations $\mathbf{x} = \boldsymbol{\varphi}(t)$ and $\mathbf{x} = \boldsymbol{\psi}(t)$ with different domain intervals, and $\boldsymbol{\psi}(t) = \boldsymbol{\varphi}(f(t))$ where f(t) is a monotonic respeeding function. A special case would be [a,b] = [c,d] = [0,1] to which our example $\varphi(t)$ and $\psi(t) = \varphi(t^2)$ would apply. Mappings φ and ψ are called smoothly equivalent curves and $\int_{\boldsymbol{\varphi}} \alpha_{\mathbf{x}}$ is the same for any two such curves (Buck p 386 Theorem 2 (i)). A similar but generalized reparametrization comment applies to integration of 2-forms and k-forms.

So imagine that we have a curved line hanging in R^m space and as t varies perhaps from 0 to 1 in t-space, we move along the image curve in R^m. The problem is how to integrate a 1-form along this curve.

We can define the calculational meaning of the above integral in two steps, each being a definition, as outlined in Section 10.11.

First definition:

$$\int_{\boldsymbol{\varphi}} \alpha_{\mathbf{x}} \equiv \int_{[0,1]} \varphi^*(\alpha_{\mathbf{x}})$$

$$= \text{the integral in t-space of the pullback of } \alpha_{\mathbf{x}} \text{ over the 1-cube } [0,1] \qquad (10.12.5)$$

On the left is an integral of the 1-form $\alpha_{\mathbf{x}}$ over a curve $\boldsymbol{\varphi}$ in R^m.
On the right is an integral of a *different* 1-form $\varphi^*(\alpha_{\mathbf{x}})$ (the pullback of $\alpha_{\mathbf{x}}$) over a 1-cube [0,1] in R^1.
Note that $\alpha_{\mathbf{x}}$ lies in ${}_{\mathbf{x}}\Lambda^1(R^m)$ while $\varphi^*(\alpha_{\mathbf{x}})$ lies in ${}_{t}\Lambda^1(R)$.
Since our usual 1-form pullback mapping is $\varphi^* : {}_{\mathbf{x}}\Lambda^1(R^m) \rightarrow {}_{t}\Lambda^1(R^n)$, we have n = 1 (see (10.7.18)).

The "tall" m x n R-matrix for this problem is then an m x 1 matrix which is just a column vector of m elements $\partial_t\varphi^i$,

$$R^i{}_1 = (D^{(t)}\varphi)^i{}_1 = \partial\varphi^i(t)/\partial t . \qquad // \ t^1 \equiv t, \text{ the only coordinate in t-space} \qquad (10.12.6)$$

We then compute the pullback of $\varphi^*(\alpha_{\mathbf{x}})$ of $\alpha_{\mathbf{x}}$:

$$\alpha_{\mathbf{x}} = \Sigma_i f_i(\mathbf{x})\, {}_{\mathbf{x}}\lambda^i = \Sigma_i f_i(\mathbf{x})\, dx^i = \mathbf{f}(x) \bullet \mathbf{dx}\,, \qquad \mathbf{dx} \equiv (dx^1, dx^2, \ldots.\, dx^m)$$

$$\varphi^*(\alpha_{\mathbf{x}}) = \Sigma_i \varphi^*(f_i(\mathbf{x}))\, \varphi^*(dx^i) \qquad // \ (10.9.5)\ 3$$

$$= \Sigma_i f_i(\boldsymbol{\varphi}(t))\ \Sigma_{j=1}^{n} R^i{}_j\, dt^j \qquad // \ (10.9.5)\ 1 \text{ and } 5$$

$$= \Sigma_i f_i(\boldsymbol{\varphi}(t))\ R^i{}_1\, dt^1 \qquad // \ n = 1$$

$$= \Sigma_i f_i(\boldsymbol{\varphi}(t))\, [\partial\varphi^i(t)/\partial t]\ dt \qquad // \ (10.12.6) \text{ and } t^1 = t$$

$$= g(t)\, dt \qquad (10.12.7)$$

where

$$g(t) \equiv \Sigma_i f_i(\boldsymbol{\varphi}(t))\, [\partial\varphi^i(t)/\partial t] = \Sigma_i f_i(\boldsymbol{\varphi}(t))\, \partial_t\varphi^i(t) = \mathbf{f}(\boldsymbol{\varphi}(t)) \bullet (\partial_t\boldsymbol{\varphi})\,. \qquad (10.12.8)$$

The object $\varphi^*(\alpha_{\mathbf{x}}) = g(t)\, dt$ is a 1-form in dual t-space ${}_t\Lambda^1(R)$.

Using the definition given above, one then has,

$$\int_{\varphi} \alpha_{\mathbf{x}} = \int_{\varphi} \alpha_{\mathbf{x}} = \int_{[0,1]} \varphi^*(\alpha_{\mathbf{x}}) = \int_{[0,1]} g(t)\, dt\,. \qquad (10.12.9)$$

Thus the integral of the 1-form $\alpha_{\mathbf{x}}$ over the curve $\boldsymbol{\varphi}$ in x-space is defined to be equal to the integral of the 1-form $g(t)\, dt$ over a 1-cube in t-space. So far no regular calculus integrals have appeared.

Second definition:

$$\int_{[0,1]} g(t)\, {}_t\lambda = \int_{[0,1]} g(t)\, dt \equiv \int_0^1 g(t)\, dt\ . \qquad (10.12.10)$$

On the left is the integral of a 1-form on a 1-cube, on the right is an ordinary calculus integral of a function over the interval [0,1] of the real axis. It is this second definition that motivates giving the dual space basis vector ${}_t\lambda$ the cosmetic name *dt*.

If one flips the "orientation" of the integration domain, so that [0,1] becomes [1,0], the result changes sign, and of course this fact agrees with the usual notion that $\int_1^0 g(t)\, dt = - \int_0^1 g(t)\, dt$.

We have then shown that the integral of a 1-form is described by,

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \mathbf{f(x)} \bullet \boldsymbol{dx} = \int_{[0,1]} g(t)\, dt$$

$$= \int_0^1 \mathbf{f}(\boldsymbol{\varphi}(t)) \bullet (\partial_t\boldsymbol{\varphi}(t))\, dt = \int_0^1 f_i(\boldsymbol{\varphi}(t))\, [\partial\varphi^i(t)/\partial t]\, dt \,. \qquad (10.12.11)$$

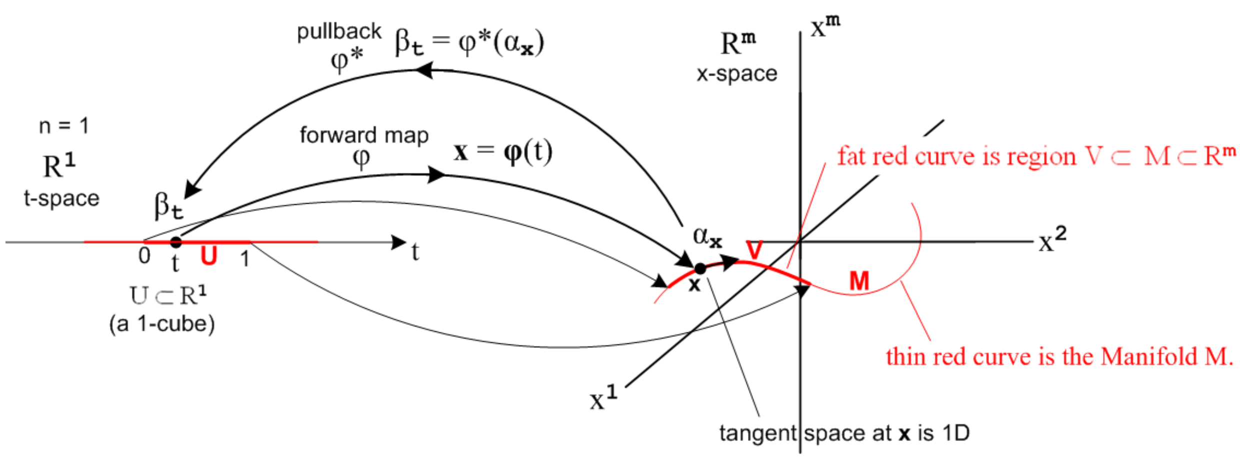
This result appears in Sjamaar Ch 4 Eq (4.1) with **φ = c** and m = n.

Notice that the only locations where **f(x)** is "sensed" in this integral are points on the curve **x** = **φ**(t).

Since d**x** = (∂_t**φ**(t)) dt, the above can be written concisely as

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \mathbf{f(x)} \bullet \boldsymbol{dx} = \int_0^1 \mathbf{f}(\boldsymbol{\varphi}(t)) \bullet d\mathbf{x} \qquad \text{where } d\mathbf{x} = (\partial_t\boldsymbol{\varphi}(t))\, dt \,. \qquad (10.12.12)$$

We redisplay the earlier Fig (10.9.3b) to illustrate the above discussion, where β_t = g(t) *dt* :



(10.12.13)

Recall now our "no differential forms" integration done in (10.10.42),

$$L'\langle B_t\rangle = \int_0^a dx\, \Sigma_{j=1}^{3}\ B_i(\mathbf{F}(x))\, R^i{}_1(x) \,. \qquad // B_t \text{ means } B_{tangent} \qquad (10.12.14)$$

In the **x** = **φ**(t) notation this reads, setting a = 1 and replacing 3 by m,

$$L\langle B_t\rangle = \int_0^1 dt\, \Sigma_{j=1}^{m}\ B_i(\boldsymbol{\varphi}(t))\, (D\varphi)^i{}_1(t)$$

$$= \int_0^1 dt\ \Sigma_{j=1}^{m}\ B_i(\boldsymbol{\varphi}(\mathbf{t}))\, \partial^i\varphi(t)$$

$$= \int_0^1 \mathbf{B}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \partial\boldsymbol{\varphi}(t)\, dt \qquad (10.12.15)$$

which is the same integral appearing in (10.12.11) with **f** = **B**. Therefore, we can interpret (10.12.15) as being the integral of the 1-form,

$$\alpha_{\mathbf{x}} = \Sigma_i B_i(\mathbf{x})\, {}_{\mathbf{x}}\lambda^i = \Sigma_i B_i(\mathbf{x})\, dx^i = \mathbf{B}(\mathbf{x}) \bullet \mathbf{dx} \qquad (10.12.16)$$

and one then has

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \mathbf{B}(\mathbf{x}) \bullet \mathbf{dx} = \int_0^1 \mathbf{B}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \partial\boldsymbol{\varphi}(t)\, dt = \int_0^1 \mathbf{B}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \mathbf{dx} \qquad (10.12.17)$$

where $\mathbf{dx} = (\partial_t\boldsymbol{\varphi}(t))\, dt$. This integral is normally written $\int_\varphi \mathbf{B}(\mathbf{x}) \bullet \mathbf{dx}$ showing again the motivation for the cosmetic functional notation ***dx***. This is the "line integral of a vector field **B** over a curve φ ".

Now return to (10.12.11),

$$\int_\varphi \alpha_{\mathbf{x}} = \int_0^1 \mathbf{f}(\boldsymbol{\varphi}(t)) \bullet (\partial_t\boldsymbol{\varphi}(t))\, dt\,. \qquad \textit{(10.12.11)}$$

Suppose the vector field $\mathbf{f}(\boldsymbol{\varphi}(t))$ happens to be tangent to the curve $\boldsymbol{\varphi}$ for all values of t. In this case

$$\mathbf{f}(\boldsymbol{\varphi}(t)) \bullet (\partial_t\boldsymbol{\varphi}(t)) = |\,\mathbf{f}(\boldsymbol{\varphi}(t))\,|\,|\,(\partial_t\boldsymbol{\varphi}(t))\,| \qquad (10.12.18)$$

since $\partial_t\boldsymbol{\varphi}(t)$ is tangent to the curve at t. Note that

$$|\,(\partial_t\boldsymbol{\varphi}(t))\,|^2 = \Sigma_{i=1}^m\ (\partial_t\varphi^i(t))^2 = \Sigma_{i=1}^m\ (R^i{}_1)^2 \qquad . \qquad (10.12.19)$$

which we recognize as the K^2 object of (10.10.41). Setting $|\,\mathbf{f}(\boldsymbol{\varphi}(t))\,| = T(\boldsymbol{\varphi}(t))$, we find that

$$\int_\varphi \alpha_{\mathbf{x}} = \int_0^1 T(\boldsymbol{\varphi}(t))\, K(t)\, dt \qquad (10.12.20)$$

and this shows how the temperature integral of (10.10.42) can be fitted into the 1-form framework.

<u>Example for $R^m = R^2$:</u> The "angle form" problem mentioned in (10.5.10). (10.12.21)

In this problem we have specific functions f_1 and f_2, a specific range [0,2π] for the t-space domain, and a specific curve (a circle) $\mathbf{x} = \boldsymbol{\varphi}(t) = (x^1,x^2)$.

$$\alpha_{\mathbf{x}} = \Sigma_{i=1}^2 f_i(\mathbf{x})\, {}_{\mathbf{x}}\lambda^i = f_1(\mathbf{x})\, dx^1 + f_2(\mathbf{x})\, dx^2$$

$$= -\,(x^2/r^2)\, dx^1 + (x^1/r^2)\, dx^2 \qquad \text{where } r^2 \equiv (x^1)^2 + (x^2)^2$$

$x^1 = \varphi^1(t) = \cos t$ $\quad\partial_t\varphi^1(t) = -\sin t$ $\quad t = [0,2\pi]$
$x^2 = \varphi^2(t) = \sin t$ $\quad\partial_t\varphi^2(t) = \cos t$ $\quad$ t is the polar angle of the vector $\mathbf{x} = (x^1,x^2)$

$r^2 = (x^1)^2 + (x^2)^2 = \cos^2 t + \sin^2 t = 1$ $\quad$ vector **x** lies on the unit circle in x-space

$f_1(\boldsymbol{\varphi}(t)) = -(x^2/r^2) = -\sin t$
$f_2(\boldsymbol{\varphi}(t)) = +(x^1/r^2) = \cos t$

$\alpha_{\mathbf{x}} = -\sin t\ dx^1 + \cos t\ dx^2$

$\varphi^*(\alpha_{\mathbf{x}}) = g(t)\ dt$ $\quad$ pullback of $\alpha_{\mathbf{x}}$ (10.12.7)

$g(t) = f_i(\boldsymbol{\varphi}(t))\,[\partial\varphi^i(t)/\partial t] = [\ f_1(\boldsymbol{\varphi}(t))\partial_t\varphi^1(t) + f_2(\boldsymbol{\varphi}(t))\partial_t\varphi^2(t)]\ = [\ (-\sin t)(-\sin t) + (\cos t)(\cos t)\ ]$

$= 1$

so

$\varphi^*(\alpha_{\mathbf{x}}) = 1\ dt$

$$\int_\varphi \alpha_{\mathbf{x}} = \int_{[0,2\pi]} \varphi^*(\alpha) = \int_{[0,2\pi]} g(t)\ dt = \int_{[0,2\pi]} dt = \int_0^{2\pi} dt$$

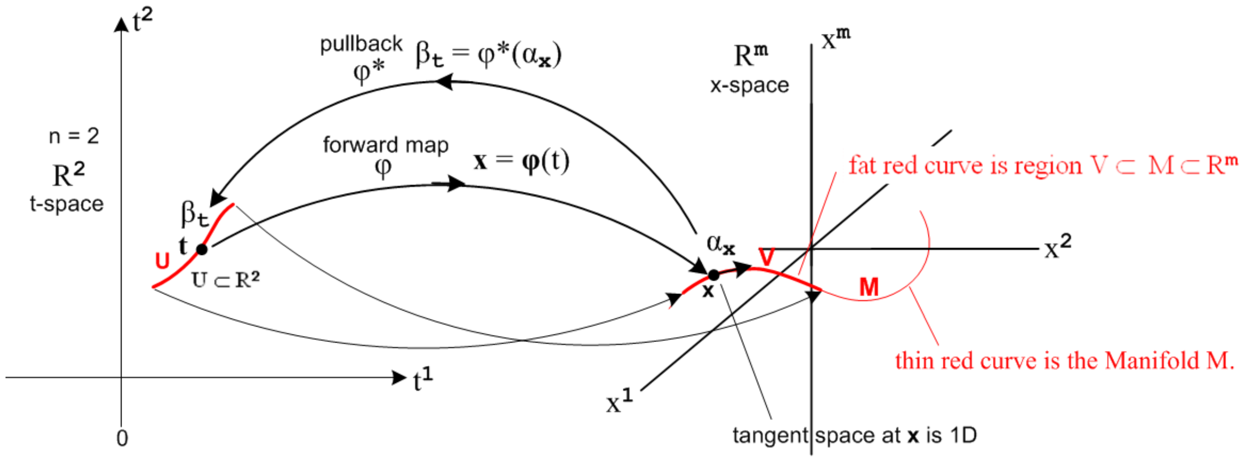
$= 2\pi$.

So the integral of this particular 1-form α around the unit circle gives the number 2π. In this example we are trying to "cover" a full circle with a single mapping $\mathbf{x} = \boldsymbol{\varphi}(t)$ and the circle has a "seam" which maps back to both $t = 0$ and $t = 2\pi$ resulting in the 2π above. See comments below (10.5.9) concerning how this 1-form example provides a counterexample to the Poincaré Lemma and shows that $\alpha_{\mathbf{x}}$ is not exact.

Integration of 1-forms over more general regions of t-space

In the general mapping picture where $\varphi : R^n \to R^m$ one is allowed to have k-forms with $k \le n$ but we are usually interested in the case that $k = n$ since this makes the most "efficient" use of t-space on the left. But there is no reason not to consider $k < n$.

Consider then this 1-form situation in the context $\varphi: R^2 \to R^3$:



(10.12.22)

Now the simple 1-cube in R^1 t-space is replaced by a general curve U in R^2, but we are still mapping a curve U to a curve V. We go through the steps above:

$$\alpha_{\mathbf{x}} = \Sigma_i f_i(\mathbf{x})\, {}_{\mathbf{x}}\lambda^i = \Sigma_i f_i(\mathbf{x})\, dx^i \quad . \qquad \textit{(10.12.2)}$$

$$\int_\varphi \alpha_{\mathbf{x}} = \int_V \Sigma_i f_i(\mathbf{x})\, dx^i = \text{integral of a 1-form over a piece of the curve } \mathbf{V} \text{ in } R^m \; . \qquad \textit{(10.12.3)}$$

<u>First definition:</u>

$$\int_V \alpha_{\mathbf{x}} \equiv \int_U \varphi^*(\alpha_{\mathbf{x}})$$

$$= \text{the integral in t-space of the pullback of } \alpha_{\mathbf{x}} \text{ over the curve U in } R^2 \qquad (10.12.23)$$

We then compute the pullback of $\varphi^*(\alpha_{\mathbf{x}})$ of $\alpha_{\mathbf{x}}$:

$$\alpha_{\mathbf{x}} = \Sigma_i f_i(\mathbf{x})\, {}_{\mathbf{x}}\lambda^i = \Sigma_i f_i(\mathbf{x})\, dx^i = \mathbf{f}(\mathbf{x}) \bullet \mathbf{dx} \,, \qquad \mathbf{dx} \equiv (dx^1, dx^2, \ldots, dx^m)$$

$$\varphi^*(\alpha_{\mathbf{x}}) = \Sigma_{i=1}{}^m \varphi^*(f_i(\mathbf{x}))\, \varphi^*(dx^i) \qquad \text{// (10.9.5) 3}$$

$$= \Sigma_{i=1}{}^m f_i(\boldsymbol{\varphi}(t))\; \Sigma_{j=1}{}^2 R^i{}_j\, dt^j \qquad \text{// (10.9.5) 1 and 5}$$

$$= \Sigma_i \Sigma_j\; f_i(\boldsymbol{\varphi}(t))\, [\partial\varphi^i(t)/\partial t^j]\; dt^j$$

$$= \Sigma_j\; g_j(\mathbf{t})\; dt^j$$

$$= \mathbf{g}(\mathbf{t}) \bullet \mathbf{dt} \qquad (10.12.24)$$

where

$$g_j(t) \equiv \Sigma_i f_i(\boldsymbol{\varphi}(t))\, [\partial\varphi^i(t)/\partial t^j] = \Sigma_i f_i(\boldsymbol{\varphi}(t))\, \partial_j\varphi^i(t) = \mathbf{f}(\boldsymbol{\varphi}(t)) \bullet (\partial_j\boldsymbol{\varphi}) \,. \qquad (10.12.25)$$

Second definition:

$$\int_U g_j(t)\ {}_t\lambda^j = \int_U \mathbf{g}(\mathbf{t}) \bullet \mathbf{dt} = \int_U \mathbf{g}(\mathbf{t}) \bullet \mathrm{d}\mathbf{t}\ . \tag{10.12.26}$$

Assembling the pieces,

$$\int_V \alpha_{\mathbf{x}} = \int_V \mathbf{f}(\mathbf{x}) \bullet \mathbf{dx} = \int_U \varphi^*(\alpha_{\mathbf{x}}) = \int_U \mathbf{g}(\mathbf{t}) \bullet \mathbf{dt} = \int_U \mathbf{g}(\mathbf{t}) \bullet \mathrm{d}\mathbf{t}\ . \tag{10.12.27}$$

When one curve is mapped into another by $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$, this result shows how to reduce the integral of the 1-form $\alpha_{\mathbf{x}}$ to a calculus line integral in t-space. In effect, the line integral $\int_V \mathbf{f}(\mathbf{x}) \bullet \mathrm{d}\mathbf{x}$ in x-space is replaced by the line integral $\int_U \mathbf{g}(\mathbf{t}) \bullet \mathrm{d}\mathbf{t}$ in t-space.

10.13 Integration of 2-forms

We wish now to look in more detail at the integration of 2-forms. The general k-form integration result is stated in (10.12.1). Once again, there is much repetition below intended to reinforce the meaning of various objects.

For the moment we set m = 3 and consider this 2-form in x-space R^3 ,

$$\alpha_{\mathbf{x}} = \Sigma'_I\ f_I(x)\ {}_x\lambda^{\wedge I} = \Sigma'_I\ f_I(x)\ dx^{\wedge I} = \Sigma_{1\le i_1<i_2\le 3}\ f_{i_1 i_2}(\mathbf{x})\ dx^{i_1} \wedge dx^{i_2}$$

$$= f_{12}(\mathbf{x})\ dx^1 \wedge dx^2 + f_{13}(\mathbf{x})\ dx^1 \wedge dx^3 + f_{23}(\mathbf{x})\ dx^2 \wedge dx^3\ . \tag{10.13.1}$$

We wish to define a meaning for the integration of this 2-form α over a piece of the surface $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$,

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \Sigma'_I\ f_I(x)\ dx^{i_1} \wedge dx^{i_2} = \text{integral of a 2-form over the surface } \varphi \text{ in } R^3\ . \tag{10.13.2}$$

The transformation $\mathbf{x} = \boldsymbol{\varphi}(t)$ is a mapping $\boldsymbol{\varphi}$: $(t_1,t_2) \rightarrow R^3$ where t_1 and t_2 are "parameters".

Comment: Officially it is the mapping $\boldsymbol{\varphi}$ which is the "surface", but we loosely refer to the image (trace) of this mapping in R^3 as "the surface". The distinction is necessary because many mappings can have the same image surface, such as $\boldsymbol{\varphi}(t_1,t_2)$ and $\boldsymbol{\varphi}(t_1^2,t_2^2)$ where the parameters are "respeeded" (reparametrized). If the integral of the 2-form $\alpha_{\mathbf{x}}$ is the same over surfaces $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ which have the same image surface, the two surfaces are called smoothly equivalent surfaces, see Buck p 386 Theorem 2 (ii).

So imagine that we have a 2D surface hanging in R^3 and as t_1 and t_2 vary (each perhaps from 0 to 1 in t-space), we move around on the image surface in R^3. The problem is how to integrate a 2-form over this surface.

We can define the calculational meaning of the above integral in two steps, each being a definition, as outlined in Section 10.11.

First definition:

$$\int_{\varphi} \alpha_{\mathbf{x}} \equiv \int_{[0,1]^2} \varphi^*(\alpha_{\mathbf{x}})$$

$$= \text{the integral in t-space of the pullback of } \alpha_{\mathbf{x}} \text{ over the 2-cube } [0,1]^2 \qquad (10.13.3)$$

On the left is an integral of the 2-form $\alpha_{\mathbf{x}}$ over a surface $\boldsymbol{\varphi}$ in R^3.
On the right is an integral of a *different* 2-form $\varphi^*(\alpha_{\mathbf{x}})$ (the pullback of $\alpha_{\mathbf{x}}$) over a 2-cube $[0,1]^2$ in R^2.
Note that $\alpha_{\mathbf{x}}$ lies in ${}_{\mathbf{x}}\Lambda^2(R^3)$ while $\varphi^*(\alpha_{\mathbf{x}})$ lies in ${}_{\mathbf{t}}\Lambda^2(R^2)$.
Since our usual mapping is $\varphi^* : {}_{\mathbf{x}}\Lambda^1(R^m) \rightarrow {}_{\mathbf{t}}\Lambda^1(R^n)$ we have m = 3 and n = 2 (see (10.7.18)).

The "tall" m x n R-matrix for this problem is then a 3 x 2 matrix,

$$R^i{}_j = (D\varphi)^i{}_j = \partial\varphi^i/\partial t^j = \partial_j\varphi^i \qquad i = 1,2,3 \qquad j = 1,2 \qquad R = \begin{pmatrix} \partial_1\varphi^1 & \partial_2\varphi^1 \\ \partial_1\varphi^2 & \partial_2\varphi^2 \\ \partial_1\varphi^3 & \partial_2\varphi^3 \end{pmatrix}. \qquad (10.13.4)$$

We then compute the pullback $\varphi^*(\alpha_{\mathbf{x}})$ of $\alpha_{\mathbf{x}}$:

$$\alpha_{\mathbf{x}} = \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\mathbf{x})\, {}_{\mathbf{x}}\lambda^{\wedge I} = \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\mathbf{x})\, dx^{i_1} \wedge dx^{i_2}$$

$$\varphi^*(\alpha_{\mathbf{x}}) = \Sigma_{1\le i_1<i_2\le 3}\, \varphi^*(f_{i_1 i_2}(\mathbf{x}))\, \varphi^*(dx^{i_1} \wedge dx^{i_2}) \qquad // (10.9.5)\ 3$$

$$= \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\boldsymbol{\varphi}(\mathbf{t}))\, \Sigma_{1\le j_1<j_2\le 2} \det \begin{pmatrix} \frac{\partial\varphi^{i_1}}{\partial t^{j_1}} & \frac{\partial\varphi^{i_1}}{\partial t^{j_2}} \\ \frac{\partial\varphi^{i_2}}{\partial t^{j_1}} & \frac{\partial\varphi^{i_2}}{\partial t^{j_2}} \end{pmatrix} dt^{j_1} \wedge dt^{j_2} \qquad // (10.9.16)$$

$$= \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\boldsymbol{\varphi}(\mathbf{t})) \det \begin{pmatrix} \frac{\partial\varphi^{i_1}}{\partial t^1} & \frac{\partial\varphi^{i_1}}{\partial t^2} \\ \frac{\partial\varphi^{i_2}}{\partial t^1} & \frac{\partial\varphi^{i_2}}{\partial t^2} \end{pmatrix} dt^1 \wedge dt^2 \qquad // \text{only one term in } \Sigma_{1\le j_1<j_2\le 2}$$

$$= \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\boldsymbol{\varphi}(\mathbf{t})) \det \begin{pmatrix} \partial_1\varphi^{i_1} & \partial_2\varphi^{i_1} \\ \partial_1\varphi^{i_2} & \partial_2\varphi^{i_2} \end{pmatrix} dt^1 \wedge dt^2 \qquad // \text{more compact notation}$$

// these determinants are the 2x2 minors of the matrix R shown above

$$= \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\boldsymbol{\varphi}(\mathbf{t}))\, \frac{\partial(\varphi^{i_1}, \varphi^{i_2})}{\partial(t^1, t^2)}\, dt^1 \wedge dt^2 \qquad // \text{Jacobian notation for determinants}$$

$$= \Sigma_{1\le i_1<i_2\le 3}\, g_{i_1 i_2}(\mathbf{t})\, dt^1 \wedge dt^2$$

where

$$g_{i_1i_2}(\mathbf{t}) = f_{i_1i_2}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^{i_1}, \varphi^{i_2})}{\partial(t^1, t^2)} . \tag{10.13.5}$$

Since in this example there are so few terms (three) in the sum $\Sigma_{1\le i_1<i_2\le 3}$, we just write them out

$$\varphi^*(\alpha_{\mathbf{x}}) = [\, f_{12}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} + f_{13}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^1, \varphi^3)}{\partial(t^1, t^2)} + f_{23}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)}\,]\ dt^1 \wedge dt^2$$

$$= G(\mathbf{t})\ dt^1 \wedge dt^2$$

where

$$G(\mathbf{t}) \equiv [\, f_{12}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} + f_{13}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^1, \varphi^3)}{\partial(t^1, t^2)} + f_{23}(\boldsymbol{\varphi}(\mathbf{t}))\frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)}\,] . \tag{10.13.6}$$

The pullback $\varphi^*(\alpha_{\mathbf{x}}) = G(\mathbf{t})\ dt^1 \wedge dt^2$ is a 2-form in dual t-space ${}_t\Lambda^2(R^2)$.

Using the definition given above, one then has,

$$\int_{\varphi} \alpha_{\mathbf{x}} \equiv \int_{[0,1]^2} \varphi^*(\alpha) = \int_{[0,1]^2} G(\mathbf{t})\ dt^1 \wedge dt^2 . \tag{10.13.7}$$

Thus the integral of the 2-form $\alpha_{\mathbf{x}}$ over the surface $\boldsymbol{\varphi}$ in x-space is defined to be equal to the integral of the 2-form $G(\mathbf{t})\ dt^1 \wedge dt^2$ over a 2-cube in t-space. So far no regular calculus integrals have appeared.

<u>Second definition:</u>

$$\int_{[0,1]^2} G(\mathbf{t})\ {}_t\lambda^1 \wedge {}_t\lambda^2 = \int_{[0,1]^2} G(\mathbf{t})\ dt^1 \wedge dt^2 \equiv \int_0^1 \int_0^1 G(t^1,t^2)\ dt^1 dt^2 . \tag{10.13.8}$$

On the left is the integral of a 2-form on a 2-cube, on the right is an ordinary calculus integral of a function over the 2-cube. It is this second definition that motivates giving the dual space basis vectors ${}_t\lambda^1$ and ${}_t\lambda^2$ the cosmetic names dt^1 and dt^2 .

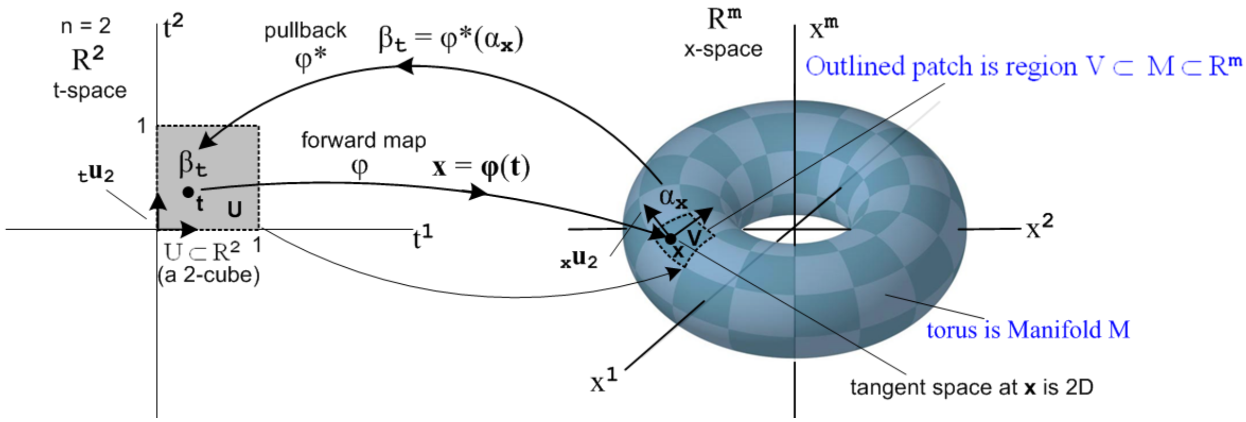
Notice that the only locations where the functions $f_{i_1i_2}(\mathbf{x})$ are "sensed" in this integral are points on the surface $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$.

Combining the two definitions gives

$$\int_{\varphi} \alpha_{\mathbf{x}} = \int_0^1 \int_0^1 G(t^1,t^2)\ dt^1 dt^2 \tag{10.13.9}$$

with G as in (10.13.6).

We redisplay the earlier Fig (10.9.3a) to illustrate the above discussion, where $\beta_t = G(\mathbf{t})\, dt^1 \wedge dt^2$



(10.13.10)

Here $R^m = R^3$ and the torus is just a sample surface to illustrate the general surface $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$.

<u>Comment on orientation</u>

Orientation of the domain surface is a tricky business in a k-form integral but can at worst cause confusion about the sign of the result. To make things more "visible", suppose [a,b] = [c,d] = [1,0]. In t-space R^2, if $[a,b] = (b-a)\hat{\mathbf{t}}_1 = \hat{\mathbf{t}}_1$ and $[c,d] = (d-c)\hat{\mathbf{t}}_2 = \hat{\mathbf{t}}_2$, then the 2-cube (unit square) integration domain $[0,1]^2$ can be regarded as being over region $\hat{\mathbf{t}}_1 \times \hat{\mathbf{t}}_2$, a vector area with two "sides" (orientations)

$$\int_{\hat{t}_1 \times \hat{t}_2} dt^1 dt^2 = 1 \qquad \text{area of the front side of a unit square is 1 area unit}$$

$$\int_{\hat{t}_2 \times \hat{t}_1} dt^1 dt^2 = -1 \qquad \text{area of the back side of a unit square is -1 area unit .} \qquad (10.13.11)$$

This is analogous to the 1D situation where [0,1] and [0,1] have opposite orientations.

$$\int_0^1 dt\ 1 = 1$$

$$\int_1^0 dt\ 1 = -1 \ .$$

The integral appearing in (10.13.8) is this

$$\int_{[0,1]^2} G(\mathbf{t})\, dt^1 \wedge dt^2 = \int_{\hat{t}_1 \times \hat{t}_2} G(\mathbf{t})\, dt^1 \wedge dt^2 = \int_0^1 \int_0^1 G(t^1,t^2)\, dt^1 dt^2 \ . \qquad (10.13.12)$$

Now consider

$$\int_{\hat{t}_1 \times \hat{t}_2} G(\mathbf{t})\, dt^1 \wedge dt^2 = \int_{\hat{t}_1 \times \hat{t}_2} [- G(\mathbf{t})\, dt^2 \wedge dt^1] = -\int_{\hat{t}_1 \times \hat{t}_2} G(\mathbf{t})\, dt^2 \wedge dt^1$$

(10.13.13)

$$\int_{\hat{t}_2 \times \hat{t}_1} G(\mathbf{t})\, dt^1 \wedge dt^2 = [- \int_{\hat{t}_1 \times \hat{t}_2}]\ G(\mathbf{t})\, dt^2 \wedge dt^1 = -\int_{\hat{t}_1 \times \hat{t}_2} G(\mathbf{t})\, dt^2 \wedge dt^1 \ .$$

In both these equations a minus sign is generated. In the first the minus sign arises because the 2-form called G(**t**) $dt^2 \wedge dt^1$ is the negative of the different 2-form called G(**t**) $dt^2 \wedge dt^1$. In the second equation the integration domain $\hat{\mathbf{t}}_1 \wedge \hat{\mathbf{t}}_2$ refers to the front side of the unit square, while $\hat{\mathbf{t}}_2 \wedge \hat{\mathbf{t}}_1$ refers to the back side of the unit square, and these domains differ by a minus sign. If both changes are made at once one gets

$$\int_{\hat{t}1 \times \hat{t}2} G(\mathbf{t})\, dt^1 \wedge dt^2 = + \int_{\hat{t}2 \times \hat{t}1} G(\mathbf{t})\, dt^2 \wedge dt^1 \qquad (10.13.14)$$

If the integration domain is written simply as $[0,1]^2$, one ends up with

$$\int_{[0,1]^2} G(\mathbf{t})\, dt^1 \wedge dt^2 = + \int_{[0,1]^2} G(\mathbf{t})\, dt^2 \wedge dt^1 \qquad (10.13.15)$$

and this seems to be a contradiction since everyone knows that $dt^2 \wedge dt^1 = - dt^1 \wedge dt^2$. The issue here is that the two $[0,1]^2$ domains are not the same, they just look the same. See Sjamaar's page 64 Remark 5.3 where he treats the domain $\hat{\mathbf{t}}_2 \times \hat{\mathbf{t}}_1$ as a reparametrization of the domain $\hat{\mathbf{t}}_1 \times \hat{\mathbf{t}}_2$ which reverses the orientation of that domain.

<u>Further processing:</u>

Using the Hodge correspondence suggested in (4.3.17) we *define* three new function names F_i

$$f_{12} \equiv F_3$$
$$f_{23} \equiv F_1$$
$$f_{13} \equiv - F_2 \qquad (10.13.16)$$

with a minus in the last line since f_{13} is in anticyclic order compared to (4.3.17). Then (10.13.9) with (10.13.6) for G(**t**) becomes,

$$\int_\varphi \alpha_x = \int_0^1 \int_0^1 [\, F_3(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} - F_2(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^1, \varphi^3)}{\partial(t^1, t^2)} + F_1(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)} \,]\; dt^1dt^2$$

$$= \int_0^1 \int_0^1 [\, F_1(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)} - F_2(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^1, \varphi^3)}{\partial(t^1, t^2)} + F_3(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} \,]\; dt^1dt^2$$

$$= \int_0^1 \int_0^1 [\, F_1(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)} + F_2(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^3, \varphi^1)}{\partial(t^1, t^2)} + F_3(\boldsymbol{\varphi}(\mathbf{t})) \frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} \,]\; dt^1dt^2$$

$$= \int_0^1 \int_0^1 \mathbf{F}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \mathbf{n}(\mathbf{t})\; dt^1dt^2 \qquad (10.13.17)$$

where **n** below is a vector normal to the surface in R^3 at point $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$,

$$\mathbf{n(t)} \equiv \left(\frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)}, \frac{\partial(\varphi^3, \varphi^1)}{\partial(t^1, t^2)}, \frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} \right) \qquad \text{// Buck p 335, 403}$$

$$= \left(\det\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}, + \det\begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix}, \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} \right) . \qquad (10.13.18)$$

That **n** really is a normal vector can be verified by showing that $\mathbf{n} \bullet {}_{\mathbf{x}}\mathbf{u}_j = 0$ for any tangent base vector ${}_{\mathbf{x}}\mathbf{u}_j$ in the tangent space $T_{\mathbf{x}}M$, where from (E.2) $({}_{\mathbf{x}}\mathbf{u}_j)^i = R^i{}_j = \partial_j\varphi^i$, see Buck p 336. But we *know* that $\mathbf{n}(\mathbf{t})$ is a normal vector because the expression for **n** in (10.13.18) is the same as **n'** in (10.10.19) (apart from change of notation) and that **n'** was *constructed* as a cross product of two vectors on the surface so it was a normal.

The vector **n(t)** is not in general a unit vector, so we define

$$\hat{\mathbf{n}}(\mathbf{t}) = \mathbf{n}(\mathbf{t}) / |\, \mathbf{n}(\mathbf{t}) \,| \qquad (10.13.19)$$

where

$$|\, \mathbf{n}(\mathbf{t}) \,|^2 = \left[\frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)} \right]^2 + \left[\frac{\partial(\varphi^3, \varphi^1)}{\partial(t^1, t^2)} \right]^2 + \left[\frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} \right]^2$$

$$= \det{}^2\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} + \det{}^2\begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix} + \det{}^2\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix}$$

$$\equiv K(\mathbf{t})^2 \qquad (10.13.20)$$

which we recognize as the same object K^2 appearing in (10.10.18). Then with

$$\mathbf{n}(\mathbf{t}) = K(\mathbf{t})\, \hat{\mathbf{n}}(\mathbf{t}) \qquad (10.13.21)$$

equation (10.13.17) may now be written,

$$\int_\varphi \alpha_{\mathbf{x}} = \int_0^1 \int_0^1 \mathbf{F}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \hat{\mathbf{n}}(\mathbf{t})\, K(\mathbf{t})\, dt^1 dt^2 . \qquad (10.13.22)$$

Going back to the original 2-form $\alpha_{\mathbf{x}}$ (10.13.1) one can write,

$$\alpha_{\mathbf{x}} = \Sigma_{1\le i_1<i_2\le 3}\, f_{i_1 i_2}(\mathbf{x})\, dx^{i_1} \wedge dx^{i_2}$$

$$= f_{12}(\mathbf{x})\, dx^1 \wedge dx^2 + f_{13}(\mathbf{x})\, dx^1 \wedge dx^3 + f_{23}(\mathbf{x})\, dx^2 \wedge dx^3$$

$$= F_1(\mathbf{x})\, dx^2 \wedge dx^3 + F_2(\mathbf{x})\, dx^3 \wedge dx^1 + F_3(\mathbf{x})\, dx^1 \wedge dx^2$$

$$= F_1(\mathbf{x})\, dA^1 + F_2(\mathbf{x})\, dA^2 + F_3(\mathbf{x})\, dA^3$$

$$= \mathbf{F}(\mathbf{x}) \bullet d\mathbf{A} \qquad \text{// } dA^1 = {*dx^1} \text{ etc, so can say } d\mathbf{A} = {*d\mathbf{x}} \qquad (10.13.23)$$

where in cyclic order we define the following differential area 2-forms

$$dA^1 \equiv dx^2 \wedge dx^3 \qquad dA^2 \equiv dx^3 \wedge dx^1 \qquad dA^3 \equiv dx^1 \wedge dx^2 \ . \tag{10.13.24}$$

Then our result (10.13.22) can be concisely written,

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \mathbf{F}(\mathbf{x}) \bullet \boldsymbol{dA} = \int_0^1 \int_0^1 [\mathbf{F}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \hat{\mathbf{n}}(\mathbf{t})\,]\, K(\mathbf{t})\, dt^1 dt^2.$$

$$= \int_0^1 \int_0^1 [\mathbf{F}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \hat{\mathbf{n}}(\mathbf{t})]\ dA \qquad dA = K(\mathbf{t})\, dt^1 dt^2$$

$$= \int_0^1 \int_0^1 \mathbf{F}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \mathbf{dA} \qquad \mathbf{dA} = dA\, \hat{\mathbf{n}} = K(\mathbf{t})\, dt^1 dt^2\ \hat{\mathbf{n}}(\mathbf{t}) \ . \tag{10.13.25}$$

Normally this is written $\int_\varphi \mathbf{F}(\mathbf{x}) \bullet \mathbf{dA}$ showing the motivation for the cosmetic functional notation $\boldsymbol{dA}$ as defined above. This is the "integral of a vector field **F** over a surface **φ** ".

Recall now our "no differential forms" surface integration done in (10.10.20)

$$A' \langle B_n \rangle = \int_{S'} dA'\ \mathbf{B}(\mathbf{x'}) \bullet \hat{\mathbf{n}}' = \int_S \mathbf{B}(\mathbf{F}(\mathbf{x})) \bullet \hat{\mathbf{n}}'\, K(\mathbf{x})\, dx^1 dx^2 \ . \tag*{(10.10.17)}$$

In the $\mathbf{x} = \boldsymbol{\varphi}(t)$ notation this reads

$$A \langle B_n \rangle = \int_S dA\ \mathbf{B}(\mathbf{x}) \bullet \hat{\mathbf{n}} = \int_S \mathbf{B}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \hat{\mathbf{n}}\, K(\mathbf{t})\, dt^1 dt^2 \tag{10.13.26}$$

which is the same integral appearing in (10.13.25) with **F** = **B**. Therefore, we can interpret (10.13.26) as being the integral of the 2-form,

$$\alpha_{\mathbf{x}} = \mathbf{B}(\mathbf{x}) \bullet \boldsymbol{dA} \tag{10.13.27}$$

and one then has

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \mathbf{B}(\mathbf{x}) \bullet \boldsymbol{dA} = \int_0^1 \int_0^1 [\mathbf{B}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \hat{\mathbf{n}}(\mathbf{t})\,]\, K(\mathbf{t})\, dt^1 dt^2$$

$$= \int_0^1 \int_0^1 \mathbf{B}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \mathbf{dA} \ . \tag{10.13.28}$$

Now return to (10.13.25),

$$\int_\varphi \alpha_{\mathbf{x}} = \int_\varphi \mathbf{F}(\mathbf{x}) \bullet \boldsymbol{dA} = \int_0^1 \int_0^1 \mathbf{F}(\boldsymbol{\varphi}(\mathbf{t})) \bullet \hat{\mathbf{n}}(\mathbf{t})\ K(\mathbf{t})\, dt^1 dt^2 \ . \tag*{(10.13.25)}$$

Suppose the vector field $\mathbf{F}(\boldsymbol{\varphi}(\mathbf{t}))$ happens to be normal to the surface φ for all values of $\mathbf{t}$. In this case,

$$[\mathbf{F}(\boldsymbol{\varphi}(\mathbf{t}))\bullet\ \hat{\mathbf{n}}(\mathbf{t})\ ]\ =\ |\ \mathbf{F}(\boldsymbol{\varphi}(\mathbf{t}))\ |\ . \tag{10.13.29}$$

Then setting $|\ \mathbf{F}(\boldsymbol{\varphi}(\mathbf{t}))\ | = T(\boldsymbol{\varphi}(\mathbf{t}))$ we find that

$$\int_\varphi \alpha_{\mathbf{x}}\ =\ \int_0^1 \int_0^1\ T(\boldsymbol{\varphi}(\mathbf{t}))\ K(\mathbf{t})\ dt^1 dt^2 \tag{10.13.30}$$

and this shows how the temperature integral of (10.10.20) can be fitted into the 2-form framework.

<u>Generalization from φ: $R^2\rightarrow R^3$ to φ: $R^2\rightarrow R^m$</u>

If the 2D surface lies in R^m instead of R^3, the above results are easily generalized. The 2-form $\alpha_{\mathbf{x}}$ is then,

$$\alpha_{\mathbf{x}}\ = \Sigma_{1\le i_1<i_2\le m}\ f_{i_1i_2}(\mathbf{x})\ {}_{\mathbf{x}}\lambda^{\wedge I} = \Sigma_{1\le i_1<i_2\le m}\ f_{i_1i_2}(\mathbf{x})\ dx^{i_1} \wedge dx^{i_2}\ =\ \Sigma'_I\ f_I(\mathbf{x})\ dx^{\wedge I} \tag{10.13.31}$$

where there are (m,2) = m(m-1)/2 terms in the ordered sum Σ'_I. The pullback also has (m,2) terms, being

$$\varphi^*(\alpha_{\mathbf{x}})\ = \Sigma_{1\le i_1<i_2\le m}\ f_{i_1i_2}(\boldsymbol{\varphi}(\mathbf{t}))\ \frac{\partial(\varphi^{i_1},\varphi^{i_2})}{\partial(t^1,t^2)}\ dt^1 \wedge dt^2\ \ = \Sigma'_I\ f_I(\boldsymbol{\varphi}(\mathbf{t}))\ \det\begin{pmatrix} R^{i_1}{}_1 & R^{i_1}{}_2 \\ R^{i_2}{}_1 & R^{i_2}{}_2\end{pmatrix} dt^1 \wedge dt^2$$

$$= \Sigma_{1\le i_1<i_2\le m}\ g_{i_1i_2}(\mathbf{t})\ dt^1 \wedge dt^2 \qquad\qquad g_{i_1i_2}(\mathbf{t}) = f_{i_1i_2}(\boldsymbol{\varphi}(\mathbf{t}))\ \frac{\partial(\varphi^{i_1},\varphi^{i_2})}{\partial(t^1,t^2)}$$

$$= G(\mathbf{t})\ \ dt^1 \wedge dt^2 \qquad\qquad G(\mathbf{t}) = \Sigma_{1\le i_1<i_2\le m}\ f_{i_1i_2}(\boldsymbol{\varphi}(\mathbf{t}))\ \frac{\partial(\varphi^{i_1},\varphi^{i_2})}{\partial(t^1,t^2)}\ . \tag{10.13.32}$$

The pullback $\varphi^*(\alpha_{\mathbf{x}}) = G(t)\ dt^1 \wedge dt^2$ is still a 2-form in dual t-space ${}_t\Lambda^2(R^2)$. Then

$$\int_\varphi \alpha_{\mathbf{x}}\ \equiv \int_{[0,1]^2} \varphi^*(\alpha_{\mathbf{x}})\ = \int_{[0,1]^2}\ G(\mathbf{t})\ dt^1 \wedge dt^2\ \equiv\ \ \int_0^1 \int_0^1\ G(\mathbf{t})\ dt^1 dt^2 \tag{10.13.33}$$

Here $G(\mathbf{t})$ is a sum of (m,2) terms each of which is a 2x2 Jacobian weighted by a function $f_{i_1i_2}$.

<u>Example: φ: $R^2\rightarrow R^4$</u>

$$\alpha_{\mathbf{x}}\ = f_{12}(\mathbf{x})\ dx^1 \wedge dx^2\ +\ f_{13}(\mathbf{x})\ dx^1 \wedge dx^3\ +\ f_{14}(\mathbf{x})\ dx^1 \wedge dx^4 \qquad // \text{ 2-form in } {}_{\mathbf{x}}\Lambda^2(R^4)$$
$$+\ f_{23}(\mathbf{x})\ dx^2 \wedge dx^3\ +\ f_{24}(\mathbf{x})\ dx^2 \wedge dx^4\ +\ f_{34}(\mathbf{x})\ dx^3 \wedge dx^4$$

$$\varphi^*(\alpha_{\mathbf{x}})\ = G(\mathbf{t})\ \ dt^1 \wedge dt^2 \qquad\qquad // \text{ 2-form in } {}_{t}\Lambda^2(R^2)$$

where

$$G(\mathbf{t}) = [\ f_{12}(\mathbf{x})\frac{\partial(\varphi^1, \varphi^2)}{\partial(t^1, t^2)} + f_{13}(\mathbf{x})\frac{\partial(\varphi^1, \varphi^3)}{\partial(t^1, t^2)} + f_{14}(\mathbf{x})\frac{\partial(\varphi^1, \varphi^4)}{\partial(t^1, t^2)}$$

$$+ f_{23}(\mathbf{x})\frac{\partial(\varphi^2, \varphi^3)}{\partial(t^1, t^2)} + f_{24}(\mathbf{x})\frac{\partial(\varphi^2, \varphi^4)}{\partial(t^1, t^2)} + f_{34}(\mathbf{x})\frac{\partial(\varphi^3, \varphi^4)}{\partial(t^1, t^2)}\ ] \qquad (10.13.34)$$

and the integrated 2-form is

$$\int_\varphi \alpha_\mathbf{x} \equiv \int_{[0,1]^2} \varphi^*(\alpha) = \int_0^1 \int_0^1 dt^1 dt^2\, G(\mathbf{t})\ . \qquad (10.13.35)$$

<u>Generalization to φ: $R^n \rightarrow R^m$ with $k = n$</u>

This generalization is discussed in Appendix G.2 where the $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ context is used. It turns out that there is still a unit vector $\hat{\mathbf{m}}$ having binomial (n,m) components which is similar to the $\hat{\mathbf{n}}$ discussed above. The Appendix G.2 discussion relies on a theorem, proven in Appendix G.1, showing that $\det(R^T R)$ is the sum of the squares of the full-width minors of R. The significance of $\det(R^T R)$ as a volume measure is presented in Appendix F.

Appendix A: Permutation Support

This is a very long and detailed appendix, so a summary is in order:

Section A.1 describes our Σ_P permutation notation and comments on the permutation group. It then proves three different "rearrangement theorems" and states various determinant expansions using Σ_P notation. It is shown that det(M) = det(M^T) = det($M^{\mathcal{T}}$) for any index positions of a rank-2 tensor M.

Section A.2 describes the action of a permutation operator on a generic function f(1,2...k), and states several theorems concerning multiple permutation operators. At the same time, the Alt operator is defined and various facts are proven concerning this operator. The notion of a totally antisymmetric generic function is directly related to the Alt operator.

Section A.3 mimics Section A.2 for the Sym operator in place of the Alt operator. The notion of a totally symmetric generic function is directly related to the Sym operator.

Section A.4 states some facts which concern both the Alt and the Sym operators together.

Up to this point, the various facts and theorems have taken place in a "generic permutation space" which consists of functions of k arguments which are a permutation of 1,2...k, such as f(2,1,3...k).

Section A.5 applies all the previous facts and theorems to the permutation space whose elements are the component indices of a rank-k tensor, so f(1,2,3...k) = $T^{i_1 i_2 ... i_k}$. The results of Sections A.2, A.3 and A.4 are adapted to the tensor world in subsections (a), (b) and (c).

Section A.6 deals with the permutation tensor $\varepsilon_{i_1 i_2 ... i_k}$ and shows how it can provide an alternative to the permutation notation in some situations associated with the Alt operator.

Section A.7 adapts the above generic results to the case f(1,2,....k) = ($v_{j_1} \otimes v_{j_2} \otimes .....\otimes v_{j_k}$) which is the tensor product of k vectors. Then the wedge product of k vectors is defined in terms of this application of the Alt operator, so that ($v_{j_1} \wedge v_{j_2} \wedge .....\wedge v_{j_k}$) ≡ Alt($v_{j_1} \otimes v_{j_2} \otimes .....\otimes v_{j_k}$).

Section A.8 is similar to Section A.5, but the facts and theorems are applied not to tensors, but to "tensor functions", so here f(1,2...k) = $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$. The permutation space is now the set of label subscripts on the k vector arguments of a tensor function. The results of Sections A.2, A.3 and A.4 are adapted to the tensor function world in subsections (a), (b) and (c). Subsection (d) then derives special-purpose theorems that apply to objects with two multiindices, and then to objects which have a factored form.

Section A.9 proves an obscure ordered permutation sum theorem that is used in (7.4.12).

Section A.10 defines the generic function space tensor product and then relates it to the tensor product of tensors and then to the tensor product of tensor functions.

A.1 Rearrangement Theorems and Determinants

Definition: A **permutation** P (of order k) reorders the list of integers [1,2,3...k] in some manner to give $[i_1,i_2,i_3...i_k]$. (A.1.1)

Including the initial ordering [1,2,3...k], there are k! possible permutations.

Fact: $\Sigma_P(1) = k!$. // there are k! equal terms in this sum (A.1.2)

The permutation group **rearrangement theorem** states the following:

$$\Sigma_P f(QP) = \Sigma_P f(PQ) = \Sigma_P f(P) \quad . \tag{A.1.3}$$

Here Σ_P is a sum over all k! permutations of [1,2...k], and Q is any one of these permutations.
The first two sums are just *reorderings* or *rearrangements* of the third sum and so equal the third sum.

Proof: This theorem is true because the permutations P of [1,2...k] form a *group* G :

- $P_1P_2 = P_3 \in G$ // closure
- $(P_1P_2)P_3 = P_1(P_2P_3)$ // associative
- $P = I$ // identity exists, permutation that does nothing to [1,2...k]
- P^{-1} exists for any P // just the inverse permutation. (A.1.4)

It is a fact that, for any group G with k elements g_i,

$$g_a [g_1, g_2, ...g_k] = [g_ag_1, g_ag_2, ...g_ag_k] = [g'_1, g'_2, ...g'_k] = \text{reordering of } [g_1, g_2, ...g_k]$$
$$[g_1, g_2, ...g_k]g_a = [g_1g_a, g_2g_a, ...g_kg_a] = [g''_1, g''_2, ...g''_k] = \text{reordering of } [g_1, g_2, ...g_k] \quad . \tag{A.1.5}$$

To show that $[g'_1, g'_2, ...g'_k]$ is a reordering of $[g_1, g_2, ...g_k]$, we have to show that no two elements of $[g'_1, g'_2, ...g'_k]$ are the same. Suppose for example $g'_1 = g'_2$. That would imply $g_ag_1 = g_ag_2$. Since g_a^{-1} exists in a group for any g_a, apply g_a^{-1} to both sides to get $g_a^{-1}g_ag_1 = g_a^{-1}g_ag_2$ or $g_1 = g_2$. But that contradicts the basic starting point that $[g_1, g_2, ...g_k]$ enumerates the distinct group elements. Therefore

$$\Sigma_i f(g_ag_i) = \Sigma_i f(g_ig_a) = \Sigma_i f(g_i) \quad . \tag{A.1.6}$$

This is valid only if the sum is over *all* elements of the group, which in the rearrangement theorem (A.1.3) means the sum Σ_P must be over all permutations P.

In any group, if g exists, so does g^{-1}, and it is just some element of the group. For the permutation group P^{-1} exists and is in fact the permutation which reverses the permutation of P :

$$P[1,2...k] = [i_1,i_2...i_k] \quad \Rightarrow \quad [1,2...k] = P^{-1}[i_1,i_2...i_k] \qquad PP^{-1} = P^{-1}P = 1 \quad . \tag{A.1.7}$$

In the above, since $[i_1,i_2...i_k]$ is a permutation of [1,2...k], one can get from [1,2...k] to $[i_1,i_2...i_k]$ by making some number of swaps of the integers in [1,2...k].

Comment: We are following a Maple convention that [a,b,c...] is a "list" where order is significant, whereas {a,b,c...} is a "set" where order is not significant.

The swap count S(P)

Any two permutations of [1,2,...k] can be linked by a number of pairwise swaps of the integers. For example, if we have $P[1,2,...k] = [i_1,i_2,...i_k]$, one can get from the first integer sequence to the second by doing some number S(P) of pairwise swaps. The integer S(P) is not unique, but whether it is an even or an odd integer is unique, so the factor $(-1)^{S(P)}$ is unique to a particular P (we leave it to the reader to prove this fact) . Sometimes $(-1)^{S(P)}$ is called the **parity** of permutation P.

Example: $[1,2,3] \rightarrow [2,1,3]$ $S(P) = 1$ $(-1)^{S(P)} = -1$

$[1,2,3] \rightarrow [1,3,2] \rightarrow [2,3,1] \rightarrow [2,1,3]$ $S(P) = 3$ $(-1)^{S(P)} = -1$ (A.1.8)

It seems clear that the number of position swaps to get from [1,2...k] to $[i_1,i_2...i_k]$ is the same as it is going the other direction, so

$$S(P^{-1}) = S(P) \quad . \qquad \text{(A.1.9)}$$

Finally, consider

$$P_1P_2[1,2...k] = P[1,2...k] = [i_1,i_2...i_k] \; . \qquad P = P_1P_2$$

If P_2 causes S_2 position swaps and then P_1 causes S_1 more, then P does S_2+S_1 total swaps. Thus

$$S(P) = S(P_1P_2) = S(P_1) + S(P_2)$$

so

$$(-1)^{S(P_1P_2)} = (-1)^{S(P_1)}\,(-1)^{S(P_2)} = (-1)^{S(P_2P_1)} \; . \qquad \text{(A.1.10)}$$

From the above these trivial corollaries follow :

$$(-1)^{S(PP)} = 1, \qquad (-1)^{S(PQ)} = (-1)^{S(QP)}.$$

$$(-1)^{S(P)}\,(-1)^{S(P)} = 1 \qquad (-1)^{S(P)} = (-1)^{S(Q)}\,(-1)^{S(PQ)} \qquad \text{(A.1.11)}$$

Fact: $\Sigma_P(-1)^{S(P)} = 0$. (A.1.12)

Proof : By the rearrangement theorem (A.1.3) and then (A.1.11) we know that, for any permutation Q,

$$\Sigma_P(-1)^{S(P)} = \Sigma_P(-1)^{S(QP)} = (-1)^{S(Q)}\Sigma_P(-1)^{S(P)} \; .$$

Select a Q which has $(-1)^{S(Q)} = -1$. Then $\Sigma_P(-1)^{S(P)} = -\,\Sigma_P(-1)^{S(P)} \Rightarrow \Sigma_P(-1)^{S(P)} = 0$. QED

Another Rearrangement Theorem

Another version of the rearrangement theorem is the following,

$$\Sigma_Q f(Q) = \Sigma_Q f(Q^{-1}) \quad . \tag{A.1.13}$$

Again, this is just a reordering of the sum. Consider,

$$\{g_1^{-1}, g_2^{-1}, ...g_k^{-1}\} = \{g_1', g_2', ...g_k'\} = \text{reordering of } \{g_1, g_2, ...g_k\} \quad .$$

To show that $\{g_1', g_2', ...g_k'\}$ is a reordering of $\{g_1, g_2, ...g_k\}$ we have to show that no two elements are the same. Suppose for example that $g_1' = g_2'$. That would say $g_1^{-1} = g_2^{-1}$ which in turn says $g_1 = g_2$, but that contradicts the basic starting point that $\{g_1, g_2, ...g_k\}$ enumerates the distinct group elements. Therefore,

$$\Sigma_i f(g_i) = \Sigma_i f(g_i^{-1}). \tag{A.1.14}$$

Comment: For continuous groups (like the rotation group SO(3)) , the rearrangement theorems become

$$\int dg\, f(g_a g) = \int dg\, f(g g_a) = \int dg\, f(g)$$

$$\int dg\, f(g) = \int dg\, f(g^{-1}) \tag{A.1.15}$$

where dg is called the invariant Haar measure. For SO(3) it is $dg = d\varphi d(\cos\theta)d\psi$ (Euler angles).

Determinants of a rank-2 tensor

It is well known that the determinant of a kxk matrix M can be written two equivalent ways in which the rows and columns are swapped (this is the statement that $\det(M) = \det(M^T)$),

$$\begin{aligned} \det(M_{**}) &= \Sigma_{i_1 i_2 \ldots i_k} \varepsilon_{i_1 i_2 \ldots i_k} M_{i_1 1} M_{i_2 2} \ldots M_{i_k k} \\ &= \Sigma_{i_1 i_2 \ldots i_k} \varepsilon_{i_1 i_2 \ldots i_k} M_{1 i_1} M_{2 i_2} \ldots M_{k i_k} \end{aligned} \tag{A.1.16}$$

where the permutation tensor ε is described below in Section A.6.

In permutation notation the above equations are written,

$$\begin{aligned} \det(M_{**}) &= \Sigma_P (-1)^{S(P)} M_{P(1)1} M_{P(2)2} \ldots M_{P(k)k} \\ &= \Sigma_P (-1)^{S(P)} M_{1P(1)} M_{2P(2)} \ldots M_{kP(k)} \quad . \end{aligned} \tag{A.1.17}$$

If we start over with the "up-tilt" matrix $M_a{}^b$ (mixed rank-2 tensor) then (A.1.16) becomes.

$$\det(M_*{}^*) = \Sigma_{i_1 i_2 \ldots i_k} \varepsilon_{i_1 i_2 \ldots i_k} M_{i_1}{}^1 M_{i_2}{}^2 \ldots M_{i_k}{}^k$$
$$= \Sigma_{i_1 i_2 \ldots i_k} \varepsilon_{i_1 i_2 \ldots i_k} M_1{}^{i_1} M_2{}^{i_2} \ldots M_k{}^{i_k} \qquad \text{(A.1.18)}$$

which in permutation notation becomes

$$\det(M_*{}^*) = \Sigma_P (-1)^{S(P)}\, M_{P(1)}{}^1 M_{P(2)}{}^2 \ldots M_{P(k)}{}^k$$
$$= \Sigma_P (-1)^{S(P)}\, M_1{}^{P(1)} M_2{}^{P(2)} \ldots M_k{}^{P(k)} \ . \qquad \text{(A.1.19)}$$

Statements for $\det(M^{**})$ and $\det(M^*{}_*)$ are similar,

$$\det(M^{**}) = \Sigma_P (-1)^{S(P)}\, M^{P(1)1} M^{P(2)2} \ldots M^{P(k)k}$$
$$= \Sigma_P (-1)^{S(P)}\, M^{1P(1)} M^{2P(2)} \ldots M^{kP(k)} \qquad \text{(A.1.20)}$$

$$\det(M^*{}_*) = \Sigma_P (-1)^{S(P)}\, M^{P(1)}{}_1 M^{P(2)}{}_2 \ldots M^{P(k)}{}_k$$
$$= \Sigma_P (-1)^{S(P)}\, M^1{}_{P(1)} M^2{}_{P(2)} \ldots M^k{}_{P(k)} \ . \qquad \text{(A.1.21)}$$

Notice that in all the equation pairs above, the second line involves the matrix transposes of elements in the first line. For example $(M^T)_{ab} = M_{ba}$ and $(M^T)_a{}^b = M_b{}^a$ where the two indices are just swapped, corresponding to a swap of rows and columns. So in all these cases we have $\det(M) = \det(M^T)$ for any position of the asterisk index position markers, for example, $\det(M^*{}_*) = \det([M^T]^*{}_*)$.

This conclusion is also true for the covariant transpose, so for example $\det(M^*{}_*) = \det([M^{\mathcal{T}}]^*{}_*)$, but this fact is less obvious. In the covariant transpose defined in (2.11.f.1) one has $M_{ab} = (M^{\mathcal{T}})_{ba}$ and $M_a{}^b = (M^{\mathcal{T}})^b{}_a$. The covariant transpose is formed not by swapping indices but by reflecting the indices in a vertical line between the indices. We shall now show that

Fact: $\det([M^{\mathcal{T}}]^*{}_*) = \det(M^*{}_*)$ and similarly for all other index positions. (A.1.22)

<u>Proof</u>: Consider that

$$\Sigma_{ij} g_{ai} M^i{}_j g^{jb} = M_a{}^b = (M^{\mathcal{T}})^b{}_a$$

where we use the raising and lowering functionality of the metric tensor shown in (2.2.1). In order to get all indices "down", define these three indices-down matrices,

$$(g_{dn})_{ab} = g_{ab} \qquad (g_{up})_{ab} = g^{ab} \qquad N_{ab} = M^a{}_b$$

so that

$$(M^{\mathcal{T}})^b{}_a = \Sigma_{ij}\, (g_{dn})_{ai}\, N_{ij}\, (g_{up})_{jb} = (g_{dn}\, N\, g_{up})_{ab} \ .$$

Then

$$\det([M^{\mathcal{T}}]^*{}_*) = \det[\, (g_{dn}\, N\, g_{up})_{**}] = \det(g_{dn}) \det(N_{**}) \det(g_{up}) \ .$$

As shown in (2.2.2), $g_{dn}g_{up} = 1$ so $\det(g_{dn})\det(g_{up}) = 1$, and therefore

$$\det([M^T]^*{}_*) = \det(N_{**}) = \det(M^*{}_*) \ .$$

A similar argument shows that $\det([M^T]_*{}^*) = \det(M_*{}^*)$. For both asterisks up or both down, there is no difference between M^T and M^{T} and we already know that $\det(M^{\mathrm{T}}) = \det(M)$.

The final conclusion is this:

Fact: $\det(M) = \det(M^{\mathrm{T}}) = \det(M^T)$ for any index positions. (A.1.23)

<u>Symmetric Sum Rearrangement Theorem</u>

For any particular permutation P of [1,2...k],

$$\Sigma_{i_1 i_2 ... i_k} f_{i_1 i_2 ... i_k} = \Sigma_{i_1 i_2 ... i_k} f_{i_{P(1)} i_{P(2)} ... i_{P(k)}}$$

or

$$\Sigma_I f_I = \Sigma_I f_{P(I)} \ . \qquad \text{// multiindex notation} \qquad \text{(A.1.24)}$$

<u>Proof</u>: Since Σ_I is a symmetric sum, one is free to shuffle the dummy summation index names at will, and this shuffle is indicated by permutation P. For example

$$\Sigma_I f_I = \Sigma_{i_1 i_2 ... i_k} f_{i_1 i_2 ... i_k} = \Sigma_{i_2 i_1 ... i_k} f_{i_2 i_1 ... i_k} = \Sigma_{i_1 i_2 ... i_k} f_{i_2 i_1 ... i_k}$$

$$= \Sigma_{i_1 i_2 ... i_k} f_{i_{P(1)} i_{P(2)} ... i_{P(k)}} = \Sigma_I f_{P(I)} \quad \text{where } P[1,2...k] = [2,1..k]$$

A.2 The Alt Operator in Generic Notation

The generic Alt operator acts on a function f of the integers [1,2....k] to create a new function, g = Alt(f), as follows:

$$g(1,2...k) = [\mathrm{Alt}(f)](1,2...k) \equiv (1/k!)\ \Sigma_P (-1)^{S(P)} f(\ P(1),P(2)...P(k)\ ) \qquad \text{(A.2.1)}$$

where f is any function and P are the k! permutations of the set of integers [1,2....k]. Informally we write the above as

$$g(1,2...k) = (1/k!)\ [\ f(1,2...k) - f(2,1...k) + \text{other signed permutations}\ ] \ .$$

<u>Examples</u>:

$$g(1,2) = [\mathrm{Alt}(f)](1,2) = (1/2)\ [\ f(1,2) - f(2,1)\ ] \qquad \text{(A.2.2)}$$

$$g(1,2,3) = [\mathrm{Alt}(f)](1,2,3) = (1/6)\ [\ f(1,2,3) - f(1,3,2) + f(3,1,2) - f(3,2,1) + f(2,3,1) - f(2,1,3)\ ] \ .$$

Now, let R be some permutation of [1,2....k]. We write

$$R[1,2....k] = [R(1),R(2).....R(k)] \qquad (A.2.3)$$

where for example R(1) gives the integer into which 1 is converted by the permutation R. We can apply the operator R to a function of 1,2...k in this manner,

$$R\, f(1,2,...k) = f(\, R(1),R(2).....R(k)\,)\ . \qquad (A.2.4)$$

Fact: Any permutation R is a linear operator, so $R(\, \Sigma_i a_i f_i) = \Sigma_i a_i (Rf_i)$ (A.2.5)

Proof: Let $h(1,2,...k) \equiv \Sigma_i a_i f_i(1,2,...k)$. Then

$R\, h(1,2,...k) = h(\, R(1),R(2).....R(k)\,)$ // (A.2.4) applied to h

$= \Sigma_i a_i f_i(\, R(1),R(2).....R(k)\,)$ // definition of h

$= \Sigma_i a_i\, R\, f_i(1,2,...k)$. // (A.2.4) applied to f_i QED

Now suppose R = QP, the product of two permutations Q and P. Then starting with (A.2.3),

$$(QP)\, f(1,2,...k) = f(\ (QP)(1),(QP)(2).....(QP)(k)\ )$$

$$= f(\ Q(P(1), Q(P(2)).... Q(P(k)\,)$$

$$\equiv\ f(\, QP(1), QP(2)... .QP(k)\,)\ . \qquad (A.2.6)$$

But from (A.2.4),

$$(QP)\, f(1,2,...k) = Q\, \{\, P\, f(1,2,...k)\} \ = Q\, f(\, P(1),P(2).....P(k)\,) \qquad (A.2.7)$$

Therefore we have shown that

$$Q\, f(\, P(1),P(2).....P(k)\,) \ =\ f(\, QP(1), QP(2)... .QP(k)\,)\ . \qquad (A.2.8)$$

Definition: A function f(1,2..k) is **totally antisymmetric** if it changes sign when any two arguments are swapped. (A.2.9)

Examples: f(1,2) = - f(2,1) ⇒ f is totally antisymmetric

f(1,2,3) = -f(2,1,3)
f(1,2,3) = -f(3,2,1) ⇒ f is totally antisymmetric
f(1,2,3) = -f(1,3,1)

Fact: f(1,2..k) totally antisymmetric $\Leftrightarrow$ P f(1,2,3..k) = $(-1)^{S(P)}$ f(1,2,3..k) (A.2.10)

Proof: [$\Rightarrow$] S(P) is the number of pairwise swaps going from [1,2,3...k] to P[1,2,3...k] = [$i_1,i_2,i_3...i_n$]. If f is totally antisymmetric by the definition above, each such swap causes a minus sign, and the product of these minus signs is then $(-1)^{S(P)}$. [$\Leftarrow$] If P = any pairwise swap, $(-1)^{S(P)}$ = -1, so f(1,2..k) is then totally antisymmetric.

Fact: The function g(1,2..k) ≡ [Alt(f)](1,2...k) is totally antisymmetric in its arguments. (A.2.11)

Proof: Let Q be some permutation of [1,2....k]. Then apply Q to the function g(1,2..k),

$$Q\, g(1,2..k) = Q \{ (1/k!) \Sigma_P (-1)^{S(P)} f(P(1),P(2)...P(k)) \} \quad \text{// definition of g}$$

$$= (1/k!) \Sigma_P (-1)^{S(P)}\, Q\, f(P(1),P(2)...P(k)) \quad \text{// (A.2.5), Q is linear}$$

$$= (1/k!) \Sigma_P (-1)^{S(P)}\, f(QP(1),QP(2)...QP(k)) \quad \text{// (A.2.8)}$$

$$= (-1)^{S(Q)}\, (1/k!) \Sigma_P\, (-1)^{S(QP)}\, f(QP(1),QP(2)...QP(k)) \quad \text{// (A.1.11)}$$

$$= (-1)^{S(Q)}\, (1/k!) \Sigma_P\, (-1)^{S(P)}\, f(P(1),P(2)...P(k)) \quad \text{// (A.1.3), rearrangement thm.}$$

$$= (-1)^{S(Q)}\, g(1,2..k) \ . \quad \text{// definition of g}$$

By (A.2.10$\Leftarrow$) it follows that g(1,2,..k) is totally antisymmetric. QED

Fact: Alt is a linear operator, so Alt($\Sigma_i a_i f_i$) = $\Sigma_i a_i$ Alt(f_i) . (A.2.12)

Proof: Let h(1,2,...k) ≡ $\Sigma_i a_i f_i$(1,2,...k). Then

$$[\text{Alt}(h)](1,2,...k) = (1/k!) \Sigma_P (-1)^{S(P)} h(P(1),P(2)...P(k)) \quad \text{// (A.2.1) def of Alt(h)}$$

$$= (1/k!) \Sigma_P (-1)^{S(P)} \{ \Sigma_i a_i f_i(P(1),P(2)...P(k)) \} \quad \text{// definition of h}$$

$$= \Sigma_i a_i\, [\, (1/k!) \Sigma_P (-1)^{S(P)}\, f_i(P(1),P(2)...P(k))\,] \quad \text{// reorder sums}$$

$$= \Sigma_i a_i\, \text{Alt}(f_i) \quad \text{// (A.2.1) def of Alt}(f_i)$$

Fact: Alt is a projection operator, so Alt(Alt(f)) = Alt(f) . (A.2.13)

Comment: This is why (1/k!) is included in the definition of Alt.

Proof: By (A.2.11) we know that Alt(f) is a totally antisymmetric function, and therefore from (A.2.10),

$$P\, [\text{Alt}(f)](1,2...k)] = (-1)^{S(P)} [\text{Alt}(f)](1,2...k)] \ . \quad \text{(A.2.14)}$$

Next, consider that

$[\text{Alt}(f)](P(1),P(2)...P(k)) = P\,[\text{Alt}(f)](1,2...k)$ // (A.2.4) applied with $f\rightarrow \text{Alt}(f)$, $R\rightarrow P$

$= (-1)^{S(P)}[\text{Alt}(f)](1,2...k)$. // (A.2.14) (A.2.15)

Now examine Alt(Alt(f)) :

$[\text{Alt}(\text{Alt}(f))](1,2...k) = (1/k!)\,\Sigma_P\,(-1)^{S(P)}[\text{Alt}(f)](P(1),P(2)...P(k))$

$= (1/k!)\,\Sigma_P\,(-1)^{S(P)}\{(-1)^{S(P)}[\text{Alt}(f)](1,2...k)]\}$ // (A.2.15)

$= (1/k!)\,\Sigma_P[\text{Alt}(f)](1,2...k)]\}$ // $(-1)^{S(P)}(-1)^{S(P)} = 1$

$= [\text{Alt}(f)](1,2...k)]\ \{(1/k!)\Sigma_P(1)\}$ // reorder factors

$= [\text{Alt}(f)](1,2...k)]\ \{1\}$. // (A.1.2) QED

Fact: If f is a totally antisymmetric function, then Alt(f) = f . (A.2.16)

Proof:

$\text{Alt}(f)(1,2...k) = (1/k!)\,\Sigma_P\,(-1)^{S(P)}f(P(1),P(2)...P(k))$ // definition of Alt(f) (A.2.1)

$= (1/k!)\,\Sigma_P\,(-1)^{S(P)}\,P\,f(1,2,...k)$ // (A.2.4) with $R\rightarrow P$

$= (1/k!)\,\Sigma_P\,(-1)^{S(P)}\,(-1)^{S(P)}\,f(1,2,...k)$ // (A.2.10)

$= (1/k!)\,\Sigma_P\,f(1,2,...k)$ // (A.1.11)

$= (1/k!)\ f(1,2,...k)\ \{\,\Sigma_P\,(1)\,\}$ // reorder

$= f(1,2,...k)$ // (A.1.2)

Fact: If f is totally antisymmetric, then

$f(1,2,3...k) = (-1)^{k-1}\ f(2, 3,... k-1, k, 1)$ forward cyclic (A.2.17)
$f(1,2,3...k) = (-1)^{k-1}\ f(k, 1, 2, 3,... k-1)$ backward cyclic

Proof: Let B be the particular permutation which does this: B[1,2,3..k-1,k] = [2,3,...k-1,k,1] (Backward cyclic). One then has S(B) = k-1 because it takes k-1 swaps to move the 1 from one end to the other. If f is totally antisymmetric, then according to (A.2.10) one has $B\,f(1,2,3..k) = (-1)^{S(B)}\ f(1,2,3..k)$ so then

$f(2,3,... k,1) = B\ f(1,2,3..k) = (-1)^{S(B)}\ f(1,2,3..k) = (-1)^{k-1}\ f(1,2,3...k)$.

On the other hand, if F[1,2,3..k-1,k] = [k, 1, 2, 3,... k-1] (Forward cyclic), S(F) = k-1 for the same reason, and then

$$f(k, 1, 2, 3,... k-1) = F\ f(1,2,3..k) = (-1)^{S(F)}\ f(1,2,3..k) = (-1)^{k-1}\ f(1,2,3...k)\ .$$

Example: A very commonly used fact is that, for k = 3, $(-1)^{k-1} = (-1)^2 = 1$ and so

$$f(1,2,3) = f(2,3,1) = f(3,1,2) \qquad \text{f totally antisymmetric} \qquad (A.2.18)$$

A.3 The Sym Operator in Generic Notation

This section is an obvious copy, paste and edit job on the previous section. We omit what would be (A.3.3) through (A.3.8) since they would be the same as (A.2.3) through (A.2.8). The changes are mainly these:

$$\text{antisymmetric} \rightarrow \text{symmetric} \qquad (-1)^{S(P)} \rightarrow 1 \qquad \text{Alt} \rightarrow \text{Sym}\ .$$

The Sym operator acts on a function of the integers [1,2....k] to create a new function, g = Sym(f), as follows:

$$g(1,2...k) = [\text{Sym}(f)](1,2...k) \equiv (1/k!)\ \Sigma_P f(\ P(1),P(2)...P(k)\ ) \qquad (A.3.1)$$

where f is any function and P are the k! permutations of the set of integers [1,2....k]. Informally we write the above as

$$g(1,2...k) = (1/k!)\ [\ g(1,2...k) + g(2,1...k) + \text{other permutations}\ ]$$

Examples:

$$g(1,2) = [\text{Sym}(f)](1,2) = (1/2)\ [\ f(1,2) + f(2,1)\ ] \qquad (A.3.2)$$

$$g(1,2,3) = [\text{Sym}(f)](1,2,3) = (1/6)\ [\ f(1,2,3) + f(1,3,2) + f(3,1,2) + f(3,2,1) + f(2,3,1) + f(2,1,3)\ ]$$

Definition: A function f(1,2..k) is **totally symmetric** if it is unchanged when any two arguments are swapped. (A.3.9)

Examples: f(1,2) = f(2,1) ⇒ f is totally symmetric

f(1,2,3) = f(2,1,3)
f(1,2,3) = f(3,2,1) ⇒ f is totally symmetric
f(1,2,3) = f(1,3,1)

Fact: f(1,2..k) totally symmetric ⇔ P f(1,2,3..k) = f(1,2,3..k) (A.3.10)

Proof: [⇒] S(P) is the number of pairwise swaps going from [1,2,3...k] to P[1,2,3...k] = $[i_1,i_2,i_3...i_n]$. If f is totally symmetric by the definition above, each such swap causes a plus sign, and the product of these plus signs is then 1. [⇐] If P = any pairwise swap, $(-1)^{S(P)}$ = 1, so f(1,2..k) is then totally symmetric.

Fact: The function g(1,2..k) ≡ [Sym(f)](1,2...k) is totally symmetric in its arguments. (A.3.11)

Proof: Let Q be some permutation of [1,2....k]. Then apply Q to the function g(1,2..k),

Q g(1,2..k) = Q { (1/k!) Σ_P f(P(1),P(2)...P(k)) } // definition of g

= (1/k!) Σ_P Q f(P(1),P(2)...P(k)) // (A.2.5), Q is linear

= (1/k!) Σ_P f(QP(1),QP(2)...QP(k)) // (A.2.8)

= (1/k!) Σ_P f(P(1),P(2)...P(k)) // (A.1.3), rearrangement thm.

= g(1,2..k) // definition of g

By (A.3.10⇐) it follows that g(1,2,..k) is totally symmetric. QED

Fact: Sym is a linear operator, so Sym($\Sigma_i a_i f_i$) = $\Sigma_i a_i$ Sym(f_i) (A.3.12)

Proof: Let h(1,2,...k) ≡ $\Sigma_i a_i f_i$(1,2,...k). Then

[Sym(h)](1,2,...k) = (1/k!) Σ_P h(P(1),P(2)...P(k)) // (A.3.1) def of Sym(h)

= (1/k!) { $\Sigma_i a_i f_i$(P(1),P(2)...P(k)) } // definition of h

= $\Sigma_i a_i$ [(1/k!) f_i(P(1),P(2)...P(k))] // reorder sums

= $\Sigma_i a_i$ Sym(f_i) // (A.3.1) def of Sym(f_i)

Fact: Sym is a projection operator, so Sym(Sym(f)) = Sym(f) . (A.3.13)

Comment: This is why (1/k!) is included in the definition of Sym.

Proof: By (A.3.11) we know that Sym is a totally symmetric function, and therefore from (A.3.10),

P [Sym(f)](1,2...k)] = [Sym(f)](1,2...k)] . (A.3.14)

Next, consider that

$[\text{Sym}(f)](P(1),P(2)...P(k)) = P\, [\text{Sym}(f)](1,2...k)]$ // (A.2.4) applied with $f \rightarrow \text{Sym}(f)$, $R \rightarrow P$

$= [\text{Sym}(f)](1,2...k)]$. // (A.3.14) (A.3.15)

Now examine Sym(Sym(f)) :

$[\text{Sym}(\text{Sym}(f))](1,2...k) = (1/k!)\, \Sigma_{\mathbf{P}}\, [\text{Sym}(f)](P(1),P(2)...P(k))$

$= (1/k!)\, \Sigma_{\mathbf{P}}\, \{[\text{Sym}(f)](1,2...k)]\}$ // (A.3.15)

$= [\text{Sym}(f)](1,2...k)]\ \{(1/k!)\Sigma_{\mathbf{P}}(1)\}$ // reorder

$= [\text{Sym}(f)](1,2...k)]\ \{1\}$ // (A.1.2) QED

Fact: If f is a totally symmetric function, then Sym(f) = f . (A.3.16)

<u>Proof</u>:

$\text{Sym}(f)(1,2...k) = (1/k!)\, \Sigma_{\mathbf{P}}\, f(P(1),P(2)...P(k))$ // definition of Sym(f) (A.3.1)

$= (1/k!)\, \Sigma_{\mathbf{P}}\ P\, f(1,2,...k)$ // (A.2.4) with $R \rightarrow P$

$= (1/k!)\, \Sigma_{\mathbf{P}}\ f(1,2,...k)$ // (A.3.10)

$= (1/k!)\ f(1,2,...k)\ \{\Sigma_{\mathbf{P}}(1)\}$ // reorder

$= f(1,2,...k)$ // (A.1.2)

Fact: If f is totally symmetric, then

$f(1,2,3...k) = f(2, 3,... k\text{-}1, k, 1)$ forward cyclic (A.3.17)
$f(1,2,3...k) = f(k, 1, 2, 3,... k\text{-}1)$ backward cyclic

This is just a special case of (A.3.10) which says Q f(1,2,3..k) = f(1,2,3..k) for *any* Q, so it certainly true for F = forward cyclic or B = backward cyclic permutations.

<u>Example</u>:

$f(1,2,3) = f(2,3,1) = f(3,1,2)$ f totally symmetric (A.3.18)

For comparison, recall (A.2.18) which said

$f(1,2,3) = f(2,3,1) = f(3,1,2)$ f totally antisymmetric *(A.2.18)*

A.4 Alt, Sym and decomposition of functions

Fact: The projection operators Alt and Sym are orthogonal, so Alt(Sym(f)) = Sym(Alt(f)) = 0 . (A.4.1)

Proof left: Alt(Sym(f)) = Alt({(1/k!) Σ_Pf(P(1),P(2)...P(k)) }

= (1/k!) Σ_P [Alt(f)](P(1),P(2)...P(k))] // (A.2.12), Alt is linear

= (1/k!) Σ_P $(-1)^{S(P)}$ [Alt(f)](1,2...k) // (A.2.15)

= {(1/k!) [Alt(f)](1,2...k)} {Σ_P $(-1)^{S(P)}$} // reorder

= {(1/k!) [Alt(f)](1,2...k)} {0} // (A.1.12)

= 0

Proof right: Sym(Alt(f)) = Sym({(1/k!) Σ_P $(-1)^{S(P)}$ f(P(1),P(2)...P(k)) }

= (1/k!) Σ_P $(-1)^{S(P)}$ [Sym(f)](P(1),P(2)...P(k))] // (A.3.12), Sym is linear

= (1/k!) Σ_P $(-1)^{S(P)}$ [Sym(f)](1,2...k) // (A.3.15)

= {(1/k!) [Sym(f)](1,2...k)} {Σ_P $(-1)^{S(P)}$} // reorder

= {(1/k!) [Sym(f)](1,2...k)} {0} // (A.1.12)

= 0

We can define a third projection operator this way,

Else() ≡ 1 - Alt() - Sym() // projection operator

Else(f) = f - Alt(f) - Sym(f) . // applied to f(1,2,3...k) (A.4.2)

One can then decompose an arbitrary function f into three pieces,

f = Alt(f) + Sym(f) + Else(f)
= a + s + e // a = a(1,2....k) etc (A.4.3)

where the "else" piece is whatever is left over, which is to say, e ≡ f - a - s. Then consider,

$$\text{Alt}(f) = \text{Alt}(a + s + e) = \text{Alt}(a) + \text{Alt}(s) + \text{Alt}(e) \qquad // \text{ (A.2.12), Alt is linear}$$

$$= \text{Alt}(\text{Alt}(f)) + \text{Alt}(\text{Sym}(f)) + \text{Alt}(\text{Else}(f)) \qquad // \text{ (A.4.3)}$$

$$= \text{Alt}(f) \quad + \quad 0 \quad + \text{Alt}(\text{Else}(f)) \qquad // \text{ (A.2.13) and (A.4.1)}$$

$$\Rightarrow \quad \text{Alt}(\text{Else}(f)) = 0 \ . \qquad \text{(A.4.4)}$$

$$\text{Sym}(f) = \text{Sym}(a + s + e) = \text{Sym}(a) + \text{Sym}(s) + \text{Sym}(e) \qquad // \text{ (A.2.12), Alt is linear}$$

$$= \text{Sym}(\text{Alt}(f)) + \text{Sym}(\text{Sym}(f)) + \text{Sym}(\text{Else}(f)) \qquad // \text{ (A.4.3)}$$

$$= \quad 0 \quad + \quad \text{Sym}(f) \quad + \text{Sym}(\text{Else}(f)) \qquad // \text{ (A.3.13) and (A.4.1)}$$

$$\Rightarrow \quad \text{Sym}(\text{Else}(f)) = 0 \ . \qquad \text{(A.4.5)}$$

This verifies that the "else" piece e of a function has neither a totally antisymmetric nor a totally symmetric component.

Example 1: For a function f(1,2) one has from (A.2.2) and (A.3.2),

$$a(1,2) = (1/2) \ [f(1,2) - f(2,1)]$$
$$s(1,2) = (1/2) [f(1,2) + f(2,1)] \qquad \Rightarrow \qquad e(1,2) = f(1,2) - a(1,2) - s(1,2) = 0 \qquad \text{(A.4.6)}$$

so the leftover else piece e(1,2) is null.

Example 2: On the other hand, for a function f(1,2,3) one has from (A.2.2) and (A.3.2)

$$a(1,2,3) = (1/6) [f(1,2,3) - f(1,3,2) + f(3,1,2) - f(3,2,1) + f(2,3,1) - f(2,1,3)] \qquad // \text{ (A.2.2)}$$
$$s(1,2,3) = (1/6) [f(1,2,3) + f(1,3,2) + f(3,1,2) + f(3,2,1) + f(2,3,1) + f(2,1,3)] \qquad // \text{ (A.3.2)}$$

$$e(1,2,3) = f(1,2,3) - (1/6) [f(1,2,3) - f(1,3,2) + f(3,1,2) - f(3,2,1) + f(2,3,1) - f(2,1,3)]$$
$$- (1/6) [f(1,2,3) + f(1,3,2) + f(3,1,2) + f(3,2,1) + f(2,3,1) + f(2,1,3)]$$

$$= f(1,2,3) - (1/3) [f(1,2,3) + f(3,1,2) + f(2,3,1)]$$

$$= (2/3) f(1,2,3) - (1/3) [f(3,1,2) + f(2,3,1)] \qquad \text{(A.4.7)}$$

so in this case the leftover piece e(1,2,3) is not null. In the case that f is either totally antisymmetric or totally symmetric, we know from (A.2.18) and (A.3.18) that all cyclic permutations of f are the same (for k = odd). In these cases, we can see explicitly from (A.4.7) that e(1,2,3) = 0, as expected.

A.5 Application to Tensors

We now restate the "generic" results of Sections A.2, A.3 and A.4 for this special case:

$$f(1,2...k) = T^{i_1 i_2 ... i_k} . \qquad \text{// a "tensor"} \qquad T \in V^k . \qquad (A.5.1)$$

Here T is any rank-k tensor (either in the weak or strong sense mentioned below (4.1.8)). This f seems perhaps an odd looking "function", but one can consider it to be just an evaluation of this more respectable mapping,

$$f(a,b,c,...q) = T^{i_a i_b i_c ... i_q} \qquad a,b,c... \in \{1,2...k\}$$

$$f: \{1,2...k\}^k \rightarrow V^k . \qquad (A.5.2)$$

This technical mapping issue is not important because we are just regarding $T^{i_1 i_2 ... i_k}$ as a "carrier" of the labels 1,2,3..k, from the point of view of doing permutations. The actual indices like i_1 could be arbitrary objects (labeled pancakes) as far as the permutation theorems are concerned, but in our applications we have in mind that i_1 is an integer in the range 1,2....n where n = dim(V) and n is unrelated to the tensor rank k.

In Section A.8 we shall instead apply our results to "tensor functions",

$$f(1,2,3...k) = T(v_{i_1}, v_{i_2}, ...v_{i_k}) .$$

Again, from a permutation point of view, $T(v_{i_1}, v_{i_2}, ...v_{i_k})$ is just a carrier of the labels 1,2...k. The permutation theorems don't care whether or not v_{i_1} happens to be a vector in V labeled by i_1, or even whether or not v_{i_1} happens to be an argument of a function T.

Here then are some Section A.2, A.3, A.4 results translated according to $f(1,2,3...k) = T^{i_1 i_2 ... i_k}$. For some of the translations, we show the actual equation from above, then its translation. For others we just state the translated result.

In all the results below, one can always specialize to the case $i_1,i_2...i_k \rightarrow 1,2,...k$. The resulting equations are then as if our mapping were $f(1,2...k) = T^{12...k}$. Note then that $i_{P(r)} \rightarrow i(r)$ in a superscript.

(a) Alt Equations (translated from Section A.2)

The basic Alt definition of (A.2.1)

$$g(1,2...k) = [Alt(f)](1,2...k) \equiv (1/k!) \Sigma_P (-1)^{S(P)} f(P(1),P(2)...P(k)) \qquad \textit{(A.2.1)}$$

becomes,

$G^{i_1i_2...i_k} = [Alt(F)]^{i_1i_2...i_k} = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, F^{i_{P(1)}i_{P(2)}...i_{P(k)}}$

or

$G = Alt(F)$. // definition of Alt acting on a tensor (A.5.3a)

We can define the object $Alt_I[F^{i_1i_2...i_k}]$ in the following obvious manner

$$Alt_I[F^{i_1i_2...i_k}] \equiv (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, F^{i_{P(1)}i_{P(2)}...i_{P(k)}}$$
$$= [Alt(F)]^{i_1i_2...i_k} \quad . \qquad (A.5.3b)$$

In multiindex notation :

$$[Alt(F)]^I = Alt_I[F^I] = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, F^{P(I)} \quad . \qquad (A.5.3c)$$

Examples:

$$G^{i_1i_2} = [Alt(F)]^{i_1i_2} = (1/2)\, [\, F^{i_1i_2} - F^{i_2i_1}\,] = Alt_I\, [F^{i_1i_2}]$$

$$G^{i_1i_2i_3} = [Alt(F)]^{i_1i_2i_3} = (1/6)\, [\, F^{i_1i_2i_3} - F^{i_2i_1i_3} + \text{the other four terms}\,] \quad . \qquad (A.5.4)$$

In practice, we might more easily write

$$G^{abc} = [Alt(F)]^{abc} = (1/6)\, (F^{abc} - F^{acb} + F^{cab} - F^{cba} + F^{bca} - F^{bac})$$

but when it comes time to prove permutation-related theorems, we use indices like ^{i_1i_2i_3} .

Continuing on, R[1,2....k] = [R(1),R(2).....R(k)] of (A.2.3) becomes, with R→ P,

$$P\, T^{i_1i_2...i_k} = T^{i_{P(1)}i_{P(2)}...i_{P(k)}} \qquad \textit{(A.2.3)} \qquad (A.5.5)$$

Fact: Any permutation P is a linear operator, so $P(\Sigma_i a_i T_i{}^{i_1i_2...i_k}) = \Sigma_i a_i (PT_i{}^{i_1i_2...i_k})$.
(A.2.5) (A.5.6)

Definition: A tensor $T^{i_1i_2...i_k}$ is **totally antisymmetric** if it changes sign when any two superscripts are swapped. *(A.2.9)* (A.5.7)

Example: $T^{i_1i_2...i_k} = -\, T^{i_2i_1...i_k}$ or $T^{abc} = -T^{bac}$

Fact: $T^{i_1i_2...i_k}$ totally antisymmetric ⇔ $P\, T^{i_1i_2...i_k} = (-1)^{S(P)}\, T^{i_1i_2...i_k}$, where P is any permutation of [1,2..k]. *(A.2.10)* (A.5.8)

Fact: The function $T^{i_1i_2...i_k} \equiv [Alt(F)]^{i_1i_2...i_k}$ is totally antisymmetric in its indices.
(A.2.11) (A.5.9)

Fact: Alt is a linear operator, so $Alt\, (\Sigma_j a_j T_j{}^{i_1i_2...i_k}) = \Sigma_j a_j\, [Alt(T_j)]^{i_1i_2...i_k}$.

(A.2.12) (A.5.10)

Fact: Alt is a projection operator, so Alt(Alt(T)) = Alt(T) . *(A.2.13)* (A.5.11)

Fact: If T is a totally antisymmetric rank-k tensor, then Alt(T) = T . *(A.2.16)* (A.5.12)

(b) Sym Equations (translated from Section A.3)

The basic Sym definition of (A.3.1)

$$g(1,2...k) = [\mathrm{Sym}\ (f)](1,2...k) \equiv (1/k!)\ \Sigma_P\ f(\ P(1),P(2)...P(k)\ ) \qquad (A.3.1)$$

becomes,

$$G^{i_1 i_2 \cdots i_k} = [\mathrm{Sym}\ (F)]^{i_1 i_2 \cdots i_k} = (1/k!)\ \Sigma_P\ F^{i_{P(1)} i_{P(2)} \cdots i_{P(k)}} \qquad (A.5.13)$$

G = Sym (F) . // definition of Sym acting on a tensor

Examples:

$$G^{i_1 i_2} = [\mathrm{Sym}(F)]^{i_1 i_2} = (1/2)\ [\ F^{i_1 i_2} + F^{i_2 i_1}\ ]$$

$$G^{i_1 i_2 i_3} = [\mathrm{Sym}(F)]^{i_1 i_2 i_3} = (1/6)\ [\ F^{i_1 i_2 i_3} + F^{i_2 i_1 i_3} + \text{the other four terms}\ ] \qquad (A.5.14)$$

$$G^{abc} = [\mathrm{Sym}(F)]^{abc} = (1/6)\ (F^{abc} + F^{acb} + F^{cab} + F^{cba} + F^{bca} + F^{bac})$$

Definition: A tensor $T^{i_1 i_2 \cdots i_k}$ is **totally symmetric** if it is unchanged when any two superscripts are swapped. *(A.3.9)* (A.5.15)

Example: $T^{i_1 i_2 \cdots i_k} = T^{i_2 i_1 \cdots i_k}$ or $T^{abc} = T^{bac} = T^{acb}$

Fact: $T^{i_1 i_2 \cdots i_k}$ totally symmetric ⇔ P $T^{i_1 i_2 \cdots i_k} = T^{i_1 i_2 \cdots i_k}$, where P is any permutation of [1,2..k]. *(A.3.10)* (A.5.16)

Fact: The function $T^{i_1 i_2 \cdots i_k} \equiv [\mathrm{Sym}(F)]^{i_1 i_2 \cdots i_k}$ is totally symmetric in its indices.
(A.3.11) (A.5.17)

Fact: Sym is a linear operator, so Sym $(\Sigma_j a_j T_j^{i_1 i_2 \cdots i_k}) = \Sigma_j a_j\ [\mathrm{Sym}(T_j)]^{i_1 i_2 \cdots i_k}$.
(A.3.12) (A.5.18)

Fact: Sym is a projection operator, so Sym (Sym (T)) = Sym (T) . *(A.3.13)* (A.5.19)

Fact: If T is a totally symmetric rank-k tensor, then Sym(T) = T . *(A.3.16)* (A.5.20)

(c) Alt, Sym and decomposition of tensors (translated from Section A.4)

Fact: The projection operators Alt and Sym are orthogonal, so Alt(Sym(f)) = Sym(Alt(f)) = 0 .
(A.4.1) (A.5.21)

Fact: A tensor $T^{i_1 i_2 \ldots i_k}$ can be decomposed in the following manner: (A.5.22)

$$T^{i_1 i_2 \ldots i_k} = A^{i_1 i_2 \ldots i_k} + S^{i_1 i_2 \ldots i_k} + E^{i_1 i_2 \ldots i_k} \qquad \textit{(A.4.3)}$$

Alt(A) = A	Sym(A) = 0	Alt(E) = 0	*(A.4.4)*
Alt(S) = 0	Sym(S) = S	Sym(E) = 0	*(A.4.5)*

where A is totally antisymmetric, S is totally symmetric, and E is whatever is left over.

A.6 The permutation tensor ε

The permutation tensor ε of rank k is written $\varepsilon_{i_1 i_2 \ldots i_k}$ where each subscript *must be* an element of {1,2...k}. The values of the tensor are these:

- $\varepsilon_{12..k} = +1$
- $\varepsilon_{i_1 i_2 \ldots i_k}$ changes sign if any two indices are swapped .
- Therefore, if two or more indices are the same, $\varepsilon_{i_1 i_2 \ldots i_k} = 0$.
- $\varepsilon^{i_1 i_2 \ldots i_k} \equiv \varepsilon_{i_1 i_2 \ldots i_k}$ (A.6.1)

The tensor $\varepsilon_{i_1 i_2 \ldots i_k}$ has k^k components, but only k! of those components are non-zero. One arrives at k! by allowing k values for i_1, then only (k-1) values for i_2, and so on.

The tensor $\varepsilon_{i_1 i_2 \ldots i_k}$ is totally antisymmetric by (A.5.7) since any index swap causes a minus sign.

Fact: Apart from scale, the $\varepsilon_{i_1 i_2 \ldots i_k}$ tensor is the *only* totally antisymmetric tensor one can construct.
(A.6.2)

Proof: From the definition of $\varepsilon_{i_1 i_2 \ldots i_k}$, we see that if $A^{i_1 i_2 \ldots i_k}$ is a *arbitrary* totally antisymmetric tensor, then one can write

$$A^{i_1 i_2 \ldots i_k} = [\, A^{12 \ldots k}] \;\; \varepsilon^{i_1 i_2 \ldots i_k} . \qquad \text{(A.6.3)}$$

Here $\varepsilon^{i_1 i_2 \ldots i_k}$ does the bookkeeping for swaps of index pairs. The scale factor is then $A^{12 \ldots k}$.

Use of the ε tensor

We noted already that all our permutation results can be specialized to $i_r \rightarrow r$. For example,

$$[\text{Alt}(T)]^{i_1 i_2 \ldots i_k} = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, T^{i_{P(1)} i_{P(2)} \ldots i_{P(k)}} \qquad \textit{(A.5.3b)}$$

then becomes

$$[\text{Alt}(T)]^{12\ldots k} = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, T^{P(1)P(2)\ldots P(k)} \quad . \qquad \text{(A.6.4)}$$

Now we make the following claim,

$$\Sigma_P\, (-1)^{S(P)}\, T^{P(1)P(2)\ldots P(k)} = \Sigma_{i_1 i_2 \ldots i_k}\, \varepsilon_{i_1 i_2 \ldots i_k}\, T^{i_1 i_2 \ldots i_k}, \quad i_r = 1,2\ldots k \quad . \qquad \text{(A.6.5)}$$

Each of the i_r sums runs from 1 to k. Notice that each side has k! non-vanishing terms in its sum.

Suppose $P[1,2,3\ldots k] = [i_1, i_2, i_3 \ldots i_k]$. Then we claim that the parity of the permutation is given by

$$(-1)^{S(P)} = \varepsilon_{i_1 i_2 \ldots i_k} \, . \qquad \text{(A.6.6)}$$

To see why this is so, start off with the identity permutation P = 1 which has $(-1)^{S(P)} = (-1)^0 = 1$. In this case $P[1,2\ldots k] = [1,2\ldots k]$ and conveniently $\varepsilon_{123\ldots k} = 1$, so both sides of (A.6.6) agree. Now swap 1↔2 and then the left side is $(-1)^1 = -1$ and the right side is $\varepsilon_{213\ldots k} = -\, \varepsilon_{123\ldots k} = -\, 1$ and again both sides agree. Now swap 2↔3. The left side is $(-1)^2$ and the right side is $-\, \varepsilon_{132\ldots k} = \varepsilon_{123\ldots k} = 1$, and again both sides agree. In this way one can exhaust all permutations P and the equation is always true.

On the left side of (A.6.5) the permutations are enumerated by P, while on the right they are enumerated by $i_1, i_2, i_3 \ldots i_k$ which is restricted by the ε tensor to be a permutation of 1,2,3...k.

Basically the notation on each side of (A.6.5) is describing the same instructions for forming the sum.

<u>Example with k = 3</u> (A.6.7)

$$\Sigma_P\, (-1)^{S(P)}\, T^{P(1)P(2)P(3)}$$

$$= T^{123} - T^{213} + T^{231} - T^{321} + T^{312} - T^{132} \, .$$

The only simple way to form this sum is to keep doing swaps. We show in red the pair that will be swapped to make the next term on the right. Compare then to

$$\Sigma_{i_1 i_2 i_3}\, \varepsilon_{i_1 i_2 i_3}\, T^{i_1 i_2 i_k} \quad .$$

To enumerate the terms, we use the 3! = 6 non-zero values of $\varepsilon_{i_1 i_2 i_3}$ in the same order as above

$$\Sigma_{i_1 i_2 i_3}\ \varepsilon_{i_1 i_2 i_3}\ T^{i_1 i_2 i_k}$$

$$= \ \varepsilon_{123}\ T^{123} + \varepsilon_{213}\ T^{213} + \varepsilon_{231}\ T^{231} + \varepsilon_{321} T^{321} + \varepsilon_{312}\ T^{312} + \varepsilon_{132}\ T^{132}$$

$$= \ (1)\ T^{123} + (-1)\ T^{213} + (1)\ T^{231} + (-1)\ T^{321} + (1)\ T^{312} + (-1) T^{132}$$

$$= T^{123} - T^{213} + T^{231} - T^{321} + T^{312} - T^{132}\ .$$

Here the signs of the ε factors alternate as shown because each one is obtained by an index pair swap on the preceding term.

Application Consider,

$$T_{i_1 i_2 \ldots i_k} = (v_{i_1} \otimes v_{i_2} \otimes \ldots.. \otimes v_{i_k})\ .$$

We can specialize this to say

$$T_{12\ldots k} = (v_1 \otimes v_2 \otimes \ldots.. \otimes v_3)$$

and then apply P using (A.2.4) with R→P to get

$$T_{P(1)P(2)\ldots P(k)} = (v_{P(1)} \otimes v_{P(2)} \otimes \ldots.. \otimes v_{P(k)})\ .$$

Then (A.6.5)

$$\Sigma_P\ (-1)^{S(P)}\ T_{P(1)P(2)\ldots P(k)} = \Sigma_{i_1 i_2 \ldots i_k}\ \varepsilon_{i_1 i_2 \ldots i_k}\ T_{i_1 i_2 \ldots i_k}, \quad i_r = 1,2\ldots k \qquad \textit{(A.6.5)}$$

becomes

$$\Sigma_P\ (-1)^{S(P)}\ (v_{P(1)} \otimes v_{P(2)} \otimes \ldots.. \otimes v_{P(k)})$$

$$= \Sigma_{i_1 i_2 \ldots i_k}\ \varepsilon_{i_1 i_2 \ldots i_k}\ (v_{i_1} \otimes v_{i_2} \otimes \ldots.. \otimes v_{i_k}), \quad i_r = 1,2\ldots k \qquad \text{(A.6.8)}$$

which appears as part of (7.1.3).

A.7 The wedge-product-of-vectors Alt equation

Here we consider a new application for our generic function f[1,2...k], namely,

$$f[1,2,....k] = (v_{j_1} \otimes v_{j_2} \otimes .....\otimes v_{j_k}) \quad . \tag{A.7.1}$$

Here the j_s label the generic objects v_{j_s} and $\otimes$ is for the moment some generic operator. Then, consider the generic Alt definition,

$$[Alt(f)](1,2...k) \equiv (1/k!)\, \Sigma_P (-1)^{S(P)} f(\, P(1),P(2)...P(k)\,) \quad . \qquad \textit{(A.2.1)}$$

Formally speaking, the left side would have to be written something like this,

$$[Alt(f)](1,2...k) = [Alt((v_{j*} \otimes v_{j*} \otimes .....\otimes v_{j*}))](1,2...k)$$

$$= Alt(v_{j_1} \otimes v_{j_2} \otimes .....\otimes v_{j_k}) \quad . \tag{A.7.2}$$

On the first line the asterisks are place holders which will get the arguments in the argument list. This lets us make a formal association $f \rightarrow (v_{j*} \otimes v_{j*} \otimes .....\otimes v_{j*})$ for a function without arguments.

On the right side of (A.2.1) just above we have

$$f(\, P(1),P(2)...P(k)\,) = (v_{j_{P(1)}} \otimes v_{j_{P(2)}} \otimes .....\otimes v_{j_{P(k)}}) \quad . \tag{A.7.3}$$

From the last three equations we end up then with this statement,

$$Alt(v_{j_1} \otimes v_{j_2} \otimes .....\otimes v_{j_k}) = (1/k!)\, \Sigma_P (-1)^{S(P)} (v_{j_{P(1)}} \otimes v_{j_{P(2)}} \otimes .....\otimes v_{j_{P(k)}}) . \tag{A.7.4}$$

If it happens that $\otimes$ means the tensor product, and if the v_{j_1} happen to be vectors in V, then the above expression happens to be our definition (7.1.2) for the wedge product of k vectors:

$$(v_{j_1} \wedge v_{j_2} \wedge .....\wedge v_{j_k}) = Alt(v_{j_1} \otimes v_{j_2} \otimes .....\otimes v_{j_k}) \quad . \tag{A.7.5}$$

As noted earlier, we can always specialize replacing $j_r \rightarrow r$. Then

$$f[1,2,....k] = (v_1 \otimes v_2 \otimes .....\otimes v_k) \quad .$$

$$Alt(v_1 \otimes v_2 \otimes .....\otimes v_k) = (1/k!)\, \Sigma_P (-1)^{S(P)} (v_{P(1)} \otimes v_{P(2)} \otimes .....\otimes v_{P(k)}) \tag{A.7.6}$$

$$(v_1 \wedge v_2 \wedge .....\wedge v_k) = Alt(v_1 \otimes v_2 \otimes .....\otimes v_k) \quad . \tag{A.7.7}$$

A.8 Application to Tensor Functions

We now restate the "generic" results of Sections A.2, A.3 and A.4 for this special case:

$$f(1,2...k) = \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) \qquad \text{// a "tensor function"} \qquad \mathcal{T} \in V^{*k} \ . \qquad (A.8.1)$$

We apologize for copy, paste and edit, but things really are exactly parallel to the tensor discussion above.

Here $\mathcal{T}$ is any rank-k tensor function. This f seems perhaps an odd looking "function", but one can consider it to be just an evaluation of this more respectable mapping,

$$f(a,b,c,...q) = \mathcal{T}(v_{i_a},v_{i_b}....v_{i_q}) \qquad a,b,c... \in \{1,2...k\}$$

$$f: \{1,2...k\}^k \rightarrow V_f{}^{*k} \ . \qquad (A.8.2)$$

This technical mapping issue is not important because we are just regarding $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$ as a "carrier" of the labels 1,2,3..k, from the point of view of doing permutations. The actual indices like i_1 could be arbitrary objects (labeled cupcakes) as far as the permutation theorems are concerned, but in our applications we have in mind that i_1 is an integer in the range 1,2....n where n = dim(V) and n is unrelated to the tensor rank k.

Here then are some Section A.2, A.3 ,A.4 results translated according to $f(1,2,3...k) = \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$. For some of the translations, we show the actual equation from above, then its translation. For others we just state the translated result.

In all the results below, one can always specialize to the case $i_1,i_2...i_k \rightarrow 1,2,...k$. The resulting equations are then as if our mapping were $f(1,2...k) = \mathcal{T}(v_1,v_2....v_k)$. Note then that $i_{P(r)} \rightarrow i(r)$ in a subscript.

(a) Alt Equations (translated from Section A.2)

The basic Alt definition of (A.2.1)

$$g(1,2...k) = [Alt(f)](1,2...k) \equiv (1/k!)\ \Sigma_P\ (-1)^{S(P)} f(\ P(1),P(2)...P(k)\ ) \qquad \textit{(A.2.1)}$$

becomes,

$$\mathcal{G}(v_{i_1},v_{i_2}....v_{i_k}) = [Alt(\mathcal{F})](v_{i_1},v_{i_2}....v_{i_k}) = (1/k!)\ \Sigma_P\ (-1)^{S(P)} \mathcal{F}(v_{i_{P(1)}},v_{i_{P(2)}}....v_{i_{P(k)}}) \qquad (A.8.3)$$

$\mathcal{G} = Alt(\mathcal{F})$. // definition of Alt acting on a tensor function

Examples:

$\mathcal{G}(v_{i_1},v_{i_2}) = [\mathrm{Alt}(\mathcal{F})](v_{i_1},v_{i_2}) = (1/2)[\ \mathcal{F}(v_{i_1},v_{i_2}) - \mathcal{F}(v_{i_2},v_{i_1})\ ]$ (A.8.4)

$\mathcal{G}(v_{i_1},v_{i_2},v_{i_3}) = [\mathrm{Alt}(\mathcal{F})](v_{i_1},v_{i_2},v_{i_3}) = (1/6)[\mathcal{F}(v_{i_1},v_{i_2},v_{i_3}) - \mathcal{F}(v_{i_2},v_{i_1},v_{i_3})+$ the other four terms]

In practice, we might more easily write

$\mathcal{G}(v_a,v_b,v_c) = [\mathrm{Alt}(\mathcal{F})](v_a,v_b,v_c) = (1/6)\ [\ \mathcal{F}(v_a,v_b,v_c) - \mathcal{F}(v_b,v_a,v_c) +$ the other four terms]

but when it comes time to prove permutation-related theorems, we use subscripts like i_1.

Continuing on, R[1,2....k] = [R(1),R(2).....R(k)] of (A.2.3) becomes, with R→ P,

$P\ \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = \mathcal{T}(v_{i_{P(1)}},v_{i_{P(2)}}....v_{i_{P(k)}})$. *(A.2.3)* (A.8.5)

Fact: Any permutation P is a linear operator, so $P(\ \Sigma_r a_r \mathcal{T}_r(v_{i_1},v_{i_2}....v_{i_k})) = \Sigma_r a_r (P\mathcal{T}_r(v_{i_1},v_{i_2}....v_{i_k}))$
(A.2.5) (A.8.6)

Definition: A tensor function $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$ is **totally antisymmetric** if it changes sign when any two arguments are swapped. *(A.2.9)* (A.8.7)

Example: $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = - \mathcal{T}(v_{i_2},v_{i_1}....v_{i_k})$ or $\mathcal{T}(v_a,v_b....v_q) = - \mathcal{T}(v_b,v_a....v_q)$

Fact: $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$ totally antisymmetric ⇔

$P\ \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = (-1)^{S(P)}\ \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$, where P is any permutation of [1,2..k].
(A.2.10) (A.8.8)

Fact: The function $\mathcal{G}(v_{i_1},v_{i_2}....v_{i_k}) \equiv [\mathrm{Alt}(\mathcal{T})](v_{i_1},v_{i_2}....v_{i_k})$ is totally antisymmetric in its labels.
(A.2.11) (A.8.9)

Fact: Alt is a linear operator, so $\mathrm{Alt}\ (\Sigma_j a_j \mathcal{T}_j(v_{i_1},v_{i_2}....v_{i_k})) = \Sigma_j a_j\ [\mathrm{Alt}(\mathcal{T}_j)](v_{i_1},v_{i_2}....v_{i_k})$.
(A.2.12) (A.8.10)

Fact: Alt is a projection operator, so $\mathrm{Alt}(\mathrm{Alt}(\mathcal{T})) = \mathrm{Alt}(\mathcal{T})$. *(A.2.13)* (A.8.11)

Fact: If $\mathcal{T}$ is a totally antisymmetric rank-k tensor, then $\mathrm{Alt}(\mathcal{T}) = \mathcal{T}$. *(A.2.16)* (A.8.12)

(b) Sym Equations (translated from Section A.3)

The basic Sym definition of (A.3.1)

$$g(1,2...k) = [\mathrm{Sym}\ (f)](1,2...k) \equiv (1/k!)\ \Sigma_P\ f(\ P(1),P(2)...P(k)\ ) \qquad \textit{(A.3.1)}$$

becomes,

$$\mathcal{G}(v_{i_1},v_{i_2}....v_{i_k}) = [\mathrm{Sym}(\mathcal{F})](v_{i_1},v_{i_2}....v_{i_k}) = (1/k!)\ \Sigma_P\ \mathcal{F}(v_{i_{P(1)}},v_{i_{P(2)}}....v_{i_{P(k)}}) \qquad \text{(A.8.13)}$$

$\mathcal{G} = \mathrm{Sym}(\mathcal{F})$. // definition of Alt acting on a tensor

<u>Examples</u>:

$$\mathcal{G}(v_{i_1},v_{i_2}) = [\mathrm{Sym}(\mathcal{F})](v_{i_1},v_{i_2}) = (1/2)[\ \mathcal{F}(v_{i_1},v_{i_2}) + \mathcal{F}(v_{i_2},v_{i_1})\ ] \qquad \text{(A.8.14)}$$

$$\mathcal{G}(v_{i_1},v_{i_2},v_{i_3}) = [\mathrm{Sym}(\mathcal{F})](v_{i_1},v_{i_2},v_{i_3}) = (1/6)[\mathcal{F}(v_{i_1},v_{i_2},v_{i_3})+\mathcal{F}(v_{i_2},v_{i_1},v_{i_3}+ \text{the other four terms}\ ]$$

In practice, we might more easily write

$$\mathcal{G}(v_a,v_b,v_c) = [\mathrm{Sym}(\mathcal{F})](v_a,v_b,v_c) = (1/6)\ [\ \mathcal{F}(v_a,v_b,v_c) + \mathcal{F}(v_b,v_a,v_c) + \text{the other four terms}\ ]$$

but when it comes time to prove permutation-related theorems, we use subscripts like i_1.

<u>Definition</u>: A tensor function $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$ is **totally symmetric** if it is unchanged when any two arguments are swapped. *(A.3.9)* (A.8.15)

<u>Example</u>: $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = \mathcal{T}(v_{i_2},v_{i_1}....v_{i_k})$ or $\mathcal{T}(v_a,v_b....v_q) = \mathcal{T}(v_b,v_a....v_q)$

Fact: $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$ totally symmetric $\Leftrightarrow$ $P\ \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$, where P is any permutation of [1,2..k]. *(A.3.10)* (A.8.16)

Fact: The function $\mathcal{G}(v_{i_1},v_{i_2}....v_{i_k}) \equiv [\mathrm{Sym}(\mathcal{F})](v_{i_1},v_{i_2}....v_{i_k})$ is totally symmetric in its labels. *(A.3.11)* (A.8.17)

Fact: Sym is a linear operator, so $\mathrm{Sym}(\Sigma_j a_j \mathcal{T}_j(v_{i_1},v_{i_2}....v_{i_k})) = \Sigma_j a_j\ [\mathrm{Sym}(\mathcal{T}_j)](v_{i_1},v_{i_2}....v_{i_k})$. *(A.3.12)* (A.8.18)

Fact: Sym is a projection operator, so $\mathrm{Sym}\ (\mathrm{Sym}\ (\mathcal{T})) = \mathrm{Sym}\ (\mathcal{T})$. *(A.3.13)* (A.8.19)

Fact: If $\mathcal{T}$ is a totally symmetric rank-k tensor, then $\mathrm{Sym}(\mathcal{T}) = \mathcal{T}$. *(A.2.16)* (A.8.20)

(c) Alt/Sym and Other Equations (translated from Section A.4, A.6 and A.7)

Fact: The projection operators Alt and Sym are orthogonal, so $\mathrm{Alt}(\mathrm{Sym}(\mathcal{T})) = \mathrm{Sym}(\mathrm{Alt}(\mathcal{T})) = 0$. *(A.4.1)* (A.8.21)

Fact: A tensor function $\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$ can be decomposed in the following manner: (A.8.22)

$$\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = \mathcal{A}(v_{i_1},v_{i_2}....v_{i_k}) + \mathcal{S}(v_{i_1},v_{i_2}....v_{i_k}) + \mathcal{E}(v_{i_1},v_{i_2}....v_{i_k}) \qquad \textit{(A.4.3)}$$

$$\mathrm{Alt}(\mathcal{A}) = \mathcal{A} \qquad \mathrm{Sym}(\mathcal{A}) = 0 \qquad \mathrm{Alt}(\mathcal{E}) = 0 \qquad \textit{(A.4.4)}$$
$$\mathrm{Alt}(\mathcal{S}) = 0 \qquad \mathrm{Sym}(\mathcal{S}) = \mathcal{S} \qquad \mathrm{Sym}(\mathcal{E}) = 0 \qquad \textit{(A.4.5)}$$

where $\mathcal{A}$ is totally antisymmetric, $\mathcal{S}$ is totally symmetric, and $\mathcal{E}$ is whatever is left over.

The following are based on Section A.6 and concern use of the ε tensor with tensor functions.

If $\mathcal{A}$ is totally antisymmetric, then

$$\mathcal{A}(v_{i_1},v_{i_2}....v_{i_k}) = [\, \mathcal{A}(v_1,v_2....v_k)]\ \ \varepsilon_{i_1 i_2 ... i_k} . \qquad \textit{(A.6.3)} \qquad \text{(A.8.23)}$$

The tensor function $[\mathrm{Alt}(\mathcal{T})](v_1,v_2....v_k)$ can be expressed as,

$$\Sigma_P (-1)^{S(P)}\ \mathcal{T}(v_{P(1)},v_{P(2)}....v_{P(k)}) = \Sigma_{i_1 i_2 ... i_k}\ \varepsilon_{i_1 i_2 ... i_k}\ \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k})$$
(A.6.5) (A.8.24)

Let

$$\mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) = (\alpha_{i_1} \otimes \alpha_{i_2} \otimes ..... \otimes \alpha_{i_k})(v_{i_1},v_{i_2}....v_{i_k})$$

so

$$\mathcal{T}(v_1,v_2....v_k) = (\alpha_1 \otimes \alpha_2 \otimes ..... \otimes \alpha_k)(v_1,v_2....v_k) .$$

Then (A.8.24) gives this way to write $(\alpha_{j_1} \wedge \alpha_{j_2} \wedge .....\wedge \alpha_{j_k})$:

$$\Sigma_P (-1)^{S(P)}\ (\alpha_{P(1)} \otimes \alpha_{P(2)} \otimes ..... \otimes \alpha_{P(k)}) \qquad \textit{(A.6.8)} \qquad \text{(A.8.25)}$$

$$= \Sigma_{i_1 i_2 ... i_k}\ \varepsilon^{i_1 i_2 ... i_k} (\alpha_{i_1} \otimes \alpha_{i_2} \otimes ..... \otimes \alpha_{i_k}), \quad i_r = 1,2...k .$$

The following is based on Section A.7,

$$(\alpha_{j_1} \wedge \alpha_{j_2} \wedge .....\wedge \alpha_{j_k}) = \mathrm{Alt}(\alpha_{j_1} \otimes \alpha_{j_2} \otimes .....\otimes \alpha_{j_k}) \qquad \textit{(A.7.5)} \qquad \text{(A.8.26)}$$

(d) Alt/Sym when there are two sets of indices

It is not uncommon to encounter objects like the following

$(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$ $\qquad\qquad$ $X_J{}^I$

where the j_r are labels and the i_r are tensor component indices. An example would be the components of a wedge product of k vectors

$(v_{j_1} \wedge v_{j_2} \wedge \ldots\ldots \wedge v_{j_k})^{i_1 i_2 \ldots i_k}$.

In this situation, we have to clarify which of the two sets of indices is being acted upon by the Alt operator. We might do this as follows, using Alt_I and Alt_J ,

$$Alt_I[(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}] = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\ (X_{j_1 j_2 \ldots j_k})^{i_{P(1)} i_{P(2)} \ldots i_{P(k)}}$$

$$Alt_J[(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}] = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\ (X_{j_{P(1)} j_{P(2)} \ldots j_{P(k)}})^{i_1 i_2 \ldots i_k}$$

In general, the above two objects are different.

Now recall Fact (A.5.12) from above,

Fact: If T is a totally antisymmetric rank-k tensor, then Alt(T) = T . *(A.2.16)* *(A.5.12)*

If it happens that $(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$ is totally antisymmetric in the i_r , then

$$Alt_I[(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}] = (X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$$

If it happens that $(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$ is totally antisymmetric in the j_r , then

$$Alt_J[(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}] = (X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$$

If it happens that $(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$ is totally antisymmetric separately in the i_r *and* the j_r, then we have

$$Alt_I[(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}] = Alt_J[(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}]$$

since both are equal to $(X_{j_1 j_2 \ldots j_k})^{i_1 i_2 \ldots i_k}$.

Recalling Fact (A.5.20),

Fact: If T is a totally symmetric rank-k tensor, then Sym(T) = T , *(A.3.16)* *(A.5.20)* we conclude a similar fact for Sym_I and Sym_J . We then summarize these results

Fact: If $(X_{j_1j_2...j_k})^{i_1i_2...i_k}$ is totally antisymmetric in both sets of indices, then

$$\text{Alt}_I[(X_{j_1j_2...j_k})^{i_1i_2...i_k}] = \text{Alt}_J[(X_{j_1j_2...j_k})^{i_1i_2...i_k}] \qquad \text{(A.8.27)}$$

Fact: If $(X_{j_1j_2...j_k})^{i_1i_2...i_k}$ is totally symmetric in both sets of indices, then

$$\text{Sym}_I[(X_{j_1j_2...j_k})^{i_1i_2...i_k}] = \text{Sym}_J[(X_{j_1j_2...j_k})^{i_1i_2...i_k}] \ . \qquad \text{(A.8.28)}$$

Example: According to (7.2.9) the object $(v_{j_1}\wedge v_{j_2}\wedge \ .....\wedge v_{j_k})^{i_1i_2...i_k}$ is totally antisymmetric in both sets of indices. Therefore,

$$\text{Alt}_I\, (v_{j_1}\wedge v_{j_2}\wedge \ .....\wedge v_{j_k})^{i_1i_2...i_k} = \text{Alt}_J\, (v_{j_1}\wedge v_{j_2}\wedge \ .....\wedge v_{j_k})^{i_1i_2...i_k} \ . \qquad \text{(A.8.29)}$$

Another case of interest is when $(X_{j_1j_2...j_k})^{i_1i_2...i_k}$ is an outer product of identical rank-2 tensors,

$$(X_{j_1j_2...j_k})^{i_1i_2...i_k} = T_{j_1}{}^{i_1}\, T_{j_2}{}^{i_2} \ .....\ T_{j_k}{}^{i_k} \ .$$

Then

$$\text{Alt}_I[(X_{j_1j_2...j_k})^{i_1i_2...i_k}] = (1/k!)\ \Sigma_P(-1)^{S(P)}\ T_{j_1}{}^{i_{P(1)}}\, T_{j_2}{}^{i_{P(2)}} \ .....\ T_{j_k}{}^{i_{P(k)}}$$

$$= (1/k!)\ \det(T_{j*}{}^{i*}\ ) \ .$$

According to (A.1.19) we can move the P(..) operators from the i subscripts to the j subscripts to get

$$= (1/k!)\ \Sigma_P(-1)^{S(P)}\ T_{j_{P(1)}}{}^{i_1}\, T_{j_{P(2)}}{}^{i_2} \ .....\ T_{j_{P(k)}}{}^{i_k} = \text{Alt}_J[(X_{j_1j_2...j_k})^{i_1i_2...i_k}] \ .$$

We are just swapping the rows and columns in the determinant shown above. Thus,

Fact: $\text{Alt}_I\,[\ T_{j_1}{}^{i_1}\, T_{j_2}{}^{i_2} \ .....\ T_{j_k}{}^{i_k}\ ] = \text{Alt}_J\,[\ T_{j_1}{}^{i_1}\, T_{j_2}{}^{i_2} \ .....\ T_{j_k}{}^{i_k}\ ] = (1/k!)\ \det(T_{j*}{}^{i*}\ )$. (A.8.30)

Either form gives the same expression $(1/k!)\ [T_{j_1}{}^{i_1}\, T_{j_2}{}^{i_2} \ .....\ T_{j_k}{}^{i_k}$ + all signed permutations].

In multiindex notation, we rewrite this last Fact as

Fact: If $T_J{}^I$ has factored form, then $\text{Alt}_I\,[T_J{}^I] = \text{Alt}_J\,[T_J{}^I] = (1/k!)\ \det(T_J{}^I)$. (A.8.31)

The above is of course true for $T^I{}_J$ or any other index positions. Next,

Fact: If $T_J{}^I$ has factored form, then $T_J{}^I = T_{P(J)}{}^{P(I)}$ where P = any permutation of the index subscripts . (A.8.32)

Proof: Reordering the index subscripts this way just reorders the factors in the product of factors. For example if $T_J{}^I = T_{j_1}{}^{i_1} T_{j_2}{}^{i_2}$ then if P[1,2] = [2,1] one gets $T_{P(J)}{}^{P(I)} = T_{j_2}{}^{i_2} T_{j_1}{}^{i_1} = T_J{}^I$.

Fact : If $T_J{}^I$ has factored form, then, $\Sigma_J \det(T^I{}_J)\, x^J = \Sigma_J\, k!\, T^I{}_J\, x^{\wedge J}$. (A.8.33)

Here $x^J = x^{j1} \otimes x^{j2} \ldots\otimes x^{jk}$, $x^{\wedge J} = x^{j1} \wedge x^{j2} \ldots\wedge x^{jk}$ and each x^j is a vector labeled by j . That is to say, x^j is not the jth component of vector x.

Proof: LHS = $\Sigma_J \det(T^I{}_J)\, x^J$

$= \Sigma_J [k!\, Alt_I(T^I{}_J)]\, x^J$ // (A.8.31)

$= \Sigma_J\, k!\, [\, (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, T^{P(I)}{}_J]\, x^J$ // (A.2.1) for Alt_I

$= \Sigma_P\, (-1)^{S(P)}\, [\Sigma_J\, T^{P(I)}{}_J\, x^J]$ // reorder

$= \Sigma_P\, (-1)^{S(P)}\, [\Sigma_J\, T^{P(I)}{}_{P(J)}\, x^{P(J)}]$ // (A.1.24) that $\Sigma_J f_J = \Sigma_J f_{P(J)}$

$= \Sigma_P\, (-1)^{S(P)}\, \Sigma_J T^I{}_J\, x^{P(J)}$ // (A.8.32)

$= k!\, \Sigma_J T^I{}_J\, [\, (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, x^{P(J)}]$ // reorder

$= k!\, \Sigma_J T^I{}_J\, Alt_J(x^J)$ // (A.2.1) for Alt_J

$= k!\, \Sigma_J T^I{}_J\, x^{\wedge J}$ = RHS . // (7.4.3) ∎

Fact : $\Sigma_I\, T^I\, u_{\wedge I} = \Sigma'_I\, k!\, Alt_I(T^I)\, u_{\wedge I}$ for any tensor T^I (A.8.34)
$\Sigma_I\, T_I\, u^{\wedge I} = \Sigma'_I\, k!\, Alt_I(T_I)\, u^{\wedge I}$ for any tensor T_I .

This theorem (first line) says that if the symmetric sum Σ_I is replaced by the ordered sum Σ'_I , then the coefficients T^I get replaced by $k!\, Alt_I(T^I)$.

Example: $T^{i_1 i_2}$ gets replaced by $2!\, Alt_I(T^{i_1 i_2}) = 2!\, (\, (1/2!)[T^{i_1 i_2} - T^{i_2 i_1}]\,) = [T^{i_1 i_2} - T^{i_2 i_1}]$. Then

$$\Sigma_I\, T^I\, u_{\wedge I} = \Sigma_{i_1 i_2}\, T^{i_1 i_2}\, u_{i_1} \wedge u_{i_2} = \Sigma_{i_1 < i_2}\, [T^{i_1 i_2} - T^{i_2 i_1}]\, u_{i_1} \wedge u_{i_2}\, .$$

Proof: This theorem (first line) was proved in (7.4.4) through (7.4.16). Here we just review that proof using multiindex notation:

$T^\wedge = \Sigma_I \ T^I u^\wedge_I \qquad\qquad u^\wedge_I = u_{i_1}{}^\wedge u_{i_2}{}^\wedge \ldots\ldots{}^\wedge u_{i_k}$ *(7.4.4)*

$T^\wedge = \ \Sigma'_I \ A^I u^\wedge_I \ .$ *(7.4.7)*

$T^\wedge \ = \Sigma_{i_1 \neq i_2 \neq \ldots \neq i_k} \ T^I u^\wedge_I$ *(7.4.9)*

$T^\wedge \ = \Sigma_P \ \ \Sigma_{i_{P(1)} < i_{P(2)} < \ldots < i_{P(k)}} \ T^I u^\wedge_I$ *(7.4.12)*

$T^\wedge \ = \Sigma_{i_1 < i_2 < \ldots < i_k} \ \ \Sigma_P \, T^{P(I)} \ u^\wedge_{P(I)}$ // using (A.9.1) below *(7.4.13)*

$u^\wedge_{P(I)} = \ (-1)^{S(P)} u^\wedge_I$ *(7.4.14)*

$T^\wedge = \ \ \Sigma_{i_1 < i_2 < \ldots < i_k} [\Sigma_P \, (-1)^{S(P)} \ \ T^{P(I)}] \, u^\wedge_I \ = \ \Sigma'_I \ \ [k! \ Alt(T)^I] \, u^\wedge_I$ *(7.4.15)*

$A^I = k! \ [Alt(T)]^I$ by comparing *(7.4.7)* and *(7.4.15)* *(7.4.16)*

The theorem goes through with the I-tilt the other way, which is the second line of (A.8.34). Alternatively, we can take the first result and apply the "tilt reversal rule" (2.9.1) to get the second line from the first line. █

Fact : $\Sigma_I \ T^I x^\wedge_I = \Sigma'_I \ k! \ Alt_I(T^I) \ x^\wedge_I$ for any tensor T^I (A.8.35)
$\Sigma_I \ T_I \ x^{\wedge I} = \Sigma'_I \ k! \ Alt_I(T_I) \ x^{\wedge I}$ for any tensor T_I

Proof: In the proof of the previous Fact, the basis vectors $\mathbf{u_i}$ played a placeholder role and the same proof works with any set of vectors $\mathbf{x_i}$ where $x^\wedge_I$ is the wedge product of those vectors. █

Fact : $\Sigma_I \ T^I{}_J \, x^\wedge_I = \Sigma'_I \ k! \ Alt_I(T^I{}_J) \ x^\wedge_I$ for any tensor $T^I{}_J$ (A.8.36)
$\Sigma_I \ T^J{}_I \ x^{\wedge I} = \Sigma'_I \ k! \ Alt_I(T^J{}_I) \ x^{\wedge I}$ for any tensor $T^J{}_I$

Proof: The first line is the first line of the previous Fact where a bystander J multiindex has been added. The same idea for the second line.

Fact : If $T^I{}_J$ has factored form, then $\Sigma_I \ T^I{}_J \, x^\wedge_I = \ \Sigma'_I \ det(T^I{}_J) \ x^\wedge_I$. (A.8.37)
If $T_J{}^I$ has factored form, then $\Sigma_I \ T^J{}_I \ x^{\wedge I} = \ \Sigma'_I \ det(T^J{}_I) \ x^{\wedge I}$,

Proof: $\Sigma_I \ T^I{}_J \, x^\wedge_I = \Sigma'_I \ k! \ Alt_I(T^I{}_J) \ x^\wedge_I$ // (A.8.36)
$= \ \ \Sigma'_I \ det(T^I{}_J) \ x^\wedge_I$ // (A.8.31)

$\Sigma_I \ T^J{}_I \ x^{\wedge I} = \Sigma'_I \ k! \ Alt_I(T^J{}_I) \ x^{\wedge I}$ // (A.8.36)
$= \ \ \Sigma'_I \ det(T^J{}_I) \ x^{\wedge I}$ // (A.8.31) █

A.9 The Ordered Sum Theorem

The ordered sum theorem states that,

$$(\Sigma_P [\Sigma_{P(i_1)<P(i_2)<\ldots<P(i_k)}])\, f_{i_1 i_2 \ldots i_k} \;=\; \Sigma_{i_1<i_2<\ldots<i_k} [\Sigma_P\, f_{P(i_1)P(i_2)\ldots P(i_k)}] . \qquad (A.9.1)$$

Rather than present a formal proof, we look at the two simplest cases and the general case is then obvious.

k = 2: First, consider

$$Q \equiv \Sigma_{i_1 \neq i_2} f_{i_1 i_2} = [\Sigma_{i_1<i_2} + \Sigma_{i_1>i_2}]\, f_{i_1 i_2}$$

$$= [\Sigma_{i_1<i_2} + \Sigma_{i_2<i_1}]\, f_{i_1 i_2} = (\Sigma_P[\Sigma_{P(i_1)<P(i_2)}])\, f_{i_1 i_2} . \qquad (A.9.2)$$

On the other hand,

$$Q = [\Sigma_{i_1<i_2} + \Sigma_{i_2<i_1}]\, f_{i_1 i_2} = \Sigma_{i_1<i_2} f_{i_1 i_2} + \Sigma_{i_2<i_1} f_{i_1 i_2}$$

$$= \Sigma_{i_1<i_2} f_{i_1 i_2} + \Sigma_{i_1<i_2} f_{i_2 i_1} \qquad \text{//dummy swap } i_1 \leftrightarrow i_2 \text{ in 2nd term}$$

$$= \Sigma_{i_1<i_2} [\, f_{i_1 i_2} + f_{i_2 i_1}] = \Sigma_{i_1<i_2} [\Sigma_P\, f_{P(i_1)P(i_2)}] . \qquad (A.9.3)$$

Thus we have proven the Theorem for k = 2,

$$(\Sigma_P [\Sigma_{P(i_1)P(i_2)}])\, f_{i_1 i_2} = \Sigma_{i_1<i_2} [\Sigma_P\, f_{P(i_1)P(i_2)}] \qquad (A.9.4)$$

k = 3: First, consider

$$Q \equiv \Sigma_{i_1 \neq i_2 \neq i_3} f_{i_1 i_2 i_3}$$

$$= (\Sigma_{i_1<i_2<i_3} + \Sigma_{i_1<i_3<i_2} + \Sigma_{i_2<i_1<i_3} + \Sigma_{i_2<i_3<i_1} + \Sigma_{i_3<i_1<i_2} + \Sigma_{i_3<i_2<i_1})\, f_{i_1 i_2 i_3}$$

$$= (\Sigma_P [\Sigma_{P(i_1)<P(i_2)<P(i_3)}])\, f_{i_1 i_2 i_3} . \qquad (A.9.5)$$

On the other hand we can rename the summation indices in all but the first sum to get

$$Q = (\Sigma_{i_1<i_2<i_3} + \Sigma_{i_1<i_3<i_2} + \Sigma_{i_2<i_1<i_3} + \Sigma_{i_2<i_3<i_1} + \Sigma_{i_3<i_1<i_2} + \Sigma_{i_3<i_2<i_1})\, f_{i_1 i_2 i_3}$$

as is $\qquad$ $2\leftrightarrow 3$ $\qquad$ $1\leftrightarrow 3$

$$= \Sigma_{i_1<i_2<i_3} f_{i_1 i_2 i_3} + \Sigma_{i_1<i_2<i_3} f_{i_1 i_3 i_2} + \Sigma_{i_1<i_2<i_3} f_{i_3 i_2 i_1} + \text{3 more sums}$$

$$= \Sigma_{i_1<i_2<i_3} [\; f_{i_1 i_2 i_3} + f_{i_1 i_3 i_2} + f_{i_3 i_2 i_1} + \text{3 more terms}\;]$$

$$= \quad \Sigma_{i_1<i_2<i_3} \; [\Sigma_P f_{P(i_1)P(i_2)P(i_3)}]. \qquad \text{(A.9.6)}$$

Thus we have proven the Theorem for k =3,

$$(\Sigma_P [\Sigma_{P(i_1)<P(i_2)<P(i_3)}]) \; f_{i_1 i_2 i_3} \; = \Sigma_{i_1<i_2<i_3} \; [\Sigma_P f_{P(i_1)P(i_2)P(i_3)}] \; . \qquad \text{(A.9.7)}$$

The argument for a k-fold sum proceeds in the same manner, and we end up with

$$(\Sigma_P \; [\Sigma_{P(i_1)<P(i_2)<...<P(i_k)}]) \, f_{i_1 i_2 ... i_k} \quad = \; \Sigma_{i_1<i_2<...<i_k} [\Sigma_P \; f_{P(i_1)P(i_2)...P(i_k)}] \, . \qquad \textit{(A.9.1)}$$

A.10 Tensor Products in Generic Notation

In Section A.2 above we use a set of generic function arguments (1,2..k) and their permutations to define the Alt and Sym operators, detached from the world of tensors and tensor functions. In this same generic vein one can define a generic tensor product as follows,

Definition: Tensor product: $(f\otimes g)(1,2,....k+k') \equiv f(1,2...k) \;\; g(k+1,k+2....k+k')$. (A.10.1)

If we translate this definition in the same way we translated everything else, we arrive at this corresponding statement in the tensor world,

Definition: Tensor product: $(T\otimes S)^{i_1 i_2 ... i_{k+k'}} \equiv T^{i_1 i_2 ... i_k} \; S^{i_{k+1} i_{k+2} ... i_{k+k'}}$,
where the ranks of tensors T,S are k,k'. (A.10.2)

Since (A.10.2) is already defined to be true using the "outer product definition" of the ⊗ symbol in Section 2.8, it seems that here we are just lucky to obtain a consistent result. We put this issue on hold for a moment, and consider next the way (A.10.1) translates into the tensor function world,

Definition: Tensor product: $(\mathcal{T}\otimes\mathcal{S})(v_{i_1},v_{i_2}....v_{i_{k+k'}}) \equiv \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) \; \mathcal{S}(v_{i_{k+1}},v_{i_{k+2}}....v_{i_{k+k'}})$,
where the ranks of dual tensors $\mathcal{T}$,$\mathcal{S}$ are k,k'. (A.10.3)

In Section 6.6 it seemed that we had to do quite a bit of work to obtain (A.10.3) as stated in (6.6.13), and here suddenly this same result just drops out as some kind of definition.

Both these issues can be clarified by use of the Dirac notation which reveals the underlying "tensor product of vector spaces" structure. First, we can write (A.10.1) in the generic world as

$$(f\otimes g)(1,2,....k+k') \equiv f(1,2...k) \;\; g(k+1,k+2....k+k') \qquad \textit{(A.10.1)}$$

$${}_{k+k'}\langle f\otimes g \,|\, 1,2,.....k+k'\rangle_{k+k'} = {}_k\langle f \,|\, 1,2,.....k\rangle_k \; * \; {}_{k'}\langle g \,|\, k+1,k+2.....k+k'\rangle_{k'} \qquad \text{(A.10.4)}$$

where for example

$$| \ 1,2,.....k>_{\mathbf{k}} \ = |1>_{\mathbf{1}} \otimes |2>_{\mathbf{1}} \otimes ...\otimes |k>_{\mathbf{1}} \tag{A.10.5}$$

$$_{\mathbf{k+k'}}<f\otimes g| \ = {}_{\mathbf{k}}<f| \otimes {}_{\mathbf{k'}}<g| \ . \tag{A.10.6}$$

The subscript k labels a ket in $V^{\mathbf{k}}$ and a bra in $V^{*\mathbf{k}}$, for example. Suddenly we are interpreting the generic argument set (1,2,...k) as if it were a tensor product of "generic kets" in their own vector spaces. By itself, this does not really make much sense, but when we think of the generic description as being a stand-in for our tensor and tensor-function cases, then it does make sense. We show in Appendix D (D.1.2) and (D.1.3), and also in the main text (2.11.e.7) and (2.11.e.9), that

$$\mathcal{T}(v_{i_1},v_{i_2}, ... v_{i_k}) \ = <T \mid v_{i_1},v_{i_2}, ... v_{i_k} >$$
$$\text{where} \mid v_{i_1},v_{i_2}, ... v_{i_k} > \ = |v_{i_1}> \otimes \ |v_{i_2}> \ \otimes \ ..... \otimes \ |v_{i_k}> \tag{A.10.7}$$

$$T_{i_1 i_2 ....i_k} \ = \ <T \mid u_{i_1},u_{i_2}, ... u_{i_k} >$$
$$\text{where} \ \mid u_{i_1},u_{i_2}, ... u_{i_k} > \ = |u_{i_1}> \otimes \ |u_{i_2}> \ \otimes \ ..... \otimes \ |u_{i_k}> \tag{A.10.8}$$

and here we see the real meanings of the generic stand-in tensor product $|1> \otimes |2> \otimes ...\otimes |k>$.

For the tensor case we then write

$$(T\otimes S)^{i_1 i_2 ... i_{k+k'}} \ = \ < T\otimes S| \, u_{i_1},u_{i_2}, ... u_{i_{k+k'}} >$$

$$= [\ _{\mathbf{k}}<T| \otimes {}_{\mathbf{k'}}<S| \ ] \ [\ | \ u_{i_1},u_{i_2}, ... u_{i_k} >_{\mathbf{k}} \otimes \ | \ u_{i_{k+1}},u_{i_{k+2}}, ... u_{i_{k+k'}} >_{\mathbf{k'}} \ ]$$

$$= {}_{\mathbf{k}}< T \mid \ u_{i_1},u_{i_2}, ... u_{i_k} >_{\mathbf{k}} \ {}_{\mathbf{k'}}< S \mid \ u_{i_{k+1}},u_{i_{k+2}}, ... u_{i_{k+k'}} >_{\mathbf{k'}}$$

$$= T^{i_1 i_2 ... i_k} \ S^{i_{k+1} i_{k+2} ... i_{k+k'}} \tag{A.10.9}$$

and the Dirac tensor product space structure then directly implies the tensor "outer product" rule defined in Chapter 2.

In the tensor-function case we do exactly the same thing but with u→v,

$$(\mathcal{T}\otimes\mathcal{S})(v_{i_1},v_{i_2}....v_{i_{k+k'}}) \ = < T\otimes S \mid v_{i_1},v_{i_2}....v_{i_{k+k'}}>$$

$$= [\ _{\mathbf{k}}<T| \otimes {}_{\mathbf{k'}}<S| \ ] \ [\ | \ v_{i_1},v_{i_2}, ... v_{i_k} >_{\mathbf{k}} \otimes \ | \ v_{i_{k+1}},v_{i_{k+2}}, ... v_{i_{k+k'}} >_{\mathbf{k'}} \ ]$$

$$= {}_{\mathbf{k}}< T \mid \ v_{i_1},v_{i_2}, ... v_{i_k} >_{\mathbf{k}} \ {}_{\mathbf{k'}}< S \mid \ v_{i_{k+1}},v_{i_{k+2}}, ... v_{i_{k+k'}} >_{\mathbf{k'}}$$

$$= \mathcal{T}(v_{i_1},v_{i_2}....v_{i_k}) \, \mathcal{S}(v_{i_{k+1}},v_{i_{k+2}}....v_{i_{k+k'}}) \tag{A.10.10}$$

and again the vector space structure forces this tensor function result. In fact, this is exactly how we derived this result in (6.6.12) for a more general case.

Appendix B: Direct Sum of Vector Spaces

There are nine short numbered sections below. Here are the headings:

1. Axioms for ⊕
2. Direct Sum Space V⊕W
3. Basis for V⊕W
4. Z = V⊕W is a vector space
5. v⊕w does not commute
6. Visualization of the Direct Sum
7. Extension to multiple ⊕ products
8. Application: adding tensor products
9. Direct sum of operators (matrices): Block Diagonal Form

1. Axioms for ⊕

Let $v_i \in V$ and $w_i \in W$ where V and W are vector spaces, and $\alpha \in K$ is a scalar. The direct sum operator ⊕ can be defined by these rules (axioms) for any $k > 1$,

$$v_1\oplus w_1 + v_2\oplus w_2 + \ldots + v_k\oplus w_k = (v_1+v_2 + \ldots + v_k) \oplus (w_1+w_2 + \ldots + w_k) \qquad (B.1)$$

$$(\alpha v)\oplus(\alpha w) = \alpha(v\oplus w) . \qquad (B.2)$$

In slightly more concise notation (B.1) can be written $\Sigma_{i=1}{}^k (v_i\oplus w_i) = (\Sigma_{i=1}{}^k v_i) \oplus (\Sigma_{i=1}{}^k w_i)$.

For k = 2 (B.1) becomes,

$$v_1\oplus w_1 + v_2\oplus w_2 = (v_1+v_2) \oplus (w_1+w_2) . \qquad (B.3)$$

Since V and W are vector spaces, each has a 0 element and one can write

$$v_1\oplus 0 + 0\oplus w_2 = (0+v_1) \oplus (0+w_2) = v_1\oplus w_2 \qquad \text{// (B.3) with } w_1 = 0 \text{ and } v_2 = 0 \qquad (B.4)$$

$$(\alpha v)\oplus 0 = \alpha(v\oplus 0) \qquad \text{// } w = 0$$
$$0\oplus(\alpha w) = \alpha(0\oplus w) \qquad \text{// } v = 0 \qquad (B.5)$$

2. Direct Sum Space V⊕W

Define space Z by

$$Z \equiv V\oplus W \qquad (B.6)$$

and let

$$z_i \equiv v_i\oplus w_i \in Z. \qquad (B.7)$$

One might write

$\oplus : (V,W) \rightarrow V\oplus W \qquad \oplus : (v,w) \mapsto (v\oplus w)$

Lemma: Given some z_i, we can find v_i and w_i such that $z_i = v_i\oplus w_i$. (B.8)

Proof: When we say $Z \equiv V\oplus W$, we mean these spaces are *the same*, so there is a 1-to-1 correspondence between elements of $z_i \in Z$ and elements $v_i\oplus w_i \in V\oplus W$.

3. Basis for V⊕W

Let us assume that:
e_i form a basis of dimension n for V
e'_j form a basis of dimension n' for W

Fact: A basis for Z can be written as

$$\{e_1\oplus 0,\ e_2\oplus 0 \text{ ....} e_n\oplus 0,\ 0\oplus e'_1,\ 0\oplus e'_2, \text{ ....} 0\oplus e'_{n'}\} \tag{B.9}$$

Proof: Let v^i be components of vector v, and w^i the components of vector w. Consider:

$$\{ v^1(e_1\oplus 0) + v^2(e_2\oplus 0) + ... + v^n(e_n\oplus 0)\} + \{w^1(0\oplus e'_1) + w^2(0\oplus e'_2) + ... + w^{n'}(0\oplus e'_{n'})\}$$

$$=\{ (v^1e_1)\oplus 0 + (v^2e_2)\oplus 0 + ...+ (v^ne_n)\oplus 0\} + \{0\oplus(w^1e'_1) + 0\oplus(w^2e'_2) + ... + 0\oplus(w^{n'}e'_{n'})\} \quad \text{// } \textit{(B.5)}$$

$$= \{ (v^1e_1+ v^2e_2 + ...+ v^ne_n)\oplus(0+0+..+0)\} + \{(0+0+..+0)\oplus (w^1e'_1+ w^2e'_2 + ...+ w^ne'_n)\} \quad \text{// } \textit{(B.1)}$$

$$= (v^1e_1+ v^2e_2 + ...+ v^ne_n) \oplus 0 + 0 \oplus (w^1e'_1+ w^2e'_2 + ...+ w^ne'_n)$$

$$= (v^1e_1+ v^2e_2 + ...+ v^ne_n) \oplus (w^1e'_1+ w^2e'_2 + ...+ w^ne'_n) \quad \text{// } \textit{(B.4)}$$

$$= v \oplus w$$

This $z = v \oplus w$ is an arbitrary element of Z, and we have therefore shown that an arbitrary element of Z can be expanded on the basis shown in (B.9) and that no smaller basis will do the job. QED

Fact: If dim(V) = n and dim(W) = n'. then dim(V⊕W) = n + n' (B.10)

Proof: Count the basis elements shown in (B.9).

Compare this Fact with that shown in (4.1.1) :

Fact: If dim(V) = n and dim(W) = n', then dim(V⊗W) = n * n' . *(4.1.1)*

4. $Z = V\oplus W$ is a vector space

Fact: If V and W are vector spaces, then $Z = V\oplus W$ is a vector space. (B.11)

Proof: We just run down the required axioms listed for example on the wiki vector space page. The conclusion one reaches is that the vector space properties are "induced" from V and W into Z.

• The fact that + is commutative within V and W causes + to be commutative within Z :

$$z_1 + z_2 = v_1\oplus w_1 + v_2\oplus w_2 = (v_1+v_2) \oplus (w_1+w_2) = (v_2+v_1) \oplus (w_2+w_1) = v_2\oplus w_2 + v_1\oplus w_1$$

$$= z_2 + z_1 .$$

• Addition in Z is associative because it is associative in V and W:

$$(z_1+ z_2) + z_3 = (v_1\oplus w_1 + v_2\oplus w_2) + v_3\oplus w_3 = (v_1+v_2)\oplus(w_1+w_2) + v_3\oplus w_3$$

$$= (v_1+v_2+v_3)\oplus (w_1+w_2+w_3) = v_1\oplus w_1 + (v_2+v_3)\oplus(w_2+w_3)$$

$$= v_1\oplus w_1 + (v_2\oplus w_2 + v_3\oplus w_3) = z_1 + (z_2 + z_3) .$$

• The zero element in Z is $0 = 0\oplus 0$ since

$$v\otimes w + 0 = v\otimes w + 0\oplus 0 = (v+0)\otimes(w+0) = v\otimes w .$$

• The additive inverse of $z = v\oplus w$ is $-z = (-v)\oplus(-w)$ since

$$z + (-z) = v\oplus w + (-v)\oplus(-w) = (v-v)\oplus(w-w) = 0\oplus 0 = 0 .$$

• For scalars a,b we have $a(bz) = (ab)z$ compatibility since

$$a(bz) = a(b[v\oplus w]) = a[(bv)\oplus(bw)] = (abv)\oplus(abw) = (ab)(v\otimes w) = (ab)z .$$

• Identity for scalar multiplication requires that $1(z) = z$:

$$1(z) = 1(v\oplus w) = (1v)\oplus(1w) = v\oplus w = z .$$

• Distributive requirement #1: $a(z_1+z_2) = az_1+ az_2$ (a = scalar)

$$a(z_1+z_2) = az_3 = a(v_3\oplus w_3) = (av_3)\oplus(aw_3) = (av_1+av_2)\oplus(aw_1+aw_2)$$

$$= (av_1)\oplus(aw_1) + (av_2)\oplus(aw_2) = a(v_1\oplus w_1) + a(v_2\oplus w_2) = az_1+ az_2$$

• Distributive requirement #2 : $(a+b)z = az + bz$ (a,b = scalars)

$$(a+b)z = (a+b)(v\oplus w) = [(a+b)v]\oplus[(a+b)w] = [av+bv]\oplus[aw+bw] = (av)\oplus(aw) + (bv)\oplus(bw)$$

$$= a(v\oplus w) + b(v\oplus w) = az + bz$$ QED

5. $v\oplus w$ does not commute

Fact: $v\oplus w \neq w\oplus v$ unless $V = W$ and $v = w$. (B.12)

Proof:

$V\neq W$: If $V\neq W$, the object $w\oplus v$ makes no sense since it would require $w \in V$ and $v \in W$.

$V=W$: $v\oplus w - w\oplus v = v\oplus w + (-w)\oplus(-v) = (v-w)\oplus(w-v) \neq 0$ unless $v = w$.

Compare (B.12) to the Fact stated in and below (4.1.1),

Fact: $v\otimes w \neq w\otimes v$ unless $V = W$ and $v = w$. *(4.1.1)*

However, there is certainly an isomorphism between $V\oplus W$ and $W\oplus V$. Writing $V\oplus W \sim W\oplus V$ one could certainly then say that $v\oplus w \sim w\oplus v$. The same could be said for the $\otimes$ operator.

6. Visualization of the Direct Sum

Consider this example

$$v = \begin{pmatrix} a \\ b \end{pmatrix} \in R^2 \qquad w = \begin{pmatrix} s \\ t \\ u \end{pmatrix} \in R^3 \qquad z = v\oplus w = \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} a \\ b \\ s \\ t \\ u \end{pmatrix} \in R^5 . \qquad (B.13)$$

Here we visualize the direct sum vector z as a tall column vector which is the stacking of the two smaller column vectors v and w. In the tall column vector, v and w each occupy a private region.

Here then are the rules (B.3) and (B.2) :

$$v_1\oplus w_1 + v_2\oplus w_2 = \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} + \begin{pmatrix} v_2 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1+v_2 \\ w_1+w_2 \end{pmatrix} = (v_1+v_2) \oplus (w_1+w_2) \qquad (B.14)$$

$$(\alpha v_1)\oplus(\alpha v_2) = \begin{pmatrix} \alpha v_1 \\ \alpha w_1 \end{pmatrix} = \alpha \begin{pmatrix} v_1 \\ w_1 \end{pmatrix} = \alpha(v_1\oplus v_2) . \qquad (B.15)$$

The fact (B.10) that $\dim(V\oplus W) = \dim(V) + \dim(W)$ is demonstrated by $5 = 2+3$.

The fact (B.12) that $v\oplus w \neq w\oplus v$ is demonstrated since $(a,b,s,t,u)^T \neq (s,t,u,a,b)^T$.

7. Extension to multiple ⊕ products

The axioms for a triple direct sum are these,

$$v_1 \oplus w_1 \oplus x_1 + v_2 \oplus w_2 \oplus x_2 + ... + v_k \oplus w_k \oplus x_k$$

$$= (v_1+v_2 + ...+ v_k) \oplus (w_1+w_2 + ... + w_k) \oplus (x_1+x_2 + ... + x_k) \quad \text{(B.16)}$$

$$(\alpha v)\oplus(\alpha w)\oplus(\alpha x) = \alpha(v\oplus w\oplus x) \quad \text{(B.17)}$$

and from this one can imagine an arbitrary number of ⊕ involved in a direct sum. One can derive these two equations from (B.1) and (B.2) by assuming associativity and then grouping things for example as

$$v_1 \oplus w_1 \oplus x_1 + v_2 \oplus w_2 \oplus x_2 + ... + v_k \oplus w_k \oplus x_k$$

$$= (v_1 \oplus w_1)\oplus x_1 + (v_2 \oplus w_2)\oplus x_2 + ... + (v_k \oplus w_k)\oplus x_k$$

$$= [(v_1 \oplus w_1) + (v_2 \oplus w_2) + ... + (v_k \oplus w_k)] \oplus (x_1 + x_2 + ... + x_k)$$

$$= [(v_1+v_2+ .. +v_k) \oplus (w_1+w_2+ .. +w_k)] \oplus (x_1 + x_2 + ... + x_k)$$

$$= (v_1+v_2 + ...+ v_k) \oplus (w_1+w_2 + ... + w_k) \oplus (x_1+x_2 + ... + x_k)$$

and

$$(\alpha v)\oplus(\alpha w)\oplus(\alpha x) =[(\alpha v)\oplus(\alpha w)] \oplus (\alpha x) = [\alpha(v\oplus w)] \oplus (\alpha x) = \alpha [(v\oplus w \oplus x] = \alpha[v\oplus w\oplus x] .$$

One can define

$$Z \equiv V\oplus W\oplus X \qquad z_i \equiv v_i \oplus w_i \oplus x_1 \in Z \quad \text{(B.6)'}$$

$$\oplus : (V,W,X) \rightarrow V\oplus W\oplus X \qquad \oplus : (v,w,x) \mapsto (v\oplus w\oplus x)$$

We leave it to the reader to prove the following extended claims:

Lemma: Given some z_i, we can find v_i, w_i and x_i such that $z_i = v_i \oplus w_i \oplus x_i$. (B.8)'

Fact: If dim(V) = n, dim(W) = n' and dim(X) = n". then dim(V⊕W⊕X) = n + n' + n". (B.10)'

Fact: If V.W and X are vector spaces, then Z = V⊕W⊕X is a vector space. (B.11)'

The extension of the "tall vector" visualization to the triple ⊕ sum seems fairly obvious where one ends up stacking three vectors to make a single tall vector.

8. Application: adding tensor products

Define the vector product space V^k as in (5.1).

• If $T = V^2 \oplus V^3$ one can write

$$t = \Sigma_{ij} T^{ij} u_i \otimes u_j \ \oplus\ \Sigma_{ijk} T^{ijk} u_i \otimes u_j \otimes u_k \ \in T \qquad // \ t = v \oplus w$$

• If $T = V^1 \oplus V^2 \oplus V^3$ one can write

$$t = \Sigma_i T^i u_i \ \oplus\ \Sigma_{ij} T^{ij} u_i \otimes u_j \ \oplus\ \Sigma_{ijk} T^{ijk} u_i \otimes u_j \otimes u_k \ \in T \qquad // \ t = v \oplus w \oplus x$$

and in this manner we eventually arrive at (5.4.1) for T(V)

$$T(V) \equiv V^0 \oplus V \oplus V^2 \oplus V^3 \oplus .... \qquad (5.4.1)$$

$$t = s \oplus \Sigma_i T^i u_i \oplus \Sigma_{ij} T^{ij} u_i \otimes u_j \oplus \Sigma_{ijk} T^{ijk} u_i \otimes u_j \otimes u_k \oplus ... \quad \in T(V), \quad s \in K \qquad (5.4.2)$$

For the space V^{*k} the objects being direct-summed are functionals instead of tensors, but the formalism is exactly the same,

$$T(V^*) \equiv V^{*0} \oplus V^* \oplus V^{*2} \oplus V^{*3} \oplus .... = \Sigma^{\oplus}{}_{k=1}{}^{\infty} V^{*k} . \qquad (6.4.1)$$

$$\tau = s \oplus \Sigma_i T_i \lambda^i \oplus \Sigma_{ij} T_{ij} \lambda^i \otimes \lambda^j \oplus \Sigma_{ijk} T_{ijk} \lambda^i \otimes \lambda^j \otimes \lambda^k + ... \qquad s \in K \qquad (6.4.2)$$

9. Direct sum of operators (matrices): Block Diagonal Form

Let vector spaces V and W have dimension n and n'.
Let S, S_1 and S_2 be n x n matrices which we can regard as linear operators in V.
Let T, T_1 and T_2 be n' x n' matrices which we can regard as linear operators in W.

Then in the direct product space $V \oplus W$ we can write these operator equations,

$$(S_1+S_2) \oplus (T_1+T_2) = S_1 \oplus T_1 + S_2 \oplus T_2$$

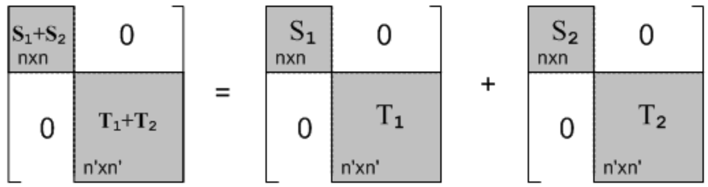
$$\alpha(S \oplus T) = \alpha S \oplus \alpha T$$

$$(S_1 \oplus T_1)(S_2 \oplus T_2) = (S_1 S_2) \oplus (T_1 T_2) \qquad (B.18)$$

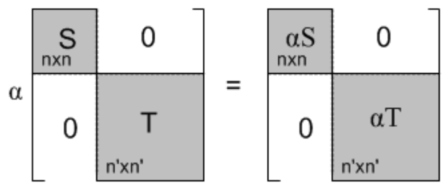
and the action of operator $S \oplus T$ of $V \otimes W$ on a vector of $V \otimes W$ is given by

$$(S \oplus T)(v \oplus w) = (Sv) \oplus (Tw) . \qquad (B.19)$$

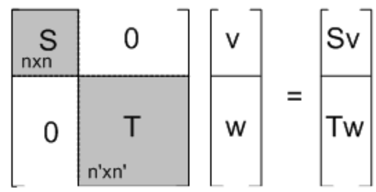
Just as we visualized the direct sum of two vectors in (B.13), it is helpful to visualize the above three matrix equations graphically:



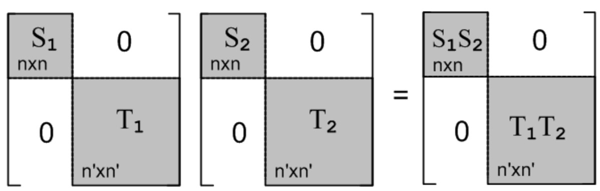
$(S_1 + S_2)\oplus(T_1 + T_2) = S_1\oplus T_1 \ + \ S_2\oplus T_2$



$\alpha(S \oplus T) \ = \alpha S \oplus \alpha T$

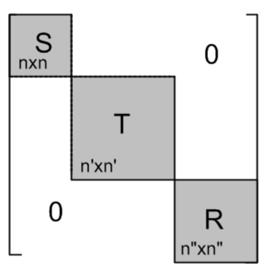


$(S \oplus T)\,(v\oplus w) = (Sv)\oplus(Tw)$



$(S_1\oplus T_1)(S_2\oplus T_2) \ = (S_1S_2)\oplus(T_1T_2)$ (B.20)

The direct sum operator S⊕T is represented as an (n+n')x(n+n') matrix that is in "block diagonal form" where the matrix outside the blocks is filled with zeros. A triple direct sum has this visualization,



$= S \oplus T \oplus R$ (B.21)

In all cases, the *entire* area outside the diagonal blocks is set to 0. The rules for such operators are,

$(S_1 + S_2)\oplus(T_1 + T_2)\oplus(R_1 + R_2) = S_1\oplus T_1\oplus R_1 \ + \ \ S_2\oplus T_2\oplus R_2$

$\alpha(S\oplus T\oplus R) \ = \alpha S \oplus \alpha T \oplus \alpha R$

$(S_1\oplus T_1\oplus R_1)(S_2\oplus T_2\oplus R_2) \ = (S_1S_2)\oplus(T_1T_2) \oplus(R_1R_2)$

$(S\oplus T\oplus R)\ (v\oplus w\oplus x) = (Sv)\oplus(Tw)\oplus(Rx)$. (B.22)

Appendix C: Theorems on Pre-Symmetrization

The Rearrangement Theorem (A.1.3) is used to prove three other theorems (One, Two and Three) where we have attempted to abstract as much as possible the "permutational nature" of the objects involved by using a generic permutation space with elements $|1,2...k\rangle$. Then in Section C.4 the theorems are summarized and are generalized to apply to arbitrary tensor products. Finally, the generic theorems are applied to tensors and tensor functions. The reader uninterested in the theorem details would do well to skip right to the summary presented in Section C.4.

It is assumed that the reader is familiar with Appendix A.1-3 and A.10.

C.1 Theorem One

Consider the following list of k+k' integers,

$$[1,2....k, k+1,k+2....k+k'] = [1,2....k+k'] . \tag{C.1.1}$$

Partition this list into a low and high group by defining

$$z \equiv [1,2....k] \qquad Z = [k+1,k+2....k+k'] . \tag{C.1.2}$$

Then

$$[1,2....k+k'] = [z,Z] . \tag{C.1.3}$$

Now let Q be a permutation of the lower integers [1,2...k] = z. There are k! possible permutations, so we know that

$$\Sigma_Q (1) = k! . \tag{C.1.4}$$

We can extend the meaning of Q so it applies to the entire list of integers [1,2....k+k'] merely by stating that this extended Q' does not alter the higher integers. Then

$$Q'[z] = Q[z] = z' = \text{some permutation of the lower integers} \tag{C.1.5a}$$

$$Q'[Z] = Z \quad \text{// since Q has no effect on the higher integers} \tag{C.1.5b}$$

$$Q'[z, Z] = [Q'(z), Q'(Z)] = [Q(z), Z] . \tag{C.1.5c}$$

Now imagine we have a function f of the lower integers and a function F of the higher ones,

$$f[z] = f[1,2....k] \qquad F[Z] = F[k+1,k+2....k+k'] . \tag{C.1.6}$$

Here are two applications we shall consider later on,

$f[z] = f[1,2....k] = T^{i_1 i_2 ... i_k}$ = components of a rank-k tensor

$f[z] = f[1,2....k] = \mathcal{T}(v_{i_1}, v_{i_2}, ... v_{i_k})$ = a rank-k tensor function . (C.1.7)

Now let P be a *general* permutation of [1,2....k+k'] = [z,Z].

$$P[z,Z] = [P(z), P(Z)] . \qquad (C.1.8)$$

Notice that QP and PQ are undefined since P and Q operate in different spaces, but Q'P and PQ' are both defined since both permutations Q' and P operate in the space of [1,2....k+k'].

Recall now the meaning of S(Q) as the number of swaps required to go from z to Q(z) . This is the same as the number of swaps required to go from [z,Z] to Q'[z,Z] = [Q(z),Z]. Therefore

$$S(Q) = S(Q') . \qquad (C.1.9)$$

We shall now prove the following theorem :

Theorem One

$$\Sigma_P (-1)^{\sigma S(P)} f[P(z)] F[P(Z)] = \Sigma_P (-1)^{\sigma S(P)} f_\wedge[P(z)] F[P(Z)] \qquad (C.1.10)$$

where

$$f_\wedge[z] \equiv (1/k!) \Sigma_Q (-1)^{\sigma S(Q)} f[Q(z)] \qquad \sigma = 1 \text{ or } 0$$

The purpose of σ is to state the theorem with and without the $(-1)^{S(P)}$ factor.

We shall use the notation $f_\wedge$ in most of this section to apply for both values of σ, but at the end, we shall distinguish these two cases by writing:

$$f_\wedge[z] \equiv (1/k!) \Sigma_Q (-1)^{S(Q)} f[Q(z)] \qquad // = \text{Alt}(f), \text{ see (A.2.1)}$$
$$f_s[z] \equiv (1/k!) \Sigma_Q f[Q(z)] \quad . \qquad // = \text{Sym}(f), \text{ see (A.3.1)} \qquad (C.1.11)$$

At the end of this section we will show that the above Theorem One with σ = 1 and σ = 0 is equivalent to the statements:

$$\text{Alt}(f\otimes F) = \text{Alt}(f_\wedge \otimes F) \qquad f_\wedge = \text{Alt}(f) \qquad \sigma = 1$$
$$\text{Sym}(f\otimes F) = \text{Sym}(f_s \otimes F) \qquad f_s = \text{Sym}(f) \qquad \sigma = 0 \quad . \qquad (C.1.12)$$

<u>Proof of Theorem One</u>: Our first task is to process the second line of (C.1.10),

$$f_\wedge[z] = (1/k!) \Sigma_Q (-1)^{\sigma S(Q)} f[Q(z)]$$
$$= (1/k!) \Sigma_Q (-1)^{\sigma S(Q')} f[Q'(z)]. \qquad // \text{(C.1.9) and (C.1.5a)} \qquad (C.1.13)$$

Apply permutation P to the above equation and use (A.2.8) to get

$$P\, f_\wedge[z] = f_\wedge[P(z)] \;=\; (1/k!)\, \Sigma_Q\; (-1)^{\sigma S(Q')}\; f\,[PQ'(z)] \;. \qquad (C.1.14)$$

Then,

$$\text{RHS (C.1.10)} \;=\; \Sigma_P\, (-1)^{\sigma S(P)}\; f_\wedge[P(z)]\, F[P(Z)]$$

$$= \Sigma_P\, (-1)^{\sigma S(P)}\; \{ (1/k!)\, \Sigma_Q\; (-1)^{\sigma S(Q')}\; f[PQ'(z)] \}\, F[P(Z)] \qquad \text{// (C.1.14) for } f_\wedge[P(z)]$$

$$= \; (1/k!)\, \Sigma_Q\; \Sigma_P\; (-1)^{\sigma S(PQ')}\; f[PQ'(z)]\, F[P(Z)] \qquad \text{// reorder and use (A.1.10)}$$

$$= \; (1/k!)\, \Sigma_Q\; \Sigma_P\; (-1)^{\sigma S(PQ')}\; f[PQ'(z)]\, F[PQ'(Z)] \qquad \text{// } Q'(Z) = Z \text{ from (C.1.5b)}$$

$$= \; (1/k!)\, \Sigma_Q\; \Sigma_P (-1)^{\sigma S(P)}\; f[P(z)]\, F[P(Z)] \qquad \text{// rearrangement theorem (A.1.3)}$$

$$= \Sigma_P (-1)^{\sigma S(P)}\; f[P(z)]\, F[P(Z)]\;\; \{ (1/k!)\, \Sigma_Q\, (1) \} \qquad \text{// reorder}$$

$$= \Sigma_P (-1)^{\sigma S(P)}\;\; f[P(z)]\, F[P(Z)]\; \{ 1 \} \qquad \text{// } \Sigma_Q\, (1) = k! \text{ from (C.1.4)}$$

$$= \text{LHS (C.1.10)} \;. \qquad \text{QED} \qquad (C.1.15)$$

Recall now definitions of the generic Alt and Sym operators,

$$[\text{Alt}(f)](1,2...k) \;\equiv\; (1/k!)\, \Sigma_P\, (-1)^{S(P)} f(\, P(1),P(2)...P(k)\,) \qquad \textit{(A.2.1)} \qquad (C.1.16)$$

$$[\text{Sym}(f)](1,2...k) \;\equiv\; (1/k!)\, \Sigma_P\;\; f(\, P(1),P(2)...P(k)\,) \qquad \textit{(A.3.1)} \qquad (C.1.17)$$

Using (C.1.16) and (C.1.17) , the second line of (C.1.10) can be restated

$$f_\wedge \;= \text{Alt}(f) \qquad \sigma = 1 \qquad \text{totally antisymmetric}$$

$$f_s \;= \text{Sym}(f) \qquad \sigma = 0 \qquad \text{totally symmetric} \;\;. \qquad (C.1.18)$$

Recall next the definition of a tensor product ⊗ in our generic function space,

$$(f\otimes g)(1,2,....k+k') \equiv f(1,2...k)\; g(k+1,k+2....k+k') \;\;. \qquad \textit{(A.10.1)} \qquad (C.1.19)$$

Then we can write

$$(f\otimes F)(1,2,....k+k') \;\;= f(1,2..k)F(k+1,k+2...k+k') \;\; = f(z)\; F(Z)$$

$$(f_\wedge\otimes F)(1,2,....k+k') \;= f_\wedge(1,2..k)F(k+1,k+2...k+k') = f_\wedge(z)\; F(Z) \;. \qquad (C.1.20)$$

Theorem One (with σ = 1) can then be stated in this manner,

$$\Sigma_P\, (-1)^{S(P)} (f\otimes F)(P(1),P(2),...P(k+k')) \;\; = \Sigma_P\, (-1)^{S(P)} (f_\wedge\otimes F)(P(1),P(2),...P(k+k')) \qquad \textit{(C.1.10)}_{\sigma=1}$$

Add a factor 1/(k+k')! to both sides and use the Alt definition (C.1.16) with k⟶ k+k' to get,

$[\text{Alt}(f\otimes F)](1,2...k+k') = [\text{Alt}(f_\wedge\otimes F)](1,2...k+k')$

or

$\text{Alt}(f\otimes F) = \text{Alt}(f_\wedge\otimes F) \qquad f_\wedge = \text{Alt}(f)$. (C.1.21)

Taking σ = 0 in (C.1.10) gives ($f_\wedge \rightarrow f_s$ as noted above),

$\Sigma_P (f\otimes F)(P(1),P(2),...P(k+k')) = \Sigma_P (f_s\otimes F)(P(1),P(2),...P(k+k'))$ *(C.1.10)*$_{\sigma=0}$

Use this with the Sym definition (C.1.17) with $k\rightarrow k+k'$ to get

$\text{Sym}(f\otimes F) = \text{Sym}(f_s\otimes F) \qquad f_s = \text{Sym}(f)$. (C.1.22)

C.2 Theorem Two

This section is a copy, paste and edit version of Section C.1. Equations that are the same have italicized equation numbers.

Consider the following set of k+k' integers,

$[1,2....k, k+1,k+2....k+k'] = [1,2....k+k']$. *(C.1.1)*

Partition this list into a low and high half by defining

$z \equiv 1,2....k \qquad Z = k+1,k+2....k+k'$ *(C.1.2)*

Then

$[1,2....k+k'] = [z,Z]$. *(C.1.3)*

Now let R be a permutation of the *upper* integers {k+1,k+2....k+k'} = Z. There are k'! possible permutations, so we know that

$\Sigma_R (1) = k'!$. (C.2.4)

We can extend the meaning of R so it applies to the entire set of integers [1,2....k+k'] merely by stating that this extended R' does not alter the lower integers. Then

$R'[Z] = R[Z] = Z'$ = some permutation of the upper integers (C.2.5a)

$R'[z] = z$ // since R has no effect on the lower integers (C.2.5b)

$R'[z, Z] = [R'(z), R'(Z)] = [z, R(Z)]$. (C.2.5c)

Now imagine we have a function f of the lower integers and a function F of the higher ones,

$f[z] = f[1,2....k] \qquad F[Z] = F[k+1,k+2....k+k']$. *(C.1.6)*

Here are two applications we shall consider later on,

$F[Z] = F\ [k+1,k+2....k+k'] = S^{i_{k+1}i_{k+2}\ldots i_{k+k'}}$ = components of a rank-k' tensor

$F[Z] = F\ [k+1,k+2....k+k'] = \mathcal{S}(v_{i_{k+1}}, v_{i_{k+2}}, ... v_{i_{k+k'}})$ = a rank-k' tensor function . (C.2.7)

Now let P be a *general* permutation of [1,2....k+k'] = [z,Z].

$P[z,Z] = [P(z), P(Z)]$ *(C.1.8)*

Notice that RP and PR are undefined since P and R operate in different spaces, but R'P and PR' are both defined since both permutations R' and P operate in the space of [1,2....k+k'].

Recall now the meaning of S(R) as the number of swaps required to go from Z to R(Z) . This is the same as the number of swaps required to go from [z,Z] to R'[z,Z] = [z,R(Z)]. Therefore

$S(R) = S(R')$. (C.2.9)

We shall now prove the following theorem :

Theorem Two

$$\Sigma_P (-1)^{\sigma S(P)} f[P(z)]\, F[P(Z)] \;=\; \Sigma_P (-1)^{\sigma S(P)} f[P(z)]\, F^{\wedge}[P(Z)] \qquad (C.2.10)$$

where

$$F^{\wedge}[Z] \;\equiv\; (1/k'!)\, \Sigma_R (-1)^{\sigma S(R)} F[R(Z)] \qquad \sigma = 1 \text{ or } 0$$

The purpose of σ is to state the theorem with and without the $(-1)^{S(P)}$ factor.

We shall use the notation $F^{\wedge}$ in most of this section to apply for both values of σ, but at the end, we shall distinguish these two cases by writing:

$$F^{\wedge}[z] \;\equiv\; (1/k!)\, \Sigma_Q (-1)^{S(Q)} F[Q(z)] \qquad // = \text{Alt}(F), \text{ see (A.2.1)}$$
$$F_s[z] \;\equiv\; (1/k!)\, \Sigma_Q\, F[Q(z)] \qquad // = \text{Sym}(F), \text{ see (A.3.1)} \qquad (C.2.11)$$

At the end of this section we will show that the above Theorem Two with σ = 1 and σ = 0 is equivalent to the statements:

$$\text{Alt}(f\otimes F) = \text{Alt}(f\otimes F^{\wedge}) \qquad F^{\wedge} = \text{Alt}(F) \qquad \sigma = 1$$
$$\text{Sym}(f\otimes F) = \text{Sym}(f\otimes F_s) \qquad F_s = \text{Sym}(F) \qquad \sigma = 0\ . \qquad (C.2.12)$$

Proof of Theorem Two: Our first task is to process the second line of (C.2.10),

$$F_{\wedge}[Z] = (1/k'!)\ \Sigma_R\ (-1)^{\sigma S(R)}\ F[R(Z)]$$
$$= (1/k'!)\ \Sigma_R\ (-1)^{\sigma S(R')}\ F[R'(Z)]. \qquad / \text{(C.2.9) and (C.2.5a)} \qquad \text{(C.2.13)}$$

Apply permutation P to the above equation and use (A.2.8) to get

$$P\ F_{\wedge}[Z] = F_{\wedge}[P(Z)] = (1/k'!)\ \Sigma_R\ (-1)^{\sigma S(R')}\ F\ [PR'(Z)]\ . \qquad \text{(C.2.14)}$$

Then,

$$\text{RHS (C.2.10)} = \Sigma_P\ (-1)^{\sigma S(P)}\ f[P(z)]\ F_{\wedge}[P(Z)]$$

$$= \Sigma_P\ (-1)^{\sigma S(P)}\ f[P(z)]\{(1/k'!)\ \Sigma_R\ (-1)^{\sigma S(R')}\ F\ [PR'(Z)]\} \qquad // \text{(C.2.14) for } F_{\wedge}[P(Z)]$$

$$= (1/k'!)\Sigma_R\ \Sigma_P(-1)^{\sigma S(PR')}\ f[P(z)]\ F\ [PR'(Z)] \qquad // \text{ reorder and (A.1.10)}$$

$$= (1/k'!)\Sigma_R\ \Sigma_P(-1)^{\sigma S(PR')}\ f[PR'(z)]\ F\ [PR'(Z)] \qquad // R'(z) = z \text{ from (C.2.5b)}$$

$$= (1/k'!)\Sigma_R\ \Sigma_P(-1)^{\sigma S(P)}\ f[P(z)]\ F\ [P(Z)] \qquad // \text{ rearrangement theorem (A.1.3)}$$

$$= \Sigma_P(-1)^{\sigma S(P)}\ f[P(z)]\ F\ [P(Z)]\ \{\ (1/k'!)\ \Sigma_R\ (1)\ \} \qquad // \text{ reorder}$$

$$= \Sigma_P(-1)^{\sigma S(P)}\ f[P(z)]\ F\ [P(Z)]\ \{1\ \} \qquad // \ \Sigma_R\ (1) = k'! \text{ from (C.2.4)}$$

$$= \text{LHS (C.2.10)}\ . \qquad \text{QED} \qquad \text{(C.2.15)}$$

Using (C.1.15) and (C.1.16) , the second line of (C.2.10) can be restated

$$F_{\wedge} = \text{Alt}(F) \qquad \sigma = 1 \qquad \text{totally antisymmetric}$$
$$F_s = \text{Sym}(F) \qquad \sigma = 0 \qquad \text{totally symmetric}\ . \qquad \text{(C.2.18)}$$

Following the same arguments used the end of Section C.1, one obtains the following equivalent restatement of Theorem Two (just move the subscript from f to F)

$$\text{Alt}(f\otimes F) = \text{Alt}(f\otimes F_{\wedge}) \qquad F_{\wedge} = \text{Alt}(F) \qquad \text{(C.1.21)}$$
$$\text{Sym}(f\otimes F) = \text{Sym}(f\otimes F_s) \qquad F_s = \text{Sym}(F)\ . \qquad \text{(C.1.22)}$$

Alternate Proof of Theorem 2

An alternate proof of Theorem Two is two start with Theorem One and just make these changes

$$z \leftrightarrow Z \qquad f\leftrightarrow F \qquad k\leftrightarrow k' \qquad Q\rightarrow R$$

Here is Theorem One

$$\Sigma_P\, (-1)^{\sigma S(P)}\, f[P(z)]\, F[P(Z)] \;=\; \Sigma_P\, (-1)^{\sigma S(P)}\, f_{\wedge}[P(z)]\, F[P(Z)]$$

where

$$f_{\wedge}[z] \;\equiv\; (1/k!)\, \Sigma_Q\, (-1)^{\sigma S(Q)}\, f[Q(z)] \qquad \sigma = 1 \text{ or } 0 \qquad \textit{(C.1.10)}$$

and here is Theorem One with the above changes applied,

$$\Sigma_P\, (-1)^{\sigma S(P)}\, F[P(Z)]\, f[P(z)] \;=\; \Sigma_P\, (-1)^{\sigma S(P)}\, F_{\wedge}[P(Z)]\, f[P(z)]$$

where

$$F_{\wedge}[z] \;\equiv\; (1/k'!)\, \Sigma_R\, (-1)^{\sigma S(R)}\, F[R(Z)] \qquad \sigma = 1 \text{ or } 0\ . \qquad \textit{(C.1.10)}_{\mathbf{swap}}$$

This is the same as Theorem Two which we quote from above,

$$\Sigma_P\, (-1)^{\sigma S(P)}\, f[P(z)]\, F[P(Z)] \;=\; \Sigma_P\, (-1)^{\sigma S(P)}\, f[P(z)]\, F_{\wedge}[P(Z)]$$

where

$$F_{\wedge}[Z] \;\equiv\; (1/k'!)\, \Sigma_R\, (-1)^{\sigma S(R)}\, F[R(Z)] \qquad \sigma = 1 \text{ or } 0\ . \qquad \textit{(C.2.10)}$$

We went ahead with the detailed proof for two reasons. First, the swap proof might not be convincing to the reader. Second, the detailed proof provides steps which are crucial to proving Theorem Three below.

C.3 Theorem Three

Now both functions have a ^ subscript :

$$\Sigma_P\, (-1)^{\sigma S(P)}\, f[P(z)]\, F[P(Z)] \;=\; \Sigma_P\, (-1)^{\sigma S(P)}\, f_{\wedge}[P(z)]\, F_{\wedge}[P(Z)] \qquad (C.3.1)$$

where

$$f_{\wedge}[z] \;\equiv\; (1/k!)\, \Sigma_Q\, (-1)^{\sigma S(Q)}\, f[Q(z)]$$
$$F_{\wedge}[Z] \;\equiv\; (1/k'!)\, \Sigma_R\, (-1)^{\sigma S(R)}\, F[R(Z)] \qquad \sigma = 1 \text{ or } 0$$

The purpose of σ is to state the theorem with and without the $(-1)^{S(P)}$ factor.

At the end of this section we will show that the above Theorem Three with σ = 1 and σ = 0 is equivalent to the statements,

$$\mathrm{Alt}(f\otimes F) = \mathrm{Alt}(f_{\wedge}\otimes F_{\wedge})\ . \qquad f_{\wedge} = \mathrm{Alt}(f) \qquad F_{\wedge} = \mathrm{Alt}(F) \qquad \textit{(C.3.5)}$$
$$\mathrm{Sym}(f\otimes F) = \mathrm{Sym}(f_s\otimes F_s)\ . \qquad f_s = \mathrm{Sym}(f) \qquad F_s = \mathrm{Sym}(F)\ . \qquad \textit{(C.3.6)}$$

This theorem will involve both R and Q, as well as R' and Q' from earlier sections. Note that

$$R'Q' = Q'R' \qquad (C.3.3)$$

because Q' acts only on the lower integers in (1,2...k+k') while R' acts only on the upper integers.

Proof of Theorem Three: Recall these results from previous sections,

$$f_\wedge[P(z)] = (1/k!)\ \Sigma_Q\ (-1)^{\sigma S(Q')}\ f\,[PQ'(z)] \qquad (C.1.13)$$

$$F_\wedge[P(Z)] = (1/k'!)\ \Sigma_R\ (-1)^{\sigma S(R')}\ F\,[PR'(Z)]\ . \qquad (C.2.13)$$

Then,

$$\text{RHS (C.3.1)} = \Sigma_P\,(-1)^{\sigma S(P)}\ f_\wedge[P(z)]\ F_\wedge[P(Z)]$$

$$= \Sigma_P\,(-1)^{\sigma S(P)}\{\,(1/k!)\ \Sigma_Q\ (-1)^{\sigma S(Q')}\ f\,[PQ'(z)]\}\{\,(1/k'!)\ \Sigma_R\ (-1)^{\sigma S(R')}\ F\,[PR'(Z)]\}$$

$$= (1/k!)(1/k'!)\Sigma_Q\Sigma_R\ \ \Sigma_P\ (-1)^{\sigma S(PR'Q')}\ f\,[PQ'(z)]\ F\,[PR'(Z)] \qquad \text{// reorder and (A.1.10)}$$

$$= (1/k!)(1/k'!)\Sigma_Q\Sigma_R\ \ \Sigma_P\ (-1)^{\sigma S(PR'Q')}\ f\,[PQ'R'(z)]\ F\,[PR'Q'(Z)] \qquad \text{// } Q'(Z) = Z \text{ from (C.1.5b)}$$
$$\text{// } R'(z) = z \text{ from (C.2.5b)}$$

$$= (1/k!)(1/k'!)\Sigma_Q\Sigma_R\ \ \Sigma_P\ (-1)^{\sigma S(PR'Q')}\ f\,[PR'Q'(z)]\ F\,[PR'Q'(Z)] \qquad \text{// (C.3.3) } R'Q' = Q'R'$$

$$= (1/k!)(1/k'!)\Sigma_Q\Sigma_R\ \ \Sigma_P\ (-1)^{\sigma S(P[R'Q'])}\ f\,[P[R'Q'](z)]\ F\,[P[R'Q'](Z)]$$

$$= (1/k!)(1/k'!)\Sigma_Q\Sigma_R\ \ \Sigma_P\ (-1)^{\sigma S(P)}\ f\,[P(z)]\ F\,[P(Z)] \qquad \text{// rearrangement theorem (A.1.3)}$$

$$= \Sigma_P\ (-1)^{\sigma S(P)}\ \ f\,[P(z)]\ F\,[P(Z)]\ \ \{\,(1/k!)\Sigma_Q(1)\,\}\{\,(1/k'!)\Sigma_R(1)\,\} \qquad \text{// reorder}$$

$$= \Sigma_P\ (-1)^{\sigma S(P)}\ \ f\,[P(z)]\ F\,[P(Z)]\ \{\,1\,\}\{\,1\,\} \qquad \text{// (C.1.4) and (C.2.4)}$$

$$= \text{LHS (C.3.1)}\ . \qquad \text{QED} \qquad (C.3.4)$$

The endgame steps of Section C.1 are identical here with the change $F \rightarrow F_\wedge$, giving

$$\text{Alt}(f\otimes F) = \text{Alt}(f_\wedge \otimes F_\wedge)\ . \qquad f_\wedge = \text{Alt}(f) \qquad F_\wedge = \text{Alt}(F) \qquad (C.3.5)$$
$$\text{Sym}(f\otimes F) = \text{Sym}(f_s \otimes F_s)\ . \qquad f_s = \text{Sym}(f) \qquad F_s = \text{Sym}(F)\ . \qquad (C.3.6)$$

C.4 Summary and Generalization

Summary of the Three Theorems

Theorems One, Two and Three have shown that, in our generic function space,

$$\text{Alt}[T\otimes S] = \text{Alt}[T_\wedge \otimes S] = \text{Alt}[T\otimes S_\wedge] = \text{Alt}[T_\wedge \otimes S_\wedge]$$
$$\text{where} \qquad T_\wedge = \text{Alt}(T) \quad S_\wedge = \text{Alt}(S) \qquad (C.4.1)$$

$$\text{Sym}[T\otimes S] = \text{Sym}[T_s \otimes S] = \text{Sym}[T\otimes S_s] = \text{Sym}[T_s \otimes S_s]$$
$$\text{where} \qquad T_s = \text{Sym}(T) \quad S_s = \text{Sym}(S)\ . \qquad (C.4.2)$$

One can of course rewrite these statements as

$$\mathrm{Alt}[T\otimes S] = \mathrm{Alt}[\mathrm{Alt}(T)\otimes S] = \mathrm{Alt}[T\otimes \mathrm{Alt}(S)] = \mathrm{Alt}[\mathrm{Alt}(T)\otimes \mathrm{Alt}(S)] \qquad (C.4.3)$$

$$\mathrm{Sym}[T\otimes S] = \mathrm{Sym}[\mathrm{Sym}(T)\otimes S] = \mathrm{Sym}[T\otimes \mathrm{Sym}(S)] = \mathrm{Sym}[\mathrm{Sym}(T)\otimes \mathrm{Sym}(S)]\ . \qquad (C.4.4)$$

Intuitively these equations are easily interpreted:

If one is going to totally antisymmetrize a tensor product, the act of pre-antisymmetrizing one or more of the tensors makes no difference. So adding any $_\wedge$ subscripts to objects inside an Alt makes no difference.

If one is going to totally symmetrize a tensor product, the act of pre-symmetrizing one or more of the tensors makes no difference. So adding any s subscripts to objects inside an Alt makes no difference.

Various "theorems" can be generated by "adding hats" to the insides of an Alt expression.

<u>Example</u>: Consider.

$$\mathrm{Alt}[A\otimes B\otimes C] = \mathrm{Alt}[(A\otimes B)\otimes C] = \mathrm{Alt}[(A\otimes B)_\wedge\otimes C] = \mathrm{Alt}[\mathrm{Alt}(A\otimes B)\otimes C]$$
$$\mathrm{Alt}[A\otimes B\otimes C] = \mathrm{Alt}[A\otimes(B\otimes C)] = \mathrm{Alt}[A\otimes(B\otimes C)_\wedge] = \mathrm{Alt}[A\otimes \mathrm{Alt}(B\otimes C)] \qquad (C.4.5)$$

Therefore

$$\mathrm{Alt}[\mathrm{Alt}(A\otimes B)\otimes C] = \mathrm{Alt}[A\otimes B\otimes C] = \mathrm{Alt}[A\otimes \mathrm{Alt}(B\otimes C)]\ . \qquad (C.4.6)$$

Replacing A,B,C with the obscure names ω,η,θ gives

$$\mathrm{Alt}[\mathrm{Alt}(\omega \otimes \eta) \otimes \theta] = \mathrm{Alt}[\omega \otimes \eta \otimes \theta] = \mathrm{Alt}[\omega \otimes \mathrm{Alt}(\eta \otimes \theta)]\ . \qquad (C.4.7)$$

This may be compared with Spivak page 80 from which we quote,

$$\begin{aligned}(2)\ \ Alt(Alt(\omega \otimes \eta) \otimes \theta) &= Alt(\omega \otimes \eta \otimes \theta)\\ &= Alt(\omega \otimes Alt(\eta \otimes \theta)).\end{aligned} \qquad (C.4.8)$$

<u>Generalization of the three theorems</u>

The theorems derived above can be generalized in the following manner. Suppose for example we have a set of integers $[1,2,3......(k_1+k_2+...+k_N)] = [1,2,3....\kappa_N]$. Instead of partitioning this into 2 groups [z,Z] as done above, we partition the integers into N groups $[z_1, z_2....z_N]$ as follows:

$\kappa_1 = k_1$ // "cumulative ranks", as in (5.6.11)
$\kappa_2 = k_1+ k_2$
$\kappa_3 = k_1+ k_2 + k_3$
...
$\kappa_N = k_1 + k_2 + ... + k_N = \Sigma_{i=1}^N k_i$. *(5.6.11)*

$$[1,2,3....\kappa_N] = [z_1, z_2....z_N] \tag{C.4.9}$$

$z_1 = [1,2,3...\kappa_1]$ // the partitions
$z_2 = [\kappa_1+1,\kappa_2+2,...\kappa_2]$
$z_2 = [\kappa_2+1,\kappa_2+2,...\kappa_3]$
...
$z_N = [\kappa_{N-1}, \kappa_{N-1} + 1, ...\kappa_N]$.

And instead of functions f and F, we have functions $f_1, f_2....f_N$.

Whereas for N = 2 we had 2^2-1 = 3 theorems, for general N there will be 2^N - 1 theorems. If we define

$$L \equiv \Sigma_P (-1)^{\sigma S(P)} f_1[P(z_1)]f_2[P(z_2)] ... f_N[P(z_N)] \quad \text{// Left side of theorems} \tag{C.4.10}$$

then here are those theorems: (exercise for the reader: use induction or brute force)

1. $L = \Sigma_P (-1)^{\sigma S(P)} (f_1)_\wedge[P(z_1)]f_2[P(z_2)] ... f_N[P(z_N)]$
2. $L = \Sigma_P (-1)^{\sigma S(P)} f_1[P(z_1)](f_2)_\wedge[P(z_2)] ... f_N[P(z_N)]$
3. $L = \Sigma_P (-1)^{\sigma S(P)} (f_1)_\wedge[P(z_1)](f_2)_\wedge[P(z_2)] ... f_N[P(z_N)]$
4. $L = \Sigma_P(-1)^{\sigma S(P)} f_1[P(z_1)]f_2[P(z_2)](f_3)_\wedge[P(z_3)] ...f_N[P(z_N)]$

......

(2^N-1). $L = \Sigma_P(-1)^{\sigma S(P)} (f_1)_\wedge[P(z_1)](f_2)_\wedge[P(z_2)](f_3)_\wedge[P(z_3)] ...(f_N)_\wedge[P(z_N)]$ (C.4.11)

Translated to Alt/Sym, notation, we then find for the case N = 3,

$$\begin{aligned} &\text{Alt}[T\otimes S\otimes R] = \text{Alt}[T_\wedge\otimes S\otimes R] = \text{Alt}[T\otimes S_\wedge\otimes R] = \text{Alt}[T\otimes S\otimes R_\wedge] \\ &= \text{Alt}[T_\wedge\otimes S_\wedge\otimes R] = \text{Alt}[T_\wedge\otimes S\otimes R_\wedge] = \text{Alt}[T\otimes S_\wedge\otimes R_\wedge] = \text{Alt}[T_\wedge\otimes S_\wedge\otimes R_\wedge] \end{aligned} \tag{C.4.12}$$

$$\begin{aligned} &\text{Sym}[T\otimes S\otimes R] = \text{Sym}[T_s\otimes S\otimes R] = \text{Sym}[T\otimes S_s\otimes R] = \text{Sym}[T\otimes S\otimes R_s] \\ &= \text{Sym}[T_s\otimes S_s\otimes R] = \text{Sym}[T_s\otimes S\otimes R_s] = \text{Sym}[T\otimes S_s\otimes R_s] = \text{Sym}[T_s\otimes S_s\otimes R_s] \end{aligned} \tag{C.4.13}$$

One can write these using $X_\wedge$ = Alt(X) and X_s = Sym(X) to obtain nested equations as we did earlier.

In general one can write

$$\text{Alt}[(T_1)\otimes(T_2) ...... \otimes(T_N)] = \text{Alt}[(T_1)_{a_1}\otimes(T_2)_{a_2} ...... \otimes(T_N)_{a_N}] \tag{C.4.14}$$

where each a_i can *independently* be a blank, (T_i) , or can be a ^, $(T_i)_\wedge$. This then is the ultimate statement that arbitrary pre-antisymmetrizing of one or more tensors in a totally antisymmetric product makes no difference. Similarly,

$$\text{Sym}[(T_1)\otimes(T_2) ...... \otimes(T_N)] = \text{Sym}[(T_1)_{a_1}\otimes(T_2)_{a_2} ...... \otimes(T_N)_{a_N}] \tag{C.4.15}$$

where each a_i can *independently* be a blank, (T_i) , or can be an s, $(T_i)_s$. This then is the ultimate statement that arbitrary pre-symmetrizing of one or more tensors in a totally symmetric product makes no difference.

Application to Tensors and Tensor Functions

All the work done above in Appendix C has been "generic", meaning the various operations are with respect to generic permutation functions like f(1,2...k). The work can be applied to tensors or tensor functions according to these simple translation rules

$$f[1,2....k] = T^{i_1 i_2 \cdots i_k} = \text{components of a rank-k tensor}$$
$$f[1,2....k] = \mathcal{T}(v_{i_1}, v_{i_2}, ... v_{i_k}) = \text{a rank-k tensor function .} \qquad \textit{(C.1.7)} \qquad \text{(C.4.16)}$$

For example, consider our result (C.4.1) above that

$$\text{Alt}(T\otimes S) = \text{Alt}(T^\wedge \otimes S) \ . \qquad \textit{(C.4.1)} \qquad \text{(C.4.17)}$$

In the generic space this equation means

$$[\text{Alt}(T\otimes S)](1,2...k+k') = [\text{Alt}(T^\wedge\otimes S)](1,2...k+k') \ . \qquad \text{(C.4.18)}$$

Translated from the generic space to the tensor space, one gets

$$[\text{Alt}(T\otimes S)]^{i_1 i_2 \cdots i_{k+k'}} = [\text{Alt}(\text{Alt}(T)\otimes S)]^{i_1 i_2 \cdots i_{k+k'}} \qquad \text{(C.4.19)}$$

where for example

$$[\text{Alt}(T)]^{i_1 i_2 \cdots i_k} = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, T^{i_{P(1)} i_{P(2)} \cdots i_{P(k)}} \qquad \text{(C.4.20)}$$

Translated from the generic space to the tensor function space, one gets instead,

$$[\text{Alt}(\mathcal{T}\otimes\mathcal{S})](v_1, v_2.....v_{k+k'}) = [\text{Alt}(\text{Alt}(\mathcal{T})\otimes\mathcal{S})](v_1, v_2.....v_{k+k'}) \qquad \text{(C.4.21)}$$

where for example

$$[\text{Alt}(\mathcal{T})](v_1, v_2....v_k) = (1/k!)\, \Sigma_P\, (-1)^{S(P)}\, \mathcal{T}(v_{P(1)}, v_{P(2)}....v_{P(k)}) \qquad \text{(C.4.22)}$$

Here we follow our convention of putting dual-space tensor names into script/italic font.

Appendix D: A Unified View of Tensors and Tensor Functions

In this section multiindex notations are shown in red to the right.

D.1 Tensor functions in Dirac notation

The vector space V has dimension n, and k ≤ n.
The vector space is real, so <a|b> = <b|a>.

In the bra-ket notation (Paul Dirac, 1947), a rank-k tensor functional $\mathcal{T}$ is represented by the bra <T| which is an element of the dual space V^{*k}. Meanwhile, elements of the space V^k are written as kets which are a tensor product of smaller kets,

$$| v_{i_1}, v_{i_2}, ... v_{i_k} \rangle = |v_{i_1}\rangle \otimes |v_{i_2}\rangle \otimes ..... \otimes |v_{i_k}\rangle . \qquad | v_I \rangle \qquad (D.1.1)$$

Here the i_r are *labels, not components*. Each v_{i_1} is a vector in V having n components $(v_{i_1})^j$.

The tensor function $\mathcal{T}(v_{i_1}, v_{i_2}, ... v_{i_k})$ is then represented by the application of the functional <T| to vectors in V^k so that,

$$\langle T | v_{i_1}, v_{i_2}, ... v_{i_k} \rangle = \mathcal{T}(v_{i_1}, v_{i_2}, ... v_{i_k}) . \qquad \mathcal{T}(v_I) = \langle T | v_I \rangle \qquad (D.1.2)$$

Due to the tensor product (of vector spaces) construction of the "ket" shown in (D.1.1), the function shown in (D.1.2) is manifestly **k-multilinear**. We call this a "tensor function".

The bra-ket notation represents an inner product (scalar product) so the spaces here are Hilbert spaces, not just vector spaces.

As shown in (6.2.2), the covariant tensor $T_{i_1 i_2 .... i_k}$ may be written (each label i_r ranges from 1 to n),

$$T_{i_1 i_2 .... i_k} = \langle T | u_{i_1}, u_{i_2}, ... u_{i_k} \rangle . \qquad T_I = \langle T | u_I \rangle \qquad (D.1.3)$$

The vectors $|u_i\rangle$ for i=1 to n form a basis for V, and the n*k kets $| u_{i_1}, u_{i_2}, ... u_{i_k} \rangle$ form a basis for V^k.

From (D.1.2) and (D.1.3), one concludes that

$$T_{i_1 i_2 .... i_k} = \mathcal{T}(u_{i_1}, u_{i_2}, ... u_{i_k}) \qquad i_r = 1 \text{ to } n \qquad T_I = \mathcal{T}(u_I) \qquad (D.1.4)$$

in agreement with (6.2.5). The contravariant form is then

$$T^{i_1 i_2 .... i_k} = \mathcal{T}(u^{i_1}, u^{i_2}, ... u^{i_k}) . \qquad T^I = \mathcal{T}(u^I) \qquad (D.1.5)$$

Let us now assume that the n vectors $|\mathbf{v}_i>$ for i=1 to n form some alternative basis for V, and then the n*k kets $| v_{i_1},v_{i_2}, ... v_{i_k} >$ form an alternative basis for V^k. The dual basis is $\{v^i\}$ where $v^i \bullet v_j = \delta^i{}_j$ as in (2.11.c.1) for the u_i basis and its dual u^j.

Looking at our two equations from above,

$$<T | v_{i_1},v_{i_2}, ... v_{i_k} > \ = \mathcal{T}(v_{i_1},v_{i_2}, ... v_{i_k}) \qquad i_r = 1 \text{ to } n \qquad (D.1.2)$$

$$<T | u_{i_1},u_{i_2}, ... u_{i_k} > \ = \ T_{i_1 i_2 ....i_k} \qquad i_r = 1 \text{ to } n \qquad (D.1.3)$$

one can say that the tensor $T_{i_1 i_2 ....i_k}$ and the tensor function $\mathcal{T}(v_{i_1},v_{i_2}, ... v_{i_k})$ are both representations of the same abstract tensor $<T|$ in two different V^k bases, $|v_I>$ and $|u_I>$. Both bases have dimension n*k. Recall

$$T = \Sigma_I T^I u_I \quad \in \quad V^k \quad = \text{a tensor} \qquad |T> = \Sigma_I T^I |u_I >$$

$$\mathcal{T} = \Sigma_I T_I \lambda^I \quad \in \quad V^{*k} \ = \text{a tensor functional} \qquad <T| \ = \ \Sigma_I T_I <u^I|$$

$$\mathcal{T} \sim T \text{ by the isomorphism } V^{*k} \sim V^k \ \text{[see below (2.11.b.1)]} . \qquad (D.1.6)$$

Notice that for the basis $\{v_i\}$,

$$<v^I|v_J> = < v^{i_1}|v_{j_1}>< v^{i_2}|v_{j_2}> \ .... < v^{i_k}|v_{j_k}>$$

$$= (v^{i_1} \bullet v_{j_1})(v^{i_2} \bullet v_{j_2}) \ .... (v^{i_k} \bullet v_{j_k})$$

$$= \delta^{i_1}{}_{j_1}\delta^{i_2}{}_{j_2}...\delta^{i_k}{}_{j_k} \qquad \text{// see (2.3.2) for basis } \{v_r\} \text{ with dual basis } \{v^r\}$$

$$= \ \delta^I{}_J \ . \qquad \text{// orthonormal basis in the multiindex notation} \qquad (D.1.7)$$

This result applies as well to the basis $|u_I>$, so

$$<u^I|u_J> \ = <v^I|v_J> \ = \ \delta^I{}_J \ . \qquad (D.1.8)$$

D.2 Basis change matrix

The basis-change transformation matrix between the $|v_I>$ and $|e_I>$ bases is given by,

$$M^I{}_J \ \equiv \ < u^{i_1},u^{i_2}, ... u^{i_k} | v_{j_1},v_{j_2}, ... v_{j_k} > \qquad M^I{}_J \ = <u^I|v_J> \qquad (D.2.1)$$

$$= < u^{i_1}|v_{j_1}>< u^{i_2}|v_{j_2}> \ .... < u^{i_k}|v_{j_k}> \qquad \text{// see (2.9.17)}$$

$= (u^{i_1} \bullet v_{j_1})(u^{i_2} \bullet v_{j_2}) \ .... \ (u^{i_k} \bullet v_{j_k})$

$= \lambda^{i_1}(v_{j_1})\lambda^{i_2}(v_{j_2}) \ .... \ \lambda^{i_k}(v_{j_k})$ // see (2.11.c.5)

$= (v_{j_1})^{i_1} (v_{j_2})^{i_2} ...(v_{j_k})^{i_k}$

$= (v_J)^I$. // multiindex notation (D.2.2)

There are n*k values for I and n*k values for J, so matrix M has dimension nk x nk.

Entirely in multiindex notation (see (2.1.7)) ,

$M^{IJ} = \langle u^I|v^J\rangle = (v^J)^I$ // pure contravariant
$M^I{}_J = \langle u^I|v_J\rangle = (v_J)^I$ // mixed
$M_I{}^J = \langle u_I|v^J\rangle = (v^J)_I$ // mixed
$M_{IJ} = \langle u_I|v_J\rangle = (v_J)_I$. // purecovariant (D.2.3)

The covariant transpose (see Section 2.11 (f)) is then,

$(M^T)_{JI} = M_{IJ} = \langle u_I|v_J\rangle = \langle v_J|\, u_I\rangle = (v_J)_I$ // Hilbert Space is real

$(M^T)_J{}^I = M^I{}_J = \langle u^I|v_J\rangle = \langle v_J|\, u^I\rangle = (v_J)^I$. (D.2.4)

In the bra-ket notation completeness of an orthonormal basis is expressed this way:

$1 = \Sigma_J |u_J\rangle\langle u^J| = \Sigma_J |u^J\rangle\langle u_J|$ // see Section 2.11 (h)
$= \Sigma_J |v_J\rangle\langle v^J| = \Sigma_J |v^J\rangle\langle v_J|$. (D.2.5)

<u>Proof</u>: (example) Consider a general V^k tensor T :

(1) $|T\rangle = 1|T\rangle = \Sigma_J |u_J\rangle\langle u^J|\, T\rangle = \Sigma_J T^J |u_J\rangle$ // so basis $|u_J\rangle$ must be complete

(2) $|u_I\rangle = 1|u_I\rangle = \Sigma_J |u_J\rangle\langle u^J|u_I\rangle = \Sigma_J |u_J\rangle\delta^J{}_I = |u_I\rangle$ // why orthonormal is needed

Therefore the basis-change matrix M has the property $MM^T = 1$ or $M^T = M^{-1}$:

$(MM^T)_I{}^K = \Sigma_J M_I{}^J(M^T)_J{}^K = \Sigma_J \langle u_I|v^J\rangle\langle v_J|\, u^K\rangle = \langle u_I|\ (\Sigma_J|v^J\rangle\langle v_J|\,)\, |u^K\rangle$

$= \langle u_I \,|\, 1 \,|\, u^K\rangle = \langle u_I \,|\, u^K\rangle = u_I \bullet u^K = \delta_I{}^K$ // see (2.11.2)

so (D.2.6)

$MM^T = 1$.

The connection then between the tensors and tensor functions is given by,

$$T_I = \langle u_I | T\rangle = \langle u_I | 1 | T\rangle = \langle u_I | \Sigma_J |v^J\rangle\langle v_J| T\rangle$$

$$= \Sigma_J \langle u_I|v^J\rangle\langle v_J| T\rangle = \Sigma_J \langle u_I|v^J\rangle\langle T| v_J\rangle$$

$$= \Sigma_J M_I{}^J \mathcal{T}(v_J) . \qquad \text{(D.2.7)}$$

Going the other direction,

$$\mathcal{T}(v_I) = \langle v_I | T \rangle = \langle v_I | 1 | T \rangle = \langle v_I | \Sigma_J |u^J\rangle\langle u_J| T \rangle$$

$$= \Sigma_J \langle v_I|u^J\rangle \langle u_J|T \rangle = \Sigma_J \langle u^J|v_I\rangle \langle T| u_J \rangle$$

$$= \Sigma_J M^J{}_I T_J = \Sigma_J (M^T)_I{}^J T_J . \qquad \text{(D.2.8)}$$

Example of (D.2.7):

$$T_{i_1 i_2} = \Sigma_{j_1,j_2=1}{}^n (v^{j_1})_{i_1}(v^{j_2})_{i_2} \mathcal{T}(v_{j_1},v_{j_2}) \qquad i_r = 1 \text{ to } n$$

or

$$T_{ij} = \Sigma_{a,b=1}{}^n (v^a)_i(v^b)_j \mathcal{T}(v_a,v_b) . \qquad // \ n^2 \text{ terms in the sum} \qquad \text{(D.2.7a)}$$

Example of (D.2.8):

$$\mathcal{T}(v_{i_1},v_{i_2}) = \Sigma_{j_1,j_2=1}{}^n (v_{i_1})^{j_1} (v_{i_2})^{j_2} T_{j_1 j_2} \qquad i_r = 1 \text{ to } n$$

or

$$\mathcal{T}(v_i,v_j) = \Sigma_{a,b=1}{}^n (v_i)^a (v_j)^b T_{ab} . \qquad // \ n^2 \text{ terms in the sum} \qquad \text{(D.2.8a)}$$

Comment: These examples can be compared to a simple quantum mechanics case. Let |x> be a basis vector describing a 1D particle at location x (coordinate representation), and let |p> be a basis vector describing a plane-wave particle having momentum p (momentum representation). Then it turns out that the basis change matrix is $\langle x|p\rangle = \psi_p(x) = C\, e^{ip\bullet x}$ where C is a normalization constant. So the basis change "matrix" (continuous matrix subscripts p and x) is a function of p, just as the basis change matrix in (D.2.8a) is a function of v_i and v_j.

D.3 Transformations of tensors and tensor functions

In this section we write vectors in bold font.

Consider two sets of n vectors $\mathbf{v}_i$ and $\mathbf{v'}_i$ where $\mathbf{v}_i$ form a basis for V. One can then write,

$$\mathbf{v'}_i = Q_i{}^j \mathbf{v}_j \qquad i = 1,2...n \qquad \text{implied sum on } j \qquad \text{(D.3.1)}$$

where $Q_i{}^j$ is a matrix describing the linear combinations of the $\mathbf{v}_i$ that make up the $\mathbf{v'}_i$. Since the tensor function $\mathcal{T}(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, ... \mathbf{v}_{i_k})$ is k-multilinear, one can certainly write

$$\mathcal{T}(\mathbf{v'}_{i_1},\mathbf{v'}_{i_2}, ... \mathbf{v'}_{i_k}) = Q_{i_1}{}^{j_1}Q_{i_2}{}^{j_2} .... Q_{i_k}{}^{j_k}\, \mathcal{T}(\mathbf{v}_{j_1},\mathbf{v}_{j_2}, ... \mathbf{v}_{j_k}) \qquad \text{(D.3.2)}$$

or just showing the ket part,

$$|\, \mathbf{v'}_{i_1},\mathbf{v'}_{i_2}, ... \mathbf{v'}_{i_k} > \; = \; Q_{i_1}{}^{j_1}Q_{i_2}{}^{j_2} .... Q_{i_k}{}^{j_k} \,|\, \mathbf{v}_{j_1},\mathbf{v}_{j_2}, ... \mathbf{v}_{j_k} > \quad . \qquad \text{(D.3.3)}$$

Equation (D.3.2) vaguely resembles the Chapter 2 transformation of a covariant tensor field,

$$T'_{i_1 i_2 ... i_k}(\mathbf{x'}) = R_{i_1}{}^{j_1}R_{i_2}{}^{j_2} .... R_{i_k}{}^{j_k}\, T_{j_1 j_2 ... j_k}(\mathbf{x}) \qquad \text{(D.3.4)}$$

where

$\mathbf{x'} = \mathbf{F}(\mathbf{x})$ and $d\mathbf{x'} = R\, d\mathbf{x}$. // R is the differential of **F**.

The resemblance is perhaps closer if we restrict $\mathbf{x'} = \mathbf{F}(\mathbf{x})$ to be a linear transformation, so then

$$\mathbf{x'} = R\,\mathbf{x} \qquad \text{or} \qquad x'^i = R^i{}_j x^j \quad . \qquad \text{(D.3.5)}$$

The resemblance between (D.3.2) and (D.3.4) we claim is really superficial and misleading, which is the main reason for bringing it up. We just make a few comments on this matter.

• The linearized transformations of Chapter 2 like $v'^i = R^i{}_j v^j$ for a vector v^i are *component* transformations. The j on v^j is a component index, and $v'^i = R^i{}_j v^j$ ($\mathbf{v'} = R\mathbf{v}$) is an instruction for creating a new vector **v'** by linearly combining the components of **v**. Transformation (D.3.5) is such a component transformation.

• In contrast, the transformation (D.3.1) that $\mathbf{v'}_i = Q_i{}^j \mathbf{v}_j$ is not a component transformation. It constructs n new vectors $\mathbf{v'}_i$ by linearly combining the n vectors $\mathbf{v}_i$. The j on $\mathbf{v}_j$ is a label, not a component index.

• In (D.3.4), the left-side object $T'_{i_1 i_2 ... i_k}(\mathbf{x'})$ has a prime on T. It is a tensor different from $T_{j_1 j_2 ... j_k}(\mathbf{x})$, and this would be true even if there were no **x** dependence of the field.

• In (D.3.2), the left-side object $\mathcal{T}(\mathbf{v'}_{i_1},\mathbf{v'}_{i_2}, ... \mathbf{v'}_{i_k})$ has no prime, it is the same $\mathcal{T}$ as on the right.

• In fact, as was shown in (2.11.e.8), the object $\mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k)$ under any Chapter 2 component transformation transforms as a scalar, so there are no $R_i{}^j$ or $Q_i{}^j$ matrices involved,

$$\mathcal{T}'(\mathbf{v'}_1, \mathbf{v'}_2, ...\mathbf{v'}_k) = \mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k) \; . \qquad \mathcal{T}'(\mathbf{v'}_z) = \mathcal{T}(\mathbf{v}_z) \qquad \textit{(2.11.f.8)}$$

where

$(\mathbf{v'}_r)^i = R^i{}_j\, (\mathbf{v}_r)^j$ $\quad r = 1...k \quad$ implied sum on j $\quad i = 1...n$

$\mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k)$ is a scalar because it is the scalar product of a rank-k tensor functional $\mathcal{T}$ = <T| with a rank-k tensor $|\, \mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k >$, just as $<\mathbf{a}\,|\,\mathbf{b}> = \mathbf{a} \bullet \mathbf{b}$ is a scalar.

• (D.3.2) is nothing more than a statement that the tensor function $\mathcal{T}(\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k)$ is k-multilinear.

D.4 Tensor Functions and Quantum Mechanics

Eq. (D.1.2) defining a tensor function as a bra-ket combination

$$<T \mid \mathbf{v}_{i_1}, \mathbf{v}_{i_2}, ... \mathbf{v}_{i_k} > \; = \mathcal{T}(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, ... \mathbf{v}_{i_k}) \qquad \textit{(D.1.2)}$$

has the following quantum mechanics incarnation, which was the original use Dirac intended for his bra-ket notation,

$$<\psi \mid \mathbf{r}_1, \mathbf{r}_2...\mathbf{r}_k> \; = \psi(\mathbf{r}_1, \mathbf{r}_2...\mathbf{r}_k) \; . \qquad \text{(D.4.1)}$$

Here object $<\psi|$ plays the role of the abstract tensor $<T|$, and the generic arguments $\mathbf{v}_{i_r}$ become the physical positions $\mathbf{r}_i$ of k particles. The object ψ is a functional in V^{*k} which gets applied to $|\mathbf{r}_1,\mathbf{r}_2....\mathbf{r}_k>$ = $|\mathbf{r}_1>\otimes|\mathbf{r}_2>....\otimes|\mathbf{r}_k>$ and the resulting function $\psi(\mathbf{r}_1,\mathbf{r}_2...\mathbf{r}_k)$ is called a "wave function" which describes the "probability amplitude" that the k particles are near spatial locations $\mathbf{r}_1$, $\mathbf{r}_2$, ...$\mathbf{r}_k$. The probability that the k particles are near these spatial locations is given by $|\psi(\mathbf{r}_1,\mathbf{r}_2...\mathbf{r}_k)|^2 d^n\mathbf{r}_1 d^n\mathbf{r}_2 ...d^n\mathbf{r}_k$. "Near" means that $\mathbf{r}_i$ lies somewhere in the range $\mathbf{r}_i$ to $\mathbf{r}_i + d^n\mathbf{r}_i$. Normally one uses n = 3 for 3D space.

It happens that in quantum mechanics literature it is the ket that is the functional in V^{*k} and the bra which is the element of V^k. So in a physics text one always sees equations like,

$$\psi(\mathbf{r}_1, \mathbf{r}_2...\mathbf{r}_k) = <\mathbf{r}_1, \mathbf{r}_2...\mathbf{r}_k \mid \psi>. \qquad \text{(D.4.2)}$$

This is a long-standing convention difference between the physics and math worlds. When talking about a functional f applied to a vector **x**, it seems natural to have $f(\mathbf{x}) = <f \mid \mathbf{x}>$, which is the math convention. The physics person writes $<x|\psi> = \psi(x)$ and says that the *state vector* $|\psi>$ is being projected onto the coordinate representation basis element $<x|$. Usually ψ is not called a "functional". A ket is thought of as a vector v, while the bra is a transpose vector v^T and then $<v_1|v_2> = v_1^T v_2$ in a matrix notation sense, so here it seems logical to put "the vector", whether v, v_2 or ψ, on the right.

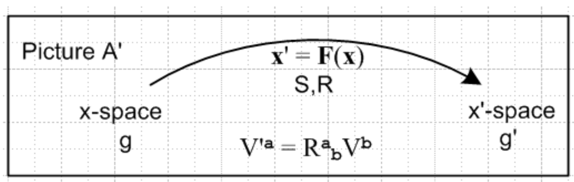
Our functional $\mathcal{T}$ maps elements of V^k to the real numbers, and $<a|b> = <b|a> = a \bullet b$, so one can "for free" switch the role of which is the functional, and which is the ket acted upon by the functional. In quantum mechanics the functional maps to complex numbers, and $<a|b> = a^* \bullet b$ where * is complex conjugation. Then $<b|a> = b^* \bullet a = (a \bullet b^*)^* = a^* \bullet b = <a|b>^*$. And $<v_1|v_2> = v_1^{T*} v_2 = v_1^\dagger v_2$. It is a crucial element of quantum mechanics that the space V^k is complex and not real. In the math world, one usually sees instead $<b|a> = b \bullet a^*$.

If the k particles are electrons or other half-integral spin particles which are in a "symmetric spin state", then the wavefunction (D.1.4) must be replaced by $[\mathrm{Alt}(\psi)](\mathbf{r}_1, \mathbf{r}_2...\mathbf{r}_k)$ in order to make it be totally antisymmetric in the coordinates $\mathbf{r}_i$, as required by "Fermi statistics" for half-integral spin particles. We mention this just to show that both the Alt operator and more generally the permutation group of Appendix A have important applications in quantum mechanics.

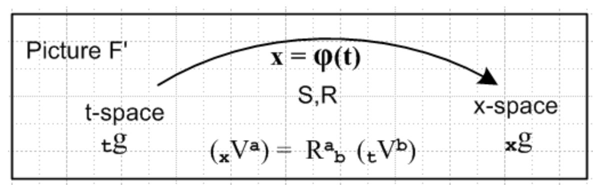
Appendix E: Kinematics Package with x' = F(x) changed to x = φ(t)

The material here is just for completeness and is intended only for perusal. It shows how the development of Chapter 10 appears for $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ in place of $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. In some ways, the $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ *results* concerning differential forms are simpler that those expressed in the $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ notation. The less pleasant aspect is that tensors (including metric tensors), basis vectors, and their spaces need an extra x or t label to distinguish the two spaces (now t-space and x-space), whereas in the $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ approach this distinction is accomplished by a prime versus no prime. We do use part of this notation in Section 10.9 since it brings our results into a more standard form for comparison with other sources.

Translation Table

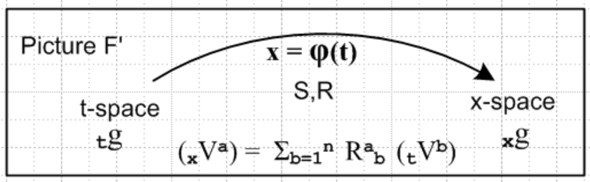


→



x-space	→	t-space	
x'-space	→	x-space	
$\mathbf{F}$	→	$\boldsymbol{\varphi}$	general transformation name
$\mathbf{x}' = \mathbf{F}(\mathbf{x})$	→	$\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$	general transformation equation
R,S	→	R,S	differential matrices (no change in name)
F*	→	φ*	pullback function
$\mathbf{V}$	→	${}_t\mathbf{V}$	vector in t-space
$\mathbf{V}'$	→	${}_x\mathbf{V}$	vector in x-space
$\mathbf{e}$	→	${}_t\mathbf{e}$	tangent base vectors in t-space
$\mathbf{u}$	→	${}_t\mathbf{u}$	axis-aligned basis vectors in t-space
g	→	${}_t g$	metric tensor in t-space
$\mathbf{u}'$	→	${}_x\mathbf{u}$	tangent base vectors in x-space
$\mathbf{e}'$	→	${}_x\mathbf{e}$	axis-aligned basis vectors in x-space
g'	→	${}_x g$	metric tensor in x-space
Λ'^k	→	${}_x\Lambda^k$	dual space to R^m
Λ^k	→	${}_t\Lambda^k$	dual space to R^n
$\lambda'^i = dx'^i$	→	${}_x\lambda^i = dx^i$	basis vector in dual space to R^m
$\lambda^i = dx^i$	→	${}_t\lambda^i = dt^i$	basis vector in dual space to R^m (E.1)

In this new notation, the "kinematics package" of (10.6.a.1) with adjustment (10.6.d.1) for "tall" R appears as



(a) $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ xform $\qquad R^i{}_j \equiv (\partial x^i/\partial t^j) = \partial_j^{(t)} x^i \qquad R = (D\varphi)$

$\quad {}_x\mathbf{V} = R\ {}_t\mathbf{V}$ vector $\qquad S^i{}_j \equiv (\partial t^i/\partial x^j) = \partial_j^{(x)} t^i$

(b) ${}_x\mathbf{e}_i$ with $({}_x\mathbf{e}_i)^j = \delta_i{}^j$ — axis-aligned basis vectors in x-space (i = 1..m)

$\quad {}_t\mathbf{e}_i \qquad {}_t\mathbf{e}_i = S\ {}_x\mathbf{e}_i$ — tangent base vectors in x-space (i = 1..n)

(c) ${}_t\mathbf{u}_i$ with $({}_t\mathbf{u}_i)^j = \delta_i{}^j$ — axis-aligned basis vectors in t-space (i = 1..n)

$\quad {}_x\mathbf{u}_i \qquad {}_x\mathbf{u}_i = R\ {}_t\mathbf{u}_i$ — tangent base vectors in t-space (i = 1..n)

$\qquad ({}_x\mathbf{u}_i)^j = R^j{}_k\ ({}_t\mathbf{u}_i)^k$

(d) ${}_x1 = |{}_x\mathbf{e}_i\rangle\langle{}_x\mathbf{e}^i| = |{}_x\mathbf{e}^i\rangle\langle{}_x\mathbf{e}_i| = |{}_x\mathbf{u}_i\rangle\langle{}_x\mathbf{u}^i| = |{}_x\mathbf{u}^i\rangle\langle{}_x\mathbf{u}_i|$ completeness in x-space

$\quad {}_t1 = |{}_t\mathbf{e}_i\rangle\langle{}_t\mathbf{e}^i| = |{}_t\mathbf{e}^i\rangle\langle{}_t\mathbf{e}_i| = |{}_t\mathbf{u}_i\rangle\langle{}_t\mathbf{u}^i| = |{}_t\mathbf{u}^i\rangle\langle{}_t\mathbf{u}_i|$ completeness in t-space

(e) $({}_t\mathbf{u}_j)^i = {}_t\mathbf{u}^i \bullet {}_t\mathbf{u}_j = \langle{}_t\mathbf{u}^i\,|\,{}_t\mathbf{u}_j\rangle = {}_tg^i{}_j = {}_x\mathbf{u}^i \bullet {}_x\mathbf{u}_j = \langle{}_x\mathbf{u}^i\,|\,{}_x\mathbf{u}_j\rangle$

$\quad ({}_t\mathbf{e}_j)^i = {}_t\mathbf{u}^i \bullet {}_t\mathbf{e}_j = \langle{}_t\mathbf{u}^i\,|\,{}_t\mathbf{e}_j\rangle = S^i{}_j = R_j{}^i$

$\quad ({}_x\mathbf{e}_j)^i = {}_x\mathbf{e}^i \bullet {}_x\mathbf{e}_j = \langle{}_x\mathbf{e}^i\,|\,{}_x\mathbf{e}_j\rangle = {}_xg^i{}_j = {}_t\mathbf{e}^i \bullet {}_t\mathbf{e}_j = \langle{}_t\mathbf{e}^i\,|\,{}_t\mathbf{e}_j\rangle$

$\quad ({}_x\mathbf{u}_j)^i = {}_x\mathbf{e}^i \bullet {}_x\mathbf{u}_j = \langle{}_x\mathbf{e}^i\,|\,{}_x\mathbf{u}_j\rangle = R^i{}_j = S_j{}^i$

(f) ${}_t\mathbf{e}^i = {}_xg^{ij}\ {}_t\mathbf{e}_j \qquad {}_x\mathbf{e}^i = {}_xg^{ij}\ {}_x\mathbf{e}_j \qquad {}_t\mathbf{u}^i = {}_tg^{ij}\ {}_t\mathbf{u}_j \qquad {}_x\mathbf{u}^i = {}_tg^{ij}\ {}_x\mathbf{u}_j$

$\quad {}_t\mathbf{e}_i = {}_xg_{ij}\ {}_t\mathbf{e}^j \qquad {}_x\mathbf{e}_i = {}_xg_{ij}\ {}_x\mathbf{e}^j \qquad {}_t\mathbf{u}_i = {}_tg_{ij}\ {}_t\mathbf{u}^j \qquad {}_x\mathbf{u}_i = {}_tg_{ij}\ {}_x\mathbf{u}^j$

(g) $\langle{}_t\mathbf{e}_j\,|\,\mathcal{S}\,|\,{}_x\mathbf{e}^i\rangle = \langle{}_x\mathbf{e}^i\,|\,\mathcal{R}\,|\,{}_t\mathbf{e}_j\rangle = {}_xg^i{}_j$

$\quad \langle{}_t\mathbf{e}_j\,|\,\mathcal{S}\,|\,{}_x\mathbf{u}^i\rangle = \langle{}_x\mathbf{u}^i\,|\,\mathcal{R}\,|\,{}_t\mathbf{e}_j\rangle = S^i{}_j = R_j{}^i$

$\quad \langle{}_t\mathbf{u}_j\,|\,\mathcal{S}\,|\,{}_x\mathbf{e}^i\rangle = \langle{}_x\mathbf{e}^i\,|\,\mathcal{R}\,|\,{}_t\mathbf{u}_j\rangle = R^i{}_j = S_j{}^i$

$\quad \langle{}_t\mathbf{u}_j\,|\,\mathcal{S}\,|\,{}_x\mathbf{u}^i\rangle = \langle{}_x\mathbf{u}^i\,|\,\mathcal{R}\,|\,{}_t\mathbf{u}_j\rangle = {}_tg^i{}_j\,.$

(h) $S = R^T \qquad S^i{}_j = (R^T)^i{}_j = R_j{}^i$

$\quad R = S^T \qquad R^i{}_j = (S^T)^i{}_j = S_j{}^i$

(i) $SR = 1 \qquad SS^T = R^TR = 1$ *(10.6.a.1)* (E.2)

The uniqueness table of (10.6.d.2) becomes the following,

<u>Metric tensors</u>

${}_tg_{ij}, {}_tg^{ij}$ unique

${}_xg^{ij}$ unique, since ${}_xg^{ij} = R^i{}_aR^j{}_b\ {}_tg^{ab}$

${}_xg_{ij}$ **not** unique, since ${}_xg_{ij} = R_i{}^aR_j{}^b\ {}_tg_{ab} = S^a{}_iS^b{}_j\ {}_tg_{ab}$ and $S^i{}_j$ not unique

Transformation matrices

$R^i{}_j = S_j{}^i$	unique (tall R matrix from **x**' = **F**(**x**))
$R^{ij} = S^{ji}$	unique since $R^{ij} = {}_tg^{ja}R^i{}_a$ and both ${}_tg^{ja}$ and $R^i{}_a$ are unique
$R_j{}^i = S^i{}_j$	**not** unique, see (10.6.c.3)
$R_{ij} = S_{ji}$	**not** unique, since $R_{ij} = {}_xg_{ia}\, R^a{}_j$ and ${}_xg_{ia}$ not unique

Axis-aligned basis vectors

$({}_t\mathbf{u}_j)^i$	unique since $({}_t\mathbf{u}_j)^i = {}_tg_j{}^i$	$({}_x\mathbf{e}_j)^i$	unique since $({}_x\mathbf{e}_j)^i = {}_xg^i{}_j\ (= \delta^i{}_j)$
$({}_t\mathbf{u}_j)_i$	unique since $({}_t\mathbf{u}_j)_i = {}_tg_{ji}$	$({}_x\mathbf{e}_j)_i$	**not** unique since $({}_x\mathbf{e}_j)_i = {}_xg_{ij}$
$({}_t\mathbf{u}^j)^i$	unique since $({}_t\mathbf{u}^j)^i = {}_tg^{ji}$	$({}_x\mathbf{e}^j)^i$	unique since $({}_x\mathbf{e}^j)^i = {}_xg^{ij}$
$({}_t\mathbf{u}^j)_i$	unique since $({}_t\mathbf{u}^j)_i = {}_tg^j{}_i$	$({}_x\mathbf{e}^j)_i$	unique since $({}_x\mathbf{e}^j)_i = {}_xg_i{}^j$

Tangent base vectors

$({}_t\mathbf{e}_j)^i$	**not** unique since $({}_t\mathbf{e}_j)^i = R_j{}^i$	$({}_x\mathbf{u}_j)^i$	unique since $({}_x\mathbf{u}_j)^i = R^i{}_j$
$({}_t\mathbf{e}_j)_i$	**not** unique since $({}_t\mathbf{e}_j)_i = R_{ji}$	$({}_x\mathbf{u}_j)_i$	**not** unique since $({}_x\mathbf{u}_j)_i = R_{ij}$
$({}_t\mathbf{e}^j)^i$	unique since $({}_t\mathbf{e}^j)^i = R^{ji}$	$({}_x\mathbf{u}^j)^i$	unique since $({}_x\mathbf{u}^j)^i = R^{ij}$
$({}_t\mathbf{e}^j)_i$	unique since $({}_t\mathbf{e}^j)_i = R^j{}_i$	$({}_x\mathbf{u}^j)_i$	**not** unique since $({}_x\mathbf{u}^j)_i = R_i{}^j$

(10.6.d.2) (E.3)

Other parts of the development translate as follows:

Basis Vectors

$\{{}_t\mathbf{u}_i\}$ i = 1,2...n basis for t-space, axis-aligned

$({}_t\mathbf{u}_i)^j = \delta_i{}^j$ components of these basis vectors in t-space . *(10.6.e.1)* (E.4)

$${}_x\mathbf{u}_i = \begin{cases} R\ {}_t\mathbf{u}_i & i = 1 \text{ through } n \text{ (tangent base vectors)} \\ \text{as needed} & i = n+1 \text{ through } m \end{cases} \qquad (10.6.e.4) \qquad (E.5)$$

$R^*{}_* = [{}_x\mathbf{u}_1, {}_x\mathbf{u}_2 \ldots {}_x\mathbf{u}_n]$ R has full rank n ⇒ basis for T_xM complete *(10.6.e.5)* (E.6)

Non-Dual Pull Backs

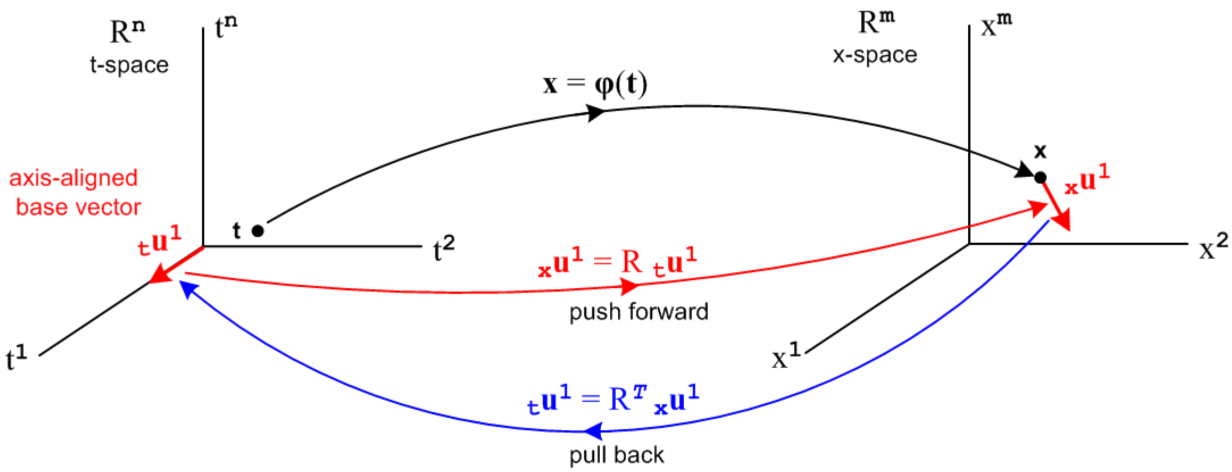
${}_x\mathbf{u}_i = R\ {}_t\mathbf{u}_i$	$\vert{}_x\mathbf{u}_i\rangle = \mathcal{R}\ \vert{}_t\mathbf{u}_i\rangle$	i = 1,2..n	push forward	*(10.7.1)*	(E.7)
${}_t\mathbf{u}_i = S\ {}_x\mathbf{u}_i$	$\vert{}_t\mathbf{u}_i\rangle = \mathcal{S}\ \vert{}_x\mathbf{u}_i\rangle$	i = 1,2..n	pull back	*(10.7.2)*	(E.8)
${}_x\mathbf{u}^i = R\ {}_t\mathbf{u}^i$	$\vert{}_x\mathbf{u}^i\rangle = \mathcal{R}\ \vert{}_t\mathbf{u}^i\rangle$	i = 1,2..n	push forward		
${}_t\mathbf{u}^i = R^T\ {}_x\mathbf{u}^i$	$\vert{}_t\mathbf{u}^i\rangle = \mathcal{R}^T\ \vert{}_x\mathbf{u}^i\rangle$	i = 1,2..n	pull back	*(10.7.4)*	(E.9)
${}_x\mathbf{e}^i = R\ {}_t\mathbf{e}^i$	$\vert{}_x\mathbf{e}^i\rangle = \mathcal{R}\ \vert{}_t\mathbf{e}^i\rangle$	i = 1,2..n	push forward		
${}_t\mathbf{e}^i = R^T\ {}_x\mathbf{e}^i$	$\vert{}_t\mathbf{e}^i\rangle = \mathcal{R}^T\ \vert{}_x\mathbf{e}^i\rangle$	i = 1,2..n	pull back	*(10.7.4)*	(E.10)

<u>Dual Pull Backs</u>

$({}_{\mathbf{x}}\mathbf{u}^i)^T = ({}_{\mathbf{t}}\mathbf{u}^i)^T R^T$ $\quad <{}_{\mathbf{x}}\mathbf{u}^i| = <{}_{\mathbf{t}}\mathbf{u}^i|\mathcal{R}^T$ $\quad$ i = 1,2..n $\quad$ push forward
$({}_{\mathbf{t}}\mathbf{u})^T = ({}_{\mathbf{x}}\mathbf{u})^T R$ $\quad <{}_{\mathbf{t}}\mathbf{u}^i| = <{}_{\mathbf{x}}\mathbf{u}^i|\mathcal{R}$ $\quad$ i = 1,2..n $\quad$ pull back $\quad$ *(10.7.6)* $\quad$ (E.11)

$({}_{\mathbf{x}}\mathbf{e}^i)^T = ({}_{\mathbf{t}}\mathbf{e}^i)^T R^T$ $\quad <{}_{\mathbf{x}}\mathbf{e}^i| = <{}_{\mathbf{t}}\mathbf{e}^i|\mathcal{R}^T$ $\quad$ i = 1,2..n $\quad$ push forward
$({}_{\mathbf{t}}\mathbf{e}^i)^T = ({}_{\mathbf{x}}\mathbf{e}^i)^T R$ $\quad <{}_{\mathbf{t}}\mathbf{e}^i| = <{}_{\mathbf{x}}\mathbf{e}^i|\mathcal{R}$ $\quad$ i = 1,2..n $\quad$ pull back $\quad$ *(10.7.6)* $\quad$ (E.12)

$\varphi^*(<{}_{\mathbf{x}}\mathbf{e}^i|) = \varphi^*({}_{\mathbf{x}}\lambda^i) = <{}_{\mathbf{x}}\mathbf{e}^i|\,\mathcal{R} = R^i{}_j <{}_{\mathbf{t}}\mathbf{u}^j| = R^i{}_j\ {}_{\mathbf{t}}\lambda^j$ $\quad$ *(10.7.19) 5*
$\varphi^*(<{}_{\mathbf{x}}\mathbf{u}^i|) = <{}_{\mathbf{x}}\mathbf{u}^i|\,\mathcal{R} = <{}_{\mathbf{t}}\mathbf{u}^i| = {}_{\mathbf{t}}\lambda^i$ $\quad$ *(E.11)* $\quad$ (E.13)



(10.9.2) $\quad$ (E.14)

Further translations of significant equations appear in Section 10.9.

Appendix F: The Volume of an n-piped embedded in R^m

First, recall these facts about the various basis vectors described in Chapter 2,

$(\mathbf{u}_j)^i = \delta_j{}^i$	axis-aligned basis vectors in x-space	// (2.5.3)
$(\mathbf{e'}_j)^i = \delta_j{}^i$	axis-aligned basis vectors in x'-space	// (2.5.3)
$(\mathbf{e}_j)^i = S^i{}_j$	tangent base vectors in x-space	// (2.3.4) and (2.1.4) $S^i{}_j = R_j{}^i$
$(\mathbf{u'}_j)^i = R^i{}_j$	inverse tangent base vectors in x'-space	// (2.5.2)

$\mathbf{e}_i = S\mathbf{e'}_i$. // (2.5.1) $d\mathbf{x} = S\, d\mathbf{x}'$

$\mathbf{u'}_i = R\mathbf{u}_i$ // (2.5.1) $d\mathbf{x'} = R\, d\mathbf{x}$. (F.1)

We claim that the volume of an m-piped in x-space = R^m spanned by the tangent base vectors $\mathbf{e}_1, \mathbf{e}_2....\mathbf{e}_m$ is given by

$$V = \det[\mathbf{e}_1, \mathbf{e}_2....\mathbf{e}_m] = \det(S) \ . \tag{F.2}$$

For the case n = 2 the fact that $V = \det[\mathbf{e}_1, \mathbf{e}_2]$ is easily shown, see text below (4.3.14).

Similarly, the volume of an m-piped in x'-space = R^m spanned by the inverse tangent base vectors $\mathbf{u'}_1$, $\mathbf{u'}_2....\mathbf{u'}_m$ is given by

$$V = \det[\mathbf{u'}_1, \mathbf{u'}_2....\mathbf{u'}_m] = \det(R) \ . \tag{F.3}$$

First of all, note from (F.1) that

$[\mathbf{e}_1, \mathbf{e}_2....\mathbf{e}_m] = S$ because $(\mathbf{e}_i)^b = S^b{}_c(\mathbf{e'}_i)^c = S^b{}_c\delta_i{}^c = S^b{}_i$, i = column index

$[\mathbf{u'}_1, \mathbf{u'}_2....\mathbf{u'}_m] = R$ because $(\mathbf{u'}_i)^b = R^b{}_c(\mathbf{u}_i)^c = R^b{}_c\delta_i{}^c = R^b{}_i$, i = column index (F.4)

Statements (F.2) and (F.3) are really the same fact stated first for the forward transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ with differential $R = S^{-1}$, and then for the inverse transformation $\mathbf{x} = F^{-1}(\mathbf{x}')$ with differential $S = R^{-1}$.

The fact (F.2) is derived in *Tensor* (with m = N) so we won't repeat that derivation here. One starts with an m-cube in x'-space which is spanned by axis-aligned unit basis vectors $\mathbf{e'}_i$ and which has volume V' = 1. One then transforms this m-cube into a skewed m-piped in x-space spanned by the tangent base vectors $\mathbf{e}_i = S\mathbf{e'}_i$ where S is the mxm matrix that maps the vectors $\mathbf{e'}_i$ into the $\mathbf{e}_i$. One then finds that the m-piped has the volume shown in (F.2). *Tensor* Appendix B concerns the geometry of N-pipeds in R^N and the result (F.2) appears as (B.5.d.13).

The different question addressed by our current Appendix F is the following:

What is the volume of an n-piped in R^m where n < m ?

The answer, to be proven below, is (new meanings for S and R) ,

$$V = \sqrt{\det(S^T S)} \quad \text{where } S = [\mathbf{e}_{i_1}, \mathbf{e}_{i_2} .... \mathbf{e}_{i_n}] \qquad \text{n-piped in x-space} = R^m \qquad (F.5)$$

$$V = \sqrt{\det(R^T R)} \quad \text{where } R = [\mathbf{u'}_{i_1}, \mathbf{u'}_{i_2} .... \mathbf{u'}_{i_n}] \qquad \text{n-piped in x'-space} = R^m \qquad (F.6)$$

where the subscripts i_r enumerate a *subset* of the tangent base vectors in each case which "span" the n-piped. For example, if in R^3 we have tangent base vectors ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) spanning a 3-piped, we know that any pair ($\mathbf{e}_{i_1}$, $\mathbf{e}_{i_2}$) with $i_1 \neq i_2$ spans a 2-piped which is a face of the 3-piped.

Notice that S and R appearing in (F.5) and (F.6) are non-square "tall" matrices because each has m rows but only n columns since each [...] has only n vectors. Thus det(S) and det(R) do not exist. In the special case that n = m, then S and R are square matrices, and for example $\sqrt{\det(S^T S)}$ = det(S), and the results (F.2) and (F.3) are recovered.

Again, (F.5) and (F.6) are really the same statement expressed for $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ with differential R and then for a transformation $\mathbf{x} = \mathbf{G}(\mathbf{x}')$ with differential S. We deal with $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ for a non-square "tall" R matrix in some detail in Section 10.6.

The result (F.5) is derived in *Tensor* [2016] Section 8.4 (h). Below, we shall prove (F.6) where the n-piped spanning vectors will be called $\mathbf{a}_1,\mathbf{a}_2,...\mathbf{a}_n$. In particular, we shall show that :

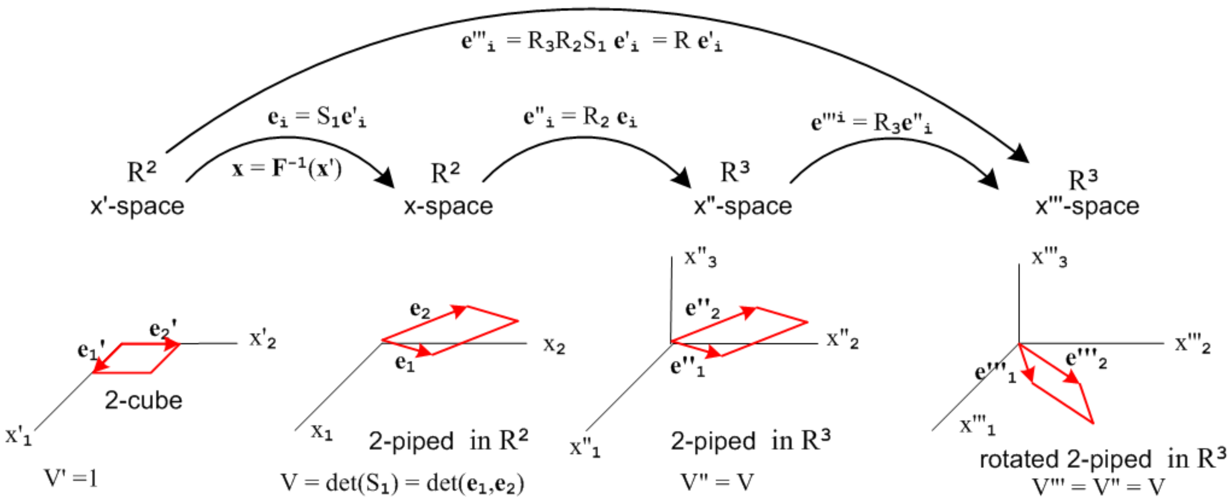
Fact : Let $\mathbf{u}_i$ for i = 1...n be n axis-aligned unit basis vectors in R^n. Let $\mathbf{a}_1,\mathbf{a}_2,...\mathbf{a}_n$ be n arbitrary vectors in R^m where $m \geq n$. These vectors span an n-piped in R^m which has a volume $V = \sqrt{\det(R^T R)}$ where $R^T R$ is a nxn matrix, $R = [\mathbf{a}_1,\mathbf{a}_2,...\mathbf{a}_n]$ is a m x n matrix, and the arbitrary n vectors may be written $\mathbf{a}_i = R\mathbf{u}_i$.
(F.4.11)

Our derivation below is a seat-of-the-pants "geometric" approach, appropriate for people like the author who like to "see" what is going on. Basically we start with simple examples and progress toward the general case. Sjamaar provides a nice "axiomatic" derivation on pp 99-102.

<u>Warning</u>: In the discussion below R_1, R_2 and R_3 are linear transformations, while R^2, R^3 are Cartesian spaces. It just happens that the same symbol R is used for both kinds of objects.

F.1 Volume of a 2-piped in R^3

The simplest case to consider is that of a 2-piped (a parallelogram) embedded in R^3. Consider then this set of four spaces connected by three transformations (this picture will be reused several times) :



(F.1.1)

The names and symbols for the left two spaces are chosen to be compatible with Chapter 2 based on Fig (2.1.1). The remaining two spaces are given the arbitrary names x"-space and x'''-space.

The leftmost picture shows a unit square (2-cube) lying in the x'_1-x'_2 plane. The square is spanned by two axis-aligned unit basis vectors $\mathbf{e'}_1$ and $\mathbf{e'}_2$ as described in (2.5.3). The volume (area) of the square is 1 unit.

The second picture is obtained from the first by a general non-linear transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x'})$ which has a linearized form $d\mathbf{x'} = R_1 d\mathbf{x}$ and correspondingly $d\mathbf{x} = S_1 d\mathbf{x'}$ where $R_1S_1 = 1$. These 2x2 R_1 and S_1 matrices are called R and S in Chapter 2, but here we add subscript 1 since this is the first of three transformations shown above. The volume of the 2-piped is $\det(S_1) = \det(\mathbf{e}_1,\mathbf{e}_2)$ as noted in (F.2). The transformed basis vectors ("tangent base vectors") are $\mathbf{e}_i = S_1\mathbf{e'}_i$ (and $\mathbf{e'}_i = R_1\mathbf{e}_i$) as shown in (2.5.1).

The third picture is obtained from the second merely by adding a third axis called $x"_3$. The 2-piped has not moved and still lies in the $x"_1$-$x"_2$ plane. Then $\mathbf{e"}_i = (\mathbf{e}_i,0)$ where we add a third zero component to the basis vectors. The 2-piped volume is still $\det(S_1)$. Below we shall discuss the linear transformation R_2 which links x-space to x"-space.

The fourth picture is obtained from the third by an arbitrary rotation R_3 in R^3 space. The 2-piped then ends up in some *arbitrary orientation* in R^3. It's volume is still $\det(S_1)$ since shape and volume (here area) are not changed by a rotation. R_3 is a 3x3 real orthogonal matrix.

The linear transformation R_2 connecting x"-space and x-space is this

$$R_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{for example:} \quad \mathbf{e"}_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (\mathbf{e}_i)^1 \\ (\mathbf{e}_i)^2 \end{pmatrix} = \begin{pmatrix} (\mathbf{e}_i)^1 \\ (\mathbf{e}_i)^2 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{e}_i \\ 0 \end{pmatrix}. \qquad \text{(F.1.2)}$$

This matrix R_2 is just the 2x2 identity matrix with a null third row added. This transformation simply adds a third null coordinate to a 2D vector in R^2 as shown in the example above. Notice that

$$R_2^T R_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \quad . \qquad \text{(F.1.3)}$$

As shown at the top of Fig (F.1.1), the concatenated effect of the three transformations on the basis vectors is this,

$$\mathbf{e}'''_i = R\,\mathbf{e}'_i \qquad \text{where } R \equiv R_3R_2S_1 \,. \qquad i = 1,2 \qquad \text{(F.1.4)}$$

The combined matrix R is in fact a "tall" 3 x 2 R matrix which we verify with the following schematic conformation picture,

$$R = R_3R_2S_1 = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \\ * & * \end{pmatrix} \begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \\ * & * \end{pmatrix} . \qquad \text{(F.1.5)}$$

In fact, from (F.1.4) and (2.5.2) we find that

$$(\mathbf{e}'''_i)^a = R^a{}_b\,(\mathbf{e}'_i)^b = R^a{}_b\,\delta_i{}^b = R^a{}_i \qquad i = 1,2 \quad a = 1,2 \qquad \text{(F.1.6)}$$

which tells us that the final vectors $\mathbf{e}'''_i$ are the columns of the matrix R, so

$$R = [\mathbf{e}'''_1, \mathbf{e}'''_2] \qquad \text{(F.1.7)}$$

which we then verify is a matrix with 3 rows and 2 rows as shown at the right of (F.1.5).

Note that non-square R does not have a determinant. But consider,

$$\begin{aligned} R^TR &= (R_3R_2S_1)^T R_3R_2S_1 \\ &= S_1^T R_2^T R_3^T R_3R_2S_1 && \text{// rule for transpose of a product of matrices} \\ &= S_1^T R_2^T R_2S_1 && \text{// } R_3^TR_3 = 1 \text{ because } R_3 \text{ is a rotation (real-orthogonal)} \\ &= S_1^T S_1 \,. && \text{// } R_2^TR_2 = 1 \text{ as just shown in (F.1.3)} \qquad \text{(F.1.8)} \end{aligned}$$

As shown in Fig (10.6.c.1), the matrix R^TR is a square matrix of dimension 2x2 and so R^TR *does* have a determinant. In fact, from (F.1.8),

$$\det(R^TR) = \det(S_1^TS_1) = \det(S_1^T)\det(S_1) = \det^2(S_1) \,. \qquad \text{(F.1.9)}$$

In terms of the overall vector transformation R going from the left picture to the right picture above, we have just shown that the 2-piped volume can be written

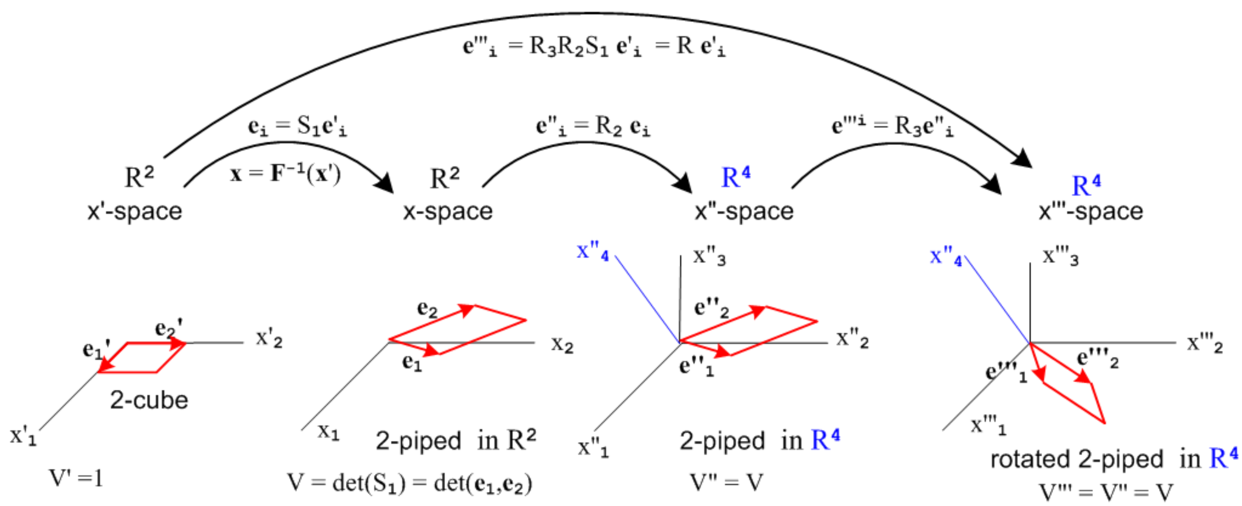
$$V = \det(S_1) = \sqrt{\det(R^TR)} \quad \text{where } R = [\mathbf{e}'''_1, \mathbf{e}'''_2] \qquad \text{(F.1.10)}$$

where both S_1 and R^TR are 2x2 matrices.

F.2 Volume of a 2-piped in R^4, R^5 and R^m

The equation numbers here mimic those of the previous section. Since some equations need not be repeated, there are missing equation numbers below.

In blue we make very slight modifications to the previous picture :



(F.2.1)

R_3 is now a real-orthogonal 4x4 rotation matrix in R^4. The new R_2 transformation has two rows of zeros added at the bottom instead of one row,

$$R_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for example: } \mathbf{e}''_i = R_2\mathbf{e}_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} (\mathbf{e}_i)^1 \\ (\mathbf{e}_i)^2 \end{pmatrix} = \begin{pmatrix} (\mathbf{e}_i)^1 \\ (\mathbf{e}_i)^2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{e}_i \\ 0 \\ 0 \end{pmatrix}. \tag{F.2.2}$$

This transformation R_2 simply adds a third and fourth null coordinate to a 2D vector in R^2 as shown in the example above. And as before,

$$R_2^T R_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \ . \tag{F.2.3}$$

The new conformation picture is this

$$R = R_3R_2S_1 = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \\ * & * \\ * & * \end{pmatrix}\begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \\ * & * \\ * & * \end{pmatrix} \tag{F.2.5}$$

and now the tall R matrix has 4 rows and 2 columns. Equation (F.1.6) still applies for this new R, so we still conclude that

$$R = [\mathbf{e}'''_1, \mathbf{e}'''_2] \tag{F.2.7}$$

but now each vector has 4 components instead of 3 as in (F.1.7).

Apart from these matrix shape changes, the steps (F.1.8) through (F.1.10) proceed exactly as above and again one concludes that

$$V = \det(S_1) = \sqrt{\det(R^T R)} \quad \text{where } R = [\mathbf{e}'''_1, \mathbf{e}'''_2] \tag{F.2.10}$$

where both S_1 and R^TR are 2x2 matrices.

In going from R^4 to R^5 the reader can see that R_2 will acquire yet another null row, one still has $R_2^TR_2$ = 1, and everything goes through as before again giving $V = \sqrt{\det(R^TR)}$ where $R = [\mathbf{e}'''_1, \mathbf{e}'''_2]$ now has 5 rows since the two vectors exist in R^5. The result clearly extends to a 2-piped in R^m for any $m \geq 2$.

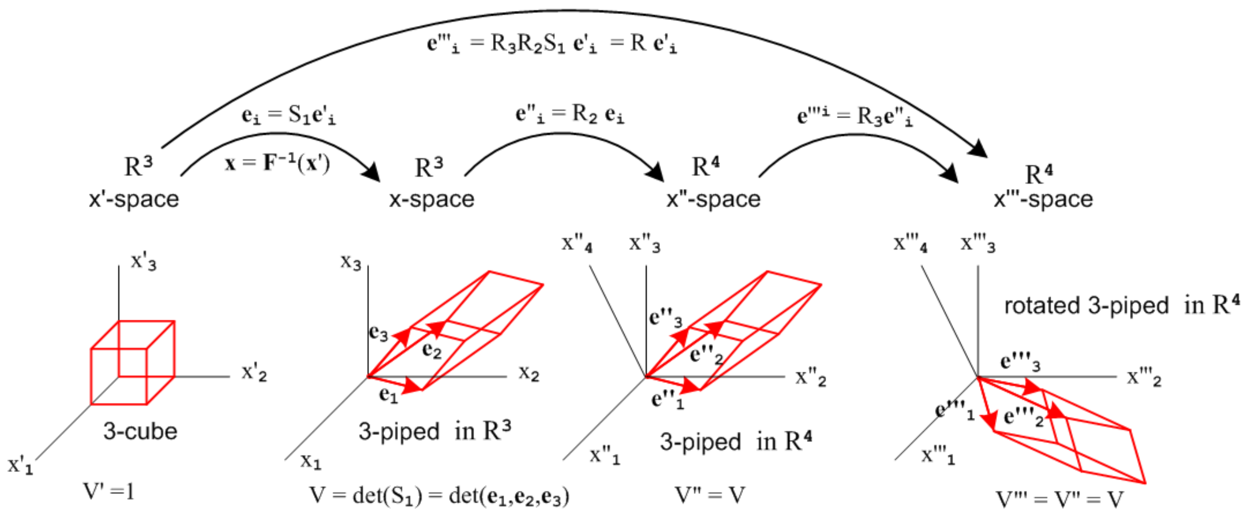
We then arrive at the following Fact, where we rename $\mathbf{e}'_i \rightarrow \mathbf{u}_i$ and $\mathbf{e}'''_{1,2} \rightarrow \mathbf{a},\mathbf{b}$:

Fact : Let $\mathbf{u}_1 = (1,0)$ and $\mathbf{u}_2 = (0,1)$ be two axis-aligned basis vectors in R^2. Let **a** and **b** be two arbitrary vectors in R^m. These vectors span a 2-piped in R^m which has a volume (area in this case) $V = \sqrt{\det(R^TR)}$ where R^TR is a 2x2 matrix, $R = [\mathbf{a},\mathbf{b}]$ is an m x 2 matrix, and the arbitrary two vectors may be written $\mathbf{a} = R\mathbf{u}_1$ and $\mathbf{b} = R\mathbf{u}_2$. (F.2.11)

Comment: Recall from (10.6.c.6) that $\text{rank}(R^TR) = \text{rank}(R)$. If **a** and **b** are linearly dependent, $R = [\mathbf{a},\mathbf{b}]$ has less than full rank 2 and so does square R^TR which means $\det(R^TR) = 0$ so $V = 0$. This is the result one would expect if **a** and **b** are collinear, so the above Fact applies to any pair of vectors **a**, **b**.

F.3 Volume of a 3-piped in R^4

The logic flows as in the previous examples, so we omit the words. We start with a new but similar transformation picture,



(F.3.1)

The R_2 matrix (3x3 identity with a null added 4th row) adds a null 4th component to any 3-vector,

$$R_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{for example: } \mathbf{e_i}" = R_2\mathbf{e_i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} (\mathbf{e_i})^1 \\ (\mathbf{e_i})^2 \\ (\mathbf{e_i})^3 \end{pmatrix} = \begin{pmatrix} (\mathbf{e_i})^1 \\ (\mathbf{e_i})^2 \\ (\mathbf{e_i})^3 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{e_i} \\ 0 \end{pmatrix}. \tag{F.3.2}$$

$$R_2^T R_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \ . \tag{F.3.3}$$

$$\mathbf{e'''_i} = R\, \mathbf{e'_i} \qquad \text{where } R \equiv R_3R_2S_1 \ . \qquad i = 1,2,3 \tag{F.3.4}$$

$$R = R_3R_2S_1 = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \tag{F.3.5}$$

$$(\mathbf{e'''_i})^a = R^a{}_b\, (\mathbf{e'_i})^b = R^a{}_b\, \delta_i{}^b = R^a{}_i \qquad i = 1,2,3 \quad a = 1,2,3 \tag{F.3.6}$$

$$R = [\mathbf{e'''_1}, \mathbf{e'''_2}, \mathbf{e'''_3}] \qquad = \text{a 4 x 3 matrix} \tag{F.3.7}$$

$$R^TR = S_1^TS_1 \ = \text{a 3x3 matrix} \tag{F.3.8}$$

$$\det(R^TR) = \det{}^2(S_1) \tag{F.3.9}$$

$$V = \det(S_1) = \sqrt{\det(R^TR)} \quad \text{where } R = [\mathbf{e'''_1}, \mathbf{e'''_2}, \mathbf{e'''_3}] \ . \tag{F.3.10}$$

Fact : Let $\mathbf{u_1}$ = (1,0,0), $\mathbf{u_2}$ = (0,1,0) and $\mathbf{u_3}$ = (0,0,1) be three axis-aligned basis vectors in R^3. Let **a**,**b**,**c** be three arbitrary vectors in R^4. These vectors span a 3-piped in R^4 which has a volume $V = \sqrt{\det(R^TR)}$ where R^TR is a 3x3 matrix, R = [**a**,**b**,**c**] is a 4 x 3 matrix, and the arbitrary three vectors may be written **a** = R$\mathbf{u_1}$, **b** = R$\mathbf{u_2}$ and **c** = R$\mathbf{u_3}$. (F.3.11)

<u>Comment</u>: Recall from (10.6.c.6) that rank(R^TR) = rank(R). If **a**,**b**,**c** are linearly dependent, R = [**a**,**b**,**c**] has less than full rank 3 and so does square R^TR which means $\det(R^TR) = 0$ so V = 0. This is the result one would expect if **a**,**b**,**c** are linearly dependent: they all lie in the same plane and thus span no 3D volume. Thus, the above Fact applies to any triplet of vectors **a**, **b, c**.

F.4 Volume of a n-piped in R^m

In the general case we have an n-piped in R^m. The method outlined in the previous examples prevails with generalizations for the objects involved. Again we omit the words.

Picture generously supplied by the reader capable of making hyperspace drawings (F.4.1)
[This is one reason mathematicians don't like geometric derivations!]

$$R_2 = \begin{pmatrix} 1_{nxn} \\ \text{m-n rows of zeros} \end{pmatrix} \quad \text{which is an m x n tall R matrix} \qquad (F.4.2)$$

$$R_2^T R_2 = 1_{nxn} \qquad (F.4.3)$$

$$\mathbf{e}'''_i = R\,\mathbf{e}'_i \qquad \text{where } R \equiv R_3R_2S_1 . \qquad i = 1,2..,n \qquad (F.4.4)$$

$$R = R_3R_2S_1 = (\text{m x m rotation matrix } R_3)\ (\text{m x n } R_2 \text{ matrix})\ (\text{n x n matrix } S_1) = \text{m x n} \qquad (F.4.5)$$

$$(\mathbf{e}'''_i)^a = R^a{}_b\,(\mathbf{e}'_i)^b = R^a{}_b\,\delta_i{}^b = R^a{}_i \qquad i = 1,2...n \quad a = 1,2...n \qquad (F.4.6)$$

$$R = [\mathbf{e}'''_1, \mathbf{e}'''_2 \ .... \ \mathbf{e}'''_n] \qquad = \text{ m x n tall R matrix} \qquad (F.4.7)$$

$$R^TR = S_1^TS_1 = \text{an nxn matrix} \qquad (F.4.8)$$

$$\det(R^TR) = \det{}^2(S_1) \qquad (F.4.9)$$

$$V = \det(S_1) = \sqrt{\det(R^TR)} \quad \text{where } R = [\mathbf{e}'''_1, \mathbf{e}'''_2 \ .... \ \mathbf{e}'''_n] \ . \qquad (F.4.10)$$

Fact : Let $\mathbf{u_i}$ for i = 1...n be n axis-aligned unit basis vectors in R^n. Let $\mathbf{a_1}$,$\mathbf{a_2}$,...$\mathbf{a_n}$ be n arbitrary vectors in R^m where $m \geq n$. These vectors span an n-piped in R^m which has a volume $V = \sqrt{\det(R^TR)}$ where R^TR is a nxn matrix, R = [$\mathbf{a_1}$,$\mathbf{a_2}$,...$\mathbf{a_n}$] is a m x n matrix, and the arbitrary n vectors may be written $\mathbf{a_i}$ = R$\mathbf{u_i}$.
(F.4.11)

Comment 1: Recall from (10.6.c.6) that rank(R^TR) = rank(R). If the $\mathbf{a}_i$ are linearly dependent, R = $[\mathbf{a}_1,\mathbf{a}_2,...\mathbf{a}_n]$ has less than full rank n and so does square R^TR which means det(R^TR) = 0 so V = 0. This is the result one would expect if $\mathbf{a}_i$ are linearly dependent: they span no nD volume. Thus, the above Fact applies to any set of n vectors $\mathbf{a}_i$.

Comment 2: The discussion above is presented for n < m. In the case n = m, things simplify. Looking at Fig (F.3.1) we can ignore the x"-space and x'"-space pictures and in effect set $R_2 = R_3 = 1$ so that $R = S_1$. The final vectors are the $\mathbf{e}_i$ in x-space which we then take to be are arbitrary vectors $\mathbf{a}_i$. The general formula still works, but now the R matrix is square, so

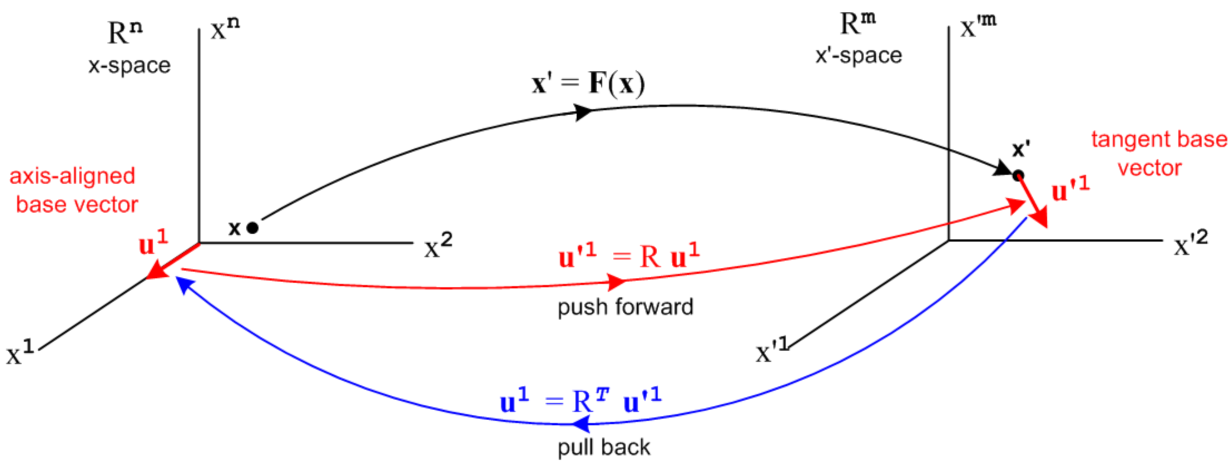
$$V = \sqrt{\det(R^TR)} = \sqrt{\det(S_1{}^TS_1)} = \sqrt{\det(S_1{}^T)\det(S_1)} = \sqrt{\det(S_1)\det(S_1)}$$

$$= \det(S_1) \quad \text{where } S_1 = [\mathbf{e}_1,\mathbf{e}_2,...\mathbf{e}_m] = R = [\mathbf{a}_1,\mathbf{a}_2,...\mathbf{a}_n]$$

in agreement with (F.2).

F.5 Application: The differential volume element of the tangent space $T_{x'}M$

Recall Fig (10.7.5) which shows how the axis-aligned basis vectors $\mathbf{u}_i$ of x-space are pushed forward to become the tangent base vectors $\mathbf{u}'_i$ which span the tangent space $T_{x'}M$ at point $\mathbf{x}'$ in x'-space,



(10.7.5)

If we take a *differential* n-cube located at position **x** in x-space, it has differential volume

$$dV = dx^1dx^2.... dx^n \quad . \tag{F.5.1}$$

This volume is mapped into an n-piped in x'-space by the mapping $\mathbf{u}'^i = R\,\mathbf{u}_i$. According to Fact (F.4.11), we may conclude that the volume of the tangent space n-piped in x'-space at point **x'** is this,

$$dV' = \sqrt{\det(R^TR)}\, dx^1dx^2.... dx^n \, . \tag{F.5.2}$$

We have already seen an example of this fact. Recall these equations from Section 10.10,

$$A' = \int_{S'} dA' = \int_S K(\mathbf{x})\, dx^1 dx^2 \,. \tag{10.10.21}$$

$$K^2 = \det(R^T R) \tag{10.10.22}$$

In the area integral, the differential "volume" is $dA' = K(\mathbf{x})\, dx^1 dx^2 = \sqrt{\det(R^T R)}\, dx^1 dx^2$.

Appendix G : The det($R^T R$) theorem and its relation to differential forms

G.1 Theorem: det($R^T R$) is the sum of the squares of the full-width minors of R

We saw and verified an example of this theorem in Section 10.10 for a 3 x 2 R matrix,

$$K^2 = \det(R^T R) \tag{10.10.22}$$

$$K^2 = \det{}^2\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} + \det{}^2\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} + \det{}^2\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} . \qquad (10.10.18)'$$

Before proving the theorem, we have Maple test it for a messy case, just to make sure it is true. Enter a generic 6x4 R matrix as follows:

```
m := 6:   n := 4:
R := matrix(m,n): print(R);
```

$$\begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} & R_{1,4} \\ R_{2,1} & R_{2,2} & R_{2,3} & R_{2,4} \\ R_{3,1} & R_{3,2} & R_{3,3} & R_{3,4} \\ R_{4,1} & R_{4,2} & R_{4,3} & R_{4,4} \\ R_{5,1} & R_{5,2} & R_{5,3} & R_{5,4} \\ R_{6,1} & R_{6,2} & R_{6,3} & R_{6,4} \end{bmatrix}$$

Compute and accumulate into "acc" the squares of all (6,4) = 15 full-width minors,

```
L := [seq(k,k=1..m)];
```

$$L := [1, 2, 3, 4, 5, 6]$$

```
acc := 0:   N := 0:
# compute all full-width minors and accumulate into "acc" the sum of their squares
for i from 1 to m-1 do
   for j from i+1 to m do    # below knock out rows i and j
      minor[i,j] := det(submatrix(R,subsop(i=NULL,j=NULL,L), 1..n));
      acc := acc + (minor[i,j])^2;   N := N+1;
   od;
od;
N;
```

$$15$$

If we take the resulting "acc" and expand it, we get a series of 4,230 terms each of which contains a product of eight matrix elements of R,

```
nops(expand(acc));   # terms in the sum of squares of the minors
```

$$4230$$

Here are four of these terms

$$+2R_{4,1}^2R_{5,2}R_{3,3}R_{6,4}R_{6,2}R_{3,4}R_{5,3}+2R_{4,1}R_{5,2}R_{3,3}R_{6,4}^2R_{5,1}R_{3,2}R_{4,3}$$
$$-2R_{4,1}R_{5,2}R_{3,3}R_{6,4}R_{5,1}R_{3,2}R_{4,4}R_{6,3}-2R_{4,1}R_{5,2}R_{3,3}^2R_{6,4}^2R_{5,1}R_{4,2}\;.$$

We next compute det(R^TR), note that it also has 4,230 terms, and then we show that det(R^TR) = acc.

```
detRTR := det(transpose(R) &* R):
nops(detRTR);
                                        4230
detRTR - acc: simplify(%);
                                          0
```

Fortified with the knowledge that the theorem seems to be true, we proceed:

Theorem: Let R be an m x n matrix with m ≥ n. There are (m,n) full-width minors.
The claim is that det(R^TR) = sum of the squares of the full-width minors. (G.1.1)

Comment: For m = n there is only one minor for square R which is det(R), the sum of the squares of the minors is then just $\det^2$(R), and indeed det(R^TR) = $\det^2$(R).

Proof for n < m :

Define a multiindex I as follows

$$I = i_1, i_2, ... i_n \tag{G.1.2}$$

where the i_r indicate which n rows of the R matrix are included in a certain minor. Each i_r takes values in the range 1 to m since R has m rows. Denote a full-width minor of R by

$$\text{minor}_I\;. \tag{G.1.3}$$

We shall need the following

Lemma:

$$\Sigma'_I\,[\text{minor}_I]^2 \;=\; (1/n!)\,\Sigma_I\,[\text{minor}_I]^2 \tag{G.1.4}$$

where

$$\Sigma'_I = \Sigma_{1\le i_1<i_2<....<i_n\le m} \qquad = \text{ordered sum}$$
$$\Sigma_I = \Sigma_{i_1,i_2,...i_n=1}^{m} \qquad = \text{symmetric sum}\;. \tag{G.1.5}$$

Proof of Lemma:

$$\Sigma_I [\text{minor}_I]^2 = \Sigma_{i_1,i_2,\ldots i_n=1}{}^m [\text{minor}_I]^2$$

$$= \Sigma_{i_1\neq i_2\neq..\neq i_n} [\text{minor}_I]^2 \qquad // \text{ minor}_I = 0 \text{ if two rows are the same}$$

$$= (\Sigma_{i_1<i_2<....<i_n} + \Sigma_{i_2<i_1<....<i_n} + \text{ n!-2 other orderings}) [\text{minor}_I]^2$$

$$= (\Sigma_P [\Sigma_{P(i_1)<P(i_2)<...<P(i_n)}]) [\text{minor}_I]^2$$

$$= \Sigma_{i_1<i_2<...<i_n} [\Sigma_P f_{P(i_1)P(i_2)...P(i_n)}] . \qquad // \text{ by (A.9.1) with } f_{i_1i_2...i_n} = [\text{minor}_I]^2$$

But $f_{i_1i_2...i_n} = [\text{minor}_I]^2$ is a totally symmetric function of the indices since row swaps don't affect a squared determinant. Thus we continue the above to get

$$= \Sigma_{i_1<i_2<...<i_n} [\Sigma_P f_{i_1i_2...i_n}] = \Sigma_{i_1<i_2<...<i_k} f_{i_1i_2...i_n} [\Sigma_P 1]$$

$$= \Sigma_{i_1<i_2<...<i_n} f_{i_1i_2...i_n} [n!] = \Sigma_{i_1<i_2<...<i_k} [\text{minor}_I]^2 [n!]$$

$$= n! \Sigma'_I [\text{minor}_I]^2 \qquad \text{QED Lemma}$$

Proof of Theorem:

We can write minor_I as the determinant of a matrix using (A.1.19)

$$\text{minor}_I = \Sigma_P(-1)^{S(P)} R^{i_1}{}_{P(1)} R^{i_2}{}_{P(2)} \ldots R^{i_n}{}_{P(n)}$$

$$= R^{i_1}{}_1 R^{i_2}{}_2 \ldots R^{i_n}{}_n + \text{all signed permutations}$$

$$\equiv \Sigma_P(-1)^{S(P)} R^I{}_{P(Z)} . \qquad // \text{ in multiindex notation, } Z = 1,2....n \qquad \text{(G.1.6)}$$

Then we first claim that

$$\text{Sum} \equiv \text{ sum of all full-width squared minors } = \Sigma'_I [\text{minor}_I]^2 . \qquad \text{(G.1.7)}$$

In this ordered sum, each full-width minor of R is included exactly once. For example, for a 3x2 R matrix we had above

$$K^2 = \underset{i_1,i_2=1,2}{\det{}^2\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix}} + \underset{i_1,i_2=1,3}{\det{}^2\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}} + \underset{i_1,i_2=2,3}{\det{}^2\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}} . \qquad \textit{(10.10.18)'}$$

Then using the above Lemma we write

$$\text{Sum} = \Sigma'_I [\text{minor}_I]^2 = (1/n!)\, \Sigma_I [\text{minor}_I]^2 = (1/n!)\, \Sigma_I [\text{minor}_I]\, [\text{minor}_I] \quad \text{// now use (G.1.6),}$$

$$= (1/n!)\, \Sigma_I\, [\, \Sigma_P(-1)^{S(P)}\, R^{i_1}{}_{P(1)}\, R^{i_2}{}_{P(2)} \ldots R^{i_n}{}_{P(n)}]\, [\, \Sigma_{P'}(-1)^{S(P')}\, R^{i_1}{}_{P'(1)}\, R^{i_2}{}_{P'(2)} .. R^{i_n}{}_{P'(n)}]$$

$$= (1/n!)\, \Sigma_I\, [\, \Sigma_P(-1)^{S(P)} R^I{}_{P(Z)}]\, [\, \Sigma_{P'}(-1)^{S(P')} R^I{}_{P'(Z)}] \qquad \text{// multiindex notation}$$

$$= (1/n!)\, \Sigma_P(-1)^{S(P)}\, \Sigma_{P'}(-1)^{S(P')} \Sigma_I\, R^I{}_{P(Z)} R^I{}_{P'(Z)} \qquad \text{// reorder}$$

$$= (1/n!)\, \Sigma_P(-1)^{S(P)}\, \Sigma_{P'}(-1)^{S(P')} \Sigma_I\, (R^T)^{P(Z)}{}_I R^I{}_{P'(Z)} \qquad \text{// matrix transposes}$$

$$= (1/n!)\, \Sigma_P(-1)^{S(P)}\, \Sigma_{P'}(-1)^{S(P')} (R^TR)^{P(Z)}{}_{P'(Z)} \qquad \text{// n matrix multiplications use up } \Sigma_I$$

$$= (1/n!)\, \Sigma_P(-1)^{S(P)}\, \Sigma_{P'}(-1)^{S(QP')} (R^TR)^{P(Z)}{}_{QP'(Z)} \qquad \text{// } \Sigma_{P'} \text{ rearrangement theorem (A.1.3)}$$

$$= (1/n!)\, \Sigma_P(-1)^{S(P)}\, \Sigma_{P'}(-1)^{S(PP')} (R^TR)^{P(Z)}{}_{PP'(Z)} \qquad \text{// select } Q = P$$

$$= (1/n!)\, \Sigma_P(-1)^{S(P)}\, \Sigma_{P'}(-1)^{S(PP')} (R^TR)^Z{}_{P'(Z)} \qquad \text{// (A.8.32) since factored form}$$

$$= (1/n!)\, \Sigma_P\, \Sigma_{P'}(-1)^{S(P')} (R^TR)^Z{}_{P'(Z)} \qquad \text{// (A.1.11)}$$

$$= (1/n!)\, \Sigma_{P'}(-1)^{S(P')} (R^TR)^Z{}_{P'(Z)} [\Sigma_P\, 1] \quad = (1/n!)\, \Sigma_{P'}(-1)^{S(P')} (R^TR)^Z{}_{P'(Z)} [n!]$$

$$= \Sigma_{P'}(-1)^{S(P')} (R^TR)^Z{}_{P'(Z)}$$

$$= \Sigma_P(-1)^{S(P)} (R^TR)^Z{}_{P(Z)}$$

$$= \det(R^TR)\ . \qquad \text{// (A.1.19)} \qquad \text{QED Theorem} \qquad \text{(G.1.8)}$$

G.2 The Connection between Theorem G.1.1 and Differential Forms

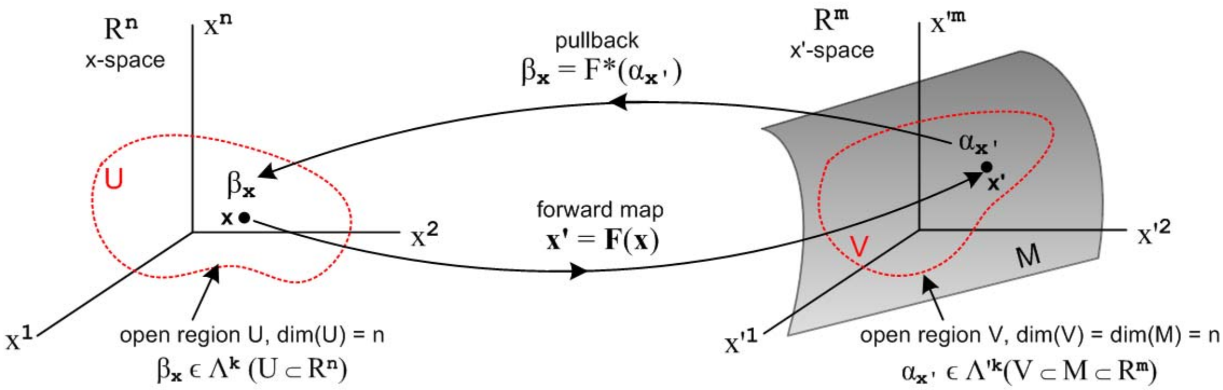
Recall from (10.11.7) that the integral of a k-form $\alpha_{\mathbf{x}'}$ for $F: R^n \rightarrow R^m$ is given by

$$\alpha_{\mathbf{x}'} = \Sigma'_I\, f_I(\mathbf{x}')\, dx'^{\wedge I} \qquad \text{(G.2.1)}$$

$$\int_{S'} \alpha_{\mathbf{x}'} = \int_S \Sigma'_J\, \Sigma'_I\, f_I(\mathbf{F}(\mathbf{x}))\, \det(R^I{}_J)\, dx^{j_1} \wedge dx^{j_2} \wedge \ldots \wedge dx^{j_k}$$

$$= \int_S \Sigma'_I\, f_I(\mathbf{F}(\mathbf{x}))\, \Sigma'_J \det(R^I{}_J)\, dx^{\wedge J}\ . \qquad \text{// } R = (DF) \qquad \text{(G.2.2)}$$

The context here is that $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is a point on a manifold M created by the mapping $F: R^n \rightarrow R^m$ as illustrated for example in (10.7.26) which we replicate here



(10.7.26)

A case of frequent interest is k = n, and in this case there is only one term in the ordered Σ'_J sum, so

$$\int_{S'} \alpha_{x'} = \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x})) \det(R^I{}_Z)\, dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n$$

$$= \int_S \Sigma'_I f_I(\mathbf{F}(\mathbf{x})) \det(R^I{}_Z)\, dx^{\wedge Z} \quad . \qquad Z \equiv 1,2...n \qquad // \ dx^{\wedge Z} = dV \qquad \text{(G.2.3)}$$

The object $\det(R^I{}_Z)$ is a full-width minor (a number) of the m x n R matrix. In (G.1.3) we called this minor_I so

$$\det(R^I{}_Z) = \text{minor}_I \equiv m_I \qquad \text{(G.2.4)}$$

where m_I is a compact notation for minor_I. Since $R^i{}_j(\mathbf{x})$ is generally a function of $\mathbf{x}$ (or $\mathbf{x}' = \mathbf{F}(\mathbf{x})$), we may regard $m_I = m_I(\mathbf{x}')$, so this minor's value is a function of $\mathbf{x}'$ on manifold M.

Then, suppressing the arguments f_I and m_I, (G.2.3) becomes

$$\int_{S'} \alpha_{x'} = \int_S [\, \Sigma'_I \ f_I \ m_I \,]\, dV \quad . \qquad \text{(G.2.5)}$$

We may consider f_I and m_I to be vectors with (m,n) components which we can dot together to get,

$$\int_{S'} \alpha_{x'} = \int_S [\, \mathbf{f} \bullet \mathbf{m} \,]\, dV \, . \qquad // \ \mathbf{f} \bullet \mathbf{m} \equiv \Sigma'_I \ f_I \ m_I \qquad \text{(G.2.6)}$$

For example, for n = 2, m = 3 we would write, showing components of each vector in "standard order",

$$\mathbf{f} \bullet \mathbf{m} = (f_{12}, f_{13}, f_{23}) \bullet (m_{12}, m_{13}, m_{23}) = f_{12}m_{12} + f_{13}m_{13} + f_{23}m_{23} = \Sigma'_I \ f_I \ m_I \, . \qquad \text{(G.2.7)}$$

One could create a minor unit vector $\hat{\mathbf{m}}$ in this manner

$$\hat{\mathbf{m}} \equiv \frac{\mathbf{m}}{|\mathbf{m}|} \tag{G.2.8}$$

where

$$|\mathbf{m}|^2 = \Sigma'_I (m_I)^2 = \Sigma'_I (\text{minor}_I)^2 = \det(R^TR) \qquad // \text{(G.1.1)} \tag{G.2.9}$$

where we just invoked the det(R^TR) theorem of Section G.1. One then has,

$$\int_{S'} \alpha_{x'} = \int_S [\ \mathbf{f} \bullet \hat{\mathbf{m}}\ ]\, |\mathbf{m}|\ \mathit{dV}$$

$$= \int_S [\ \mathbf{f} \bullet \hat{\mathbf{m}}\ ]\, [\sqrt{\det(R^TR)}\ \mathit{dV}\ ]\ . \tag{G.2.10}$$

The objects in red above are all functionals in the space $\Lambda^n(R^n)$ in our "cosmetic notation". According to the "second definition" described in (10.1.3) we are always allowed to replace the functional dV by dV to obtain a normal calculus integral that can be evaluated by standard methods,

$$\mathit{dV} \equiv \mathit{dx}^1 \wedge \mathit{dx}^2 \wedge ... \wedge \mathit{dx}^n \quad \rightarrow \quad dV \equiv dx^1 dx^2 .....dx^n\ . \tag{G.2.11}$$

Now consider the special case where the function $\mathbf{f} = \hat{\mathbf{m}}$. Then,

$$\int_{S'} \alpha_{x'} = \int_S [\ \hat{\mathbf{m}} \bullet \hat{\mathbf{m}}\ ]\, [\sqrt{\det(R^TR)}\ \mathit{dV}\ ] = \int_S \sqrt{\det(R^TR)}\ \mathit{dV}$$

$$= \int_S \sqrt{\det(R^TR)}\ dV \tag{G.2.12}$$

which displays the tangent space volume measure $dV' = \sqrt{\det(R^TR)}\, dx^1 dx^2 .... dx^n$ shown in (F.5.2). This last integral gives the volume (area) of the surface S' (example below) and for this reason we refer to the n-form α'_x with this value of **f** as the volume measure form μ'. From (G.2.1) then,

$$\mu' = \Sigma'_I f_I\, \mathit{dx'}^{\wedge I} = \mathbf{f} \bullet \mathbf{dx}^{\wedge\prime} = \hat{\mathbf{m}} \bullet \mathbf{dx}^{\wedge\prime}\ . \tag{G.2.13}$$

Here we use the same dot product idea as in (G.2.6), where f_I and $\mathit{dx'}^{\wedge I}$ are each treated as vectors having (m,n) components.

The pullback of this measure form appears in the integrand of (G.2.10),

$$F^*(\mu') = \sqrt{\det(R^TR)}\ \mathit{dV}$$

$$= \sqrt{\det((DF)^T(DF))}\ \mathit{dx}^1 \wedge \mathit{dx}^2 \wedge ... \wedge \mathit{dx}^n \qquad // \text{Sjamaar p 105 above item 8.12}\ . \tag{G.2.14}$$

Consider now a different situation where

$$\mathbf{f}(x') = |\mathbf{f}(x')|\,\hat{\mathbf{m}} \tag{G.2.15}$$

so that our new function **f**, considered as a vector with those (m,n) components, points in the $\hat{\mathbf{m}}$ direction. In this case we get

$$\int_{S'} \alpha_{x'} = \int_S [\,\mathbf{f} \bullet \hat{\mathbf{m}}\,]\,[\sqrt{\det(R^TR)}\ dV] = \int_S |\,\mathbf{f}((\mathbf{F}(\mathbf{x}))|\sqrt{\det(R^TR)}\,dV \tag{G.2.16}$$

and this is how one treats the integration of a scalar function over the surface S' such as the average temperature calculation in (10.10.20). In this case we can write the corresponding n-form in x'-space as

$$\alpha_{x'} = \Sigma'_I f_I\ dx'^{\wedge I} = \mathbf{f} \bullet \mathbf{dx}_{\wedge}' = |\,\mathbf{f}\,|\,\hat{\mathbf{m}} \bullet \mathbf{dx}_{\wedge}' = (\mathbf{f} \bullet \hat{\mathbf{m}})\,\hat{\mathbf{m}} \bullet \mathbf{dx}_{\wedge}' = (\mathbf{f} \bullet \hat{\mathbf{m}})\,\mu' \tag{G.2.17}$$

from which we extract this equation

$$\mathbf{f} \bullet \mathbf{dx}_{\wedge}' = (\mathbf{f} \bullet \hat{\mathbf{m}})\,\mu' \ . \tag{G.2.18}$$

Special case: Hypersurface where n = m - 1

Suppose now that n = m-1, so that the manifold M embedded in R^m is a hypersurface, meaning it has dimension one less than R^m. In this case, (m,n) = (m,m-1) = m and each of our "vectors" above has exactly m components.

One can define the following Hodge star objects as discussed at the start of Section 10.3 (see (H.1.13)),

$$*dx'^i = (-1)^{i-1}\ dx'^1 \wedge dx'^2\ ..\ [dx'^i] ... \wedge dx'^m \qquad \text{where } dx'^i \text{ is missing} \tag{G.2.19}$$

so that each $*dx'^i$ object is an (m-1)-form (that is, an n-form). These m Hodge dual objects can be combined to form a vector

$$*\mathbf{dx}' \equiv (*dx'^1, *dx'^2, ...*dx'^m)\ . \tag{G.2.20}$$

We now define new vectors **F** and **n'** as follows,

$$\begin{aligned} F_i &\equiv (-1)^{i-1} f_{12.[i]..m} \\ n'_i &\equiv (-1)^{i-1} m_{12.[i]..m} \end{aligned} \tag{G.2.21}$$

where the notation [i] means that index i is missing from 12...m. It follows that

$$\mathbf{F} \bullet *\mathbf{dx}' = \mathbf{f} \bullet \mathbf{dx}_{\wedge}' \tag{G.2.22}$$

$$\mathbf{n}' \bullet *\mathbf{dx}' = \mathbf{m} \bullet \mathbf{dx}_{\wedge}' \tag{G.2.23}$$

$$\mathbf{n}' \bullet \mathbf{F} = \mathbf{m} \bullet \mathbf{f}\ . \tag{G.2.24}$$

The proofs of the above three lines are basically the same, so we prove just the first line :

$$\mathbf{F} \bullet {*}\mathbf{dx'} = \Sigma_{i=1}^{m} F_i \, ({*}dx'^i)$$

$$= \Sigma_{i=1}^{m} (-1)^{i-1} f_{12.[i]..m} (-1)^{i-1} \, dx'^1 \wedge dx'^2 .. [dx'^i] ... \wedge dx'^m$$

$$= \Sigma_{i=1}^{m} f_{12.[i]..m} \, dx'^1 \wedge dx'^2 .. [dx'^i] ... \wedge dx'^m$$

$$= [\, f_{234..m} \, dx'^2 \wedge dx'^3 \wedge dx'^4 ..... \wedge dx'^m + f_{134..m} \, dx'^1 \wedge dx'^3 \wedge dx'^4 ..... \wedge dx'^m + ... \,]$$

$$= \Sigma'_I f_I \, dx'^I_{\wedge} \quad \text{with series terms reordered from standard order}$$

$$= \mathbf{f} \bullet \mathbf{dx}_{\wedge}' \ . \qquad \text{(G.2.25)}$$

Because vector **n'** is a reordering of vector **m** where certain terms have minus signs, the sum of the squares of the components of the two vectors is the same, so

$$|\mathbf{n'}| = |\mathbf{m}| \ . \qquad \text{(G.2.26)}$$

Then dividing (G.2.23) and (G.2.24) by |**n'**| one finds

$$\hat{\mathbf{n}}' \bullet {*}\mathbf{dx} = \hat{\mathbf{m}} \bullet \mathbf{dx}_{\wedge}' \qquad \text{(G.2.27)}$$
$$\hat{\mathbf{n}}' \bullet \mathbf{F} = \hat{\mathbf{m}} \bullet \mathbf{f} \ . \qquad \text{(G.2.28)}$$

Recall from above that

$$\mu' = \hat{\mathbf{m}} \bullet \mathbf{dx}_{\wedge}' \qquad \textit{(G.2.13)}$$

$$\mathbf{f} \bullet \mathbf{dx}_{\wedge}' = (\mathbf{f} \bullet \hat{\mathbf{m}}) \, \mu' \ . \qquad \textit{(G.2.18)}$$

Now consider,

$$\mu' = \hat{\mathbf{m}} \bullet \mathbf{dx}_{\wedge}' \qquad \text{// (G.2.13) just above}$$

$$= \hat{\mathbf{n}}' \bullet {*}\mathbf{dx} \qquad \text{// (G.2.27)}$$

$$\mathbf{f} \bullet \mathbf{dx}_{\wedge}' = (\mathbf{f} \bullet \hat{\mathbf{m}}) \, \mu' \qquad \text{// (G.2.18) just above}$$

$$\mathbf{F} \bullet {*}\mathbf{dx'} = (\hat{\mathbf{n}}' \bullet \mathbf{F}) \, \mu' \qquad \text{// (G.2.22) on the left and (G.2.28) on the right}$$

We therefore obtain,

$$\mu' = \hat{\mathbf{n}}' \bullet {}^*\boldsymbol{dx}' \qquad \text{// see Sjamaar p 109 item 8.17} \qquad \text{(G.2.29)}$$

$$\mathbf{F} \bullet {}^*\boldsymbol{dx}' = (\mathbf{F} \bullet \hat{\mathbf{n}}')\, \mu' . \qquad \text{// see Sjamaar p 107 item 8.16} \qquad \text{(G.2.30)}$$

Sjamaar's equations are written in the $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ context rather than $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ so have ${}^*\boldsymbol{dx}$, $\hat{\mathbf{n}}$ and μ.

<u>Example</u>: $F : R^2 \rightarrow R^3$, $k = n = 2$, $m = 3$. (G.2.31)

This example was first treated in Section 10.10 as a no-differential-forms problem, and was then reconsidered as a 2-form problem in Section 10.13 (but in the $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{t})$ context). Here we write out various objects defined above and show how equations come out with equation number references in italics.

$$\alpha_{\mathbf{x}'} = \Sigma'_I f_I(\mathbf{x}')\, dx'^{\wedge I} = f_{12}\, dx'^1 \wedge dx'^2 + f_{13}\, dx'^1 \wedge dx'^3 + f_{23}\, dx'^2 \wedge dx'^3 \qquad \text{// a general 2-form}$$

$$= \mathbf{f} \bullet \boldsymbol{dx}^{\wedge\prime} = \mathbf{F} \bullet {}^*\boldsymbol{dx}' = \mathbf{F} \bullet \boldsymbol{dA}' = (\mathbf{F} \bullet \hat{\mathbf{n}}')\, dA' \qquad \text{// see below} \qquad \textit{(G.2.1)}$$

$$\mathbf{f} = (f_{12}, f_{13}, f_{23}) \qquad \textit{(G.2.7)}$$

$$m_{12} = \text{minor}_{12} = \det(R^{12}{}_{12}) = \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} \qquad \text{// R is a 3 x 2 matrix with 3 full-width minors}$$

$$m_{13} = \text{minor}_{13} = \det(R^{13}{}_{12}) = \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}$$

$$m_{23} = \text{minor}_{23} = \det(R^{23}{}_{12}) = \det\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} \qquad \textit{(G.2.4)}$$

$$\mathbf{m} = (m_{12}, m_{13}, m_{23}) = \left(\det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix}, \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}, \det\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} \right) \qquad \textit{(G.2.7)}$$

$$\mathbf{F} = (f_{23}, -f_{13}, f_{12}) \qquad \text{// agrees with (10.13.16)} \qquad \textit{(G.2.21)}$$

$$\mathbf{n}' = (m_{23}, -m_{13}, m_{12}) = \left(\det\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}, -\det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}, \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} \right) \qquad \textit{(G.2.21)}$$

$$= \left(\det\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix}, +\det\begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix}, \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} \right)$$

= agrees with (10.10.19) where $\mathbf{n}'$ is shown as normal to the surface

$$|\,\mathbf{m}\,|^2 = |\,\mathbf{n}'\,|^2 = \det^2\begin{pmatrix} R^2{}_1 & R^2{}_2 \\ R^3{}_1 & R^3{}_2 \end{pmatrix} + \det^2\begin{pmatrix} R^3{}_1 & R^3{}_2 \\ R^1{}_1 & R^1{}_2 \end{pmatrix} + \det\begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} \qquad \textit{(G.2.26)}$$

$$\equiv K^2 = \det(R^TR) \qquad \text{// agrees with (10.10.18) and (10.10.22)} \qquad \textit{(G.2.9)}$$

$$\boldsymbol{dx}^{\wedge\prime} = (dx'^1 \wedge dx'^2,\ dx'^1 \wedge dx'^3,\ dx'^2 \wedge dx'^3) \qquad \text{// standard order}$$

$${}^*\boldsymbol{dx}' \equiv ({}^*dx'^1, {}^*dx'^2, {}^*dx'^3) = (dx'^2 \wedge dx'^3, -dx'^1 \wedge dx'^3, dx'^1 \wedge dx'^2) \equiv \boldsymbol{dA}' \qquad \textit{(10.13.23)}$$

$\mathbf{f} \bullet \mathbf{dx}_\wedge' = (f_{12}, f_{13}, f_{23}) \bullet (dx'^1 \wedge dx'^2, dx'^1 \wedge dx'^3, dx'^2 \wedge dx'^3)$

$= f_{12}\, dx'^1 \wedge dx'^2 + f_{13}\, dx'^1 \wedge dx'^3 + f_{23}\, dx'^2 \wedge dx'^3 = \Sigma'_I f_I\, dx'_\wedge{}^I$

$\mathbf{F} \bullet {}^*\mathbf{dx}' = (f_{23}, -f_{13}, f_{12}) \bullet (dx'^2 \wedge dx'^3, -dx'^1 \wedge dx'^3, dx'^1 \wedge dx'^2)$

$= f_{23}\, dx'^2 \wedge dx'^3 + f_{13}\, dx'^1 \wedge dx'^3 + f_{12}\, dx'^1 \wedge dx'^2 = \Sigma'_I f_I\, dx'_\wedge{}^I = \mathbf{f} \bullet \mathbf{dx}_\wedge'$ *(G.2.22)*

$\mathbf{n}' \bullet \mathbf{F} = (m_{23}, -m_{13}, m_{12}) \bullet (f_{23}, -f_{13}, f_{12}) = m_{23}f_{23} + m_{13}f_{13} + m_{12}f_{12}$

$= m_{12}f_{12} + m_{13}f_{13} + m_{23}f_{23} = (m_{12}, m_{13}, m_{23}) \bullet (f_{12}, f_{13}, f_{23}) = \mathbf{m} \bullet \mathbf{f}$ *(G.2.24)*

$\mu' = \hat{\mathbf{n}}' \bullet {}^*\mathbf{dx}' = \hat{\mathbf{n}}' \bullet d\mathbf{A}' = dA'$ // area measure on surface *(G.2.29)*

$\mathbf{F} \bullet d\mathbf{A}' = (\mathbf{F} \bullet \hat{\mathbf{n}}')\, dA' = (\mathbf{F} \bullet \hat{\mathbf{n}}')\, \mu'$ // this is α_x, the integration integrand *(G.2.30)*

Example: $F : R^2 \to R^4$, $k = n = 2$, $m = 4$. (G.2.32)

We leave this as a reader exercise. The exercise is to write out new versions of all the equations of the previous example. One entry is $\mathbf{f} = (f_{12}, f_{13}, f_{14}, f_{23}, f_{24}, f_{34})$.

Appendix H : Hodge Star, Differential Operators, Integral Theorems and Maxwell

Here we study the relationship between the d and * operators, differential forms, and the classical differential operators of analysis such as the Laplacian. Some classical integral theorems are derived from the generalized Stokes' Theorem, and the Maxwell Equations are reformulated in terms of differential forms as an exercise. The cosmetic notation dx^i is used throughout for the functional λ^i. Since λ^i is a functional acting on the Cartesian vector space $V = R^n$, up and down tensor indices are the same.

H.1 Properties of the Hodge star operator in R^n

Start with,

$$dx^{\wedge I} = \text{some } \textit{ordered} \text{ multi-index wedge product of k } dx^i \text{ in } R^n \text{ (a basis vector k-form)} \quad . \qquad \text{(H.1.1)}$$

This $dx^{\wedge I}$ has k vectors wedged together in "standard order". The only non-zero n-form in R^n is this,

$$dV \equiv dx^1 \wedge dx^2 \wedge \ldots dx^n \,. \qquad \text{// "the volume form" in } R^n \qquad \text{(H.1.2)}$$

The Hodge dual object $*dx^{\wedge I}$ is defined as (sign is treated below),

$$*dx^{\wedge I} \equiv (\text{sign})_{I,k}\, dx^{\wedge Ic} \qquad \text{// Ic = complement of I} \qquad \text{(H.1.3)}$$

where $dx^{\wedge Ic}$ is the full wedge product dV in which the vectors of $dx^{\wedge I}$ are *deleted*.

Fact: Since $dx^{\wedge I}$ is a k-form, $*dx^{\wedge I}$ is an (n-k)-form which is "dual" to $dx^{\wedge I}$. (H.1.4)

For example, for the k-form

$$dx^{\wedge I} = dx^a \wedge dx^b \wedge \ldots\wedge dx^q \qquad \text{// } a < b < c \ldots < q$$

one has

$$dx^{\wedge Ic} = dx^1 \wedge dx^2 \wedge \ldots[dx^a] \ldots [dx^b]\ldots[dx^q] \ldots\wedge dx^n \qquad \text{(H.1.5)}$$

where the notation $[dx^a]$ means that dx^a is missing.

<u>Example</u>: In R^3 let $dx^{\wedge I} = dx^2$. Then $dx^{\wedge Ic} = dx^1 \wedge dx^3$.

<u>Example</u>: In R^6 let $dx^{\wedge I} = dx^2 \wedge dx^4$. Then $dx^{\wedge Ic} = dx^1 \wedge dx^3 \wedge dx^5 \wedge dx^6$.

Fact: The sign $(\text{sign})_{I,k}$ in (H.1.3) is selected so that the following is true :

$$dx^{\wedge I} \wedge (*dx^{\wedge I}) = dV \ . \qquad \text{(H.1.6)}$$

Fact: For a k-form $dx_{\wedge}^{I} = dx^{a} \wedge dx^{b} \wedge ....\wedge dx^{q}$ the sign in (H.1.3) is given by

$$(\text{sign})_{I,k} = (-1)^{a+b+..+q}\ (-1)^{k(k+1)/2}\ . \qquad \text{(H.1.7)}$$

Proof:

$$dx_{\wedge}^{I} \wedge (*dx_{\wedge}^{I}) = (\text{sign})_{I,k}\ dx_{\wedge}^{I} \wedge dx_{\wedge}^{Ic}$$

$= (\text{sign})_{I,k}\ (\ dx^{a} \wedge dx^{b} \wedge ....\wedge dx^{q}\ )$ // first factor has k vectors wedged

$(\ dx^{1} \wedge dx^{2} \wedge ...[dx^{a}]\ ....\ [dx^{b}].....[dx^{q}]\ .....\wedge dx^{n}\ )\ .$ (H.1.8)

The task is then to slide each of the dx^{i} of the first factor into its corresponding "hole" in the second factor and count up the number of adjacent vector position swaps required :

slide dx^{a} to the right, number of swaps = (k-1) + (a-1).
then slide dx^{b} to the right, number of swaps = (k-2) + (b-1)
then slide dx^{c} to the right, number of swaps = (k-3) + (c-1)
....
then slide dx^{q} to the right, number of swaps = (k-k) + (q-1) . (H.1.9)

Total swaps then is

$$\text{swaps} = \Sigma_{i=1}^{k}\ (k-i) + (a+b+...+q) - k\ . \qquad \text{(H.1.10)}$$

But

$$\Sigma_{i=1}^{k}\ (k-i) = k\ \Sigma_{i=1}^{k}\ [1] - \Sigma_{i=1}^{k}\ [i] = k * k - k(k+1)/2 = k^{2}/2 - k/2$$

so

$$\Sigma_{i=1}^{k}\ (k-i)\ - k = k^{2}/2 - 3k/2 = (k-3)(k/2)$$

and

$$\text{swaps} = (k-3)(k/2) + (a+b+...+q)\ . \qquad \text{(H.1.11)}$$

Then since each adjacent pairwise vector swap creates a (-1) factor according to (8.2.4), we get

$$\text{phase} = (-1)^{a+b+...+q}\ (-1)^{(k-3)k/2}$$

But $1 = (-1)^{2k} = (-1)^{4k/2}$ so

$$(-1)^{(k-3)k/2} = (-1)^{(k-3)k/2}\ (-1)^{4k/2} = (-1)^{(k+1)k/2}$$

and the result is

$$\text{phase} = (-1)^{a+b+...+q}\ (-1)^{(k+1)k/2}\ . \qquad \text{(H.1.12)}$$

After all these "slides" are completed, equation (H.1.8) says

$$dx_\wedge^{I} \wedge (*dx_\wedge^{I}) \;=\; (\text{sign})_{I,k} * \text{phase} * dx^1 \wedge dx^2 \wedge ...\, dx^n$$

$$= (\text{sign})_{I,k} * \text{phase} * dV\,.$$

According to the requirement (H.1.6) that $dx_\wedge^{I} \wedge (*dx_\wedge^{I}) = dV$ we find

$$(\text{sign})_{I,k} = \text{phase} = (-1)^{a+b+...+q}\,(-1)^{(k+1)k/2}\,. \qquad\qquad \textbf{QED}$$

Example: $dx_\wedge^{I} = dx^i$ in R^n, so $k = 1$ and $(-1)^{(k+1)k/2} = (-1)^{(1+1)1/2} = (-1)$. Then,

$$(\text{sign})_{I,k} = (-1)^i\,(-1) = (-1)^{i+1} = (-1)^{i-1}$$

$$*dx^i = (-1)^{i-1}\, dx^1 \wedge dx^2 \wedge ...\, [dx^i] ... \wedge dx^n\,. \tag{H.1.13}$$

Verify: $dx_\wedge^{I} \wedge (*dx_\wedge^{I}) = dx^i \wedge \{\, (-1)^{i-1}\, dx^1 \wedge dx^2 \wedge ...\, [dx^i] ... \wedge dx^n \,\}$

$$= (-1)^{i-1}\, dx^i \wedge\; dx^1 \wedge dx^2 \wedge ...\, [dx^i] ... \wedge dx^n \;=\; dx^1 \wedge dx^2 \wedge ...\, dx^n = dV\,.$$

One can form an n-component vector from the $*dx^i$ objects

$$*\mathbf{dx} \equiv (\, *dx^1,\; *dx^2, ...\; *dx^n)\,. \tag{H.1.14}$$

One can think of

$$*dx^i = dA^i \qquad \text{or} \qquad *\mathbf{dx} = \mathbf{dA} \tag{H.1.15}$$

as an element of "area" in n-1 dimensions. For R^3 we have (cyclic order)

$$\begin{aligned} *dx^1 &= dx^2 \wedge dx^3 = dA^1 & \quad dx_\wedge^{I} \wedge (*dx_\wedge^{I}) &= dx^1 \wedge (\,dx^2 \wedge dx^3) = dV \\ *dx^2 &= dx^3 \wedge dx^1 = dA^2 & \quad dx_\wedge^{I} \wedge (*dx_\wedge^{I}) &= dx^2 \wedge (\,dx^3 \wedge dx^1) = dV \\ *dx^3 &= dx^1 \wedge dx^2 = dA^3 & \quad dx_\wedge^{I} \wedge (*dx_\wedge^{I}) &= dx^3 \wedge (\,dx^1 \wedge dx^2) = dV \end{aligned} \tag{H.1.16}$$

or

$$dA^k = *dx^k = (1/2)\, \varepsilon_{kij}\, dx^i \wedge dx^j\,. \tag{H.1.17}$$

Another useful example:

Fact: $*dV = 1$ (H.1.18)

Proof: Then $dx_\wedge^{I} \wedge (*dx_\wedge^{I}) = dV \wedge (*dV) = dV \wedge 1 = dV$, satisfying (H.1.6) .

Fact: $dx^i \wedge *dx^j = \delta_{i,j}\ dV$ (H.1.19)

Proof: If $i \neq j$, then dx^i appears in $*dx^j$ since $*dx^j$ only has dx^j missing. But then dx^i appears twice, and so the wedge product must vanish, hence the factor $\delta_{i,j}$. And then $dx^i \wedge *dx^i = dV$ by (H.1.6) .

Fact: $*(*dx^{\wedge I}) = (-1)^{kn+k}\ dx^{\wedge I}$ (H.1.20)

Proof: The requirement is that

$$dx^{\wedge I} \wedge (*dx^{\wedge I}) = dV\ . \qquad \textit{(H.1.6)}$$

which applied to $*dx^{\wedge I}$ says

$$*dx^{\wedge I} \wedge *(*dx^{\wedge I}) = dV\ . \qquad \text{(H.1.21)}$$

We know that there exists *some* sign such that

$$*(*dx^{\wedge I}) = (\text{sign})\ dx^{\wedge I} \qquad \text{(H.1.22)}$$

since doing the complement twice restores all the original dx_i factors. Thus, (H.1.21) says

$$*dx^{\wedge I} \wedge\ [(\text{sign})\ dx^{\wedge I}\ ] = dV$$

or

$$(\text{sign})\ *dx^{\wedge I} \wedge dx^{\wedge I} = dV\ .$$

Now consider

$$*dx^{\wedge I} \wedge\ dx^{\wedge I} = (\text{sign}')\ dx^{\wedge I} \wedge *dx^{\wedge I} = (\text{sign}')\ dV \qquad \text{(H.1.23)}$$

where (sign') arises from sliding $dx^{\wedge I}$ to the left. Once we find (sign'), we then have

$$dV = (\text{sign})\ *dx^{\wedge I} \wedge\ dx^{\wedge I} = (\text{sign})(\text{sign}')\ dV$$

so the solution to our problem is then sign = sign'. To find sign' we slide each dx^i in $dx^{\wedge I}$ to the left in (H.1.23). Doing so, we pick up a sign $(-1)^{n-k}$ since n-k is the number of vectors in $*dx^{\wedge I}$. Doing this one at a time for each of the vectors in $dx^{\wedge I}$ one gets,

$$(\text{sign}') = (-1)^{(n-k)k} = (-1)^{kn-k^2} = (-1)^{kn}\ (-1)^{k^2} = (-1)^{kn}\ (-1)^{k} = (-1)^{kn+k}\ .$$

Therefore

$$*(*dx^{\wedge I}) = (\text{sign})\ dx^{\wedge I} = (\text{sign}')\ dx^{\wedge I} = (-1)^{kn+k}\ dx^{\wedge I}\ . \qquad \textbf{QED}$$

Corollary: If α is a k-form in R^n, then $*(*\alpha) = (-1)^{kn+k}\ \alpha$. (H.1.24)

Proof: $\alpha = \Sigma'_I f_I\, dx^{\wedge I}$

$\Rightarrow\ *(*\alpha) = \Sigma'_I f_I\, *(*dx^{\wedge I}) = \Sigma'_I f_I\, (-1)^{kn+k}\, dx^{\wedge I} = (-1)^{kn+k}\, \Sigma'_I f_I\, dx^{\wedge I} = (-1)^{kn+k}\, \alpha$

in agreement with Sjamaar p 28 Exercise 2.15.

Fact: If $\alpha = a_i dx^i$ and $\beta = b_j dx^j$ are 1-forms in R^n, then $\alpha \wedge (*\beta) = \mathbf{a}\bullet\mathbf{b}\ dV$. (H.1.25)

Proof: $\alpha \wedge(*\beta) = (a_i dx^i)\wedge(b_j * dx^j) = a_i b_j dx^i \wedge (*dx^j) = a_i b_j\, \delta_{i,j} dV$ by (H.1.19) = $\mathbf{a}\bullet\mathbf{b}\ dV$

So one can say that $\alpha \wedge(*\beta)$ is "Hodge-associated" with $\mathbf{a}\bullet\mathbf{b}$. In the non-dual space $L^2(R^n)$ one would have instead $\mathbf{a} \wedge (*\mathbf{b}) = (\mathbf{a} \bullet \mathbf{b})\, u_{\wedge Z} = (\mathbf{a} \bullet \mathbf{b})\, \mathbf{u}_1 \wedge \mathbf{u}_2 \wedge ... \mathbf{u}_n$. We already saw in (4.3.18b) the Hodge association $\mathbf{a} \wedge \mathbf{b} = \mathbf{A} \bullet (\mathbf{a} \times \mathbf{b})$ in $L^2(R^3)$, where $\mathbf{A}$ = vector area = $*\mathbf{u} = (*\mathbf{u}_1, *\mathbf{u}_2, *\mathbf{u}_3) = (A_1, A_2, A_3)$. Sometimes these two Hodge correspondences are written,

$$\mathbf{a} \wedge (*\mathbf{b}) \leftrightarrow \mathbf{a} \bullet \mathbf{b} \quad \text{for } R^n \qquad \mathbf{a} \wedge \mathbf{b} \leftrightarrow \mathbf{a} \times \mathbf{b} \quad \text{for } R^3\ . \tag{H.1.26}$$

H.2 Gradient

Start with a simple 0-form and compute $d\alpha$,

$\alpha = f$ // 0-form

$d\alpha = (\partial_i f)\, dx^i = \nabla f \bullet \mathbf{dx}$. // (10.3.3) (H.2.1)

One then has the following "Hodge correspondence",

$\alpha = f$	0-form in R^n	$\alpha \leftrightarrow f$
$d\alpha = \nabla f \bullet \mathbf{dx}$	1-form in R^n	$d\alpha \leftrightarrow \nabla f$. (H.2.2)

Apply Stokes's theorem (boundary here is two oriented endpoints of a curve C)

$$\int_M d\alpha = \int_{\partial M} \alpha$$

$$\int_M \nabla f \bullet \mathbf{dx} = \int_{\partial M} f$$

$$\int_C \nabla f \bullet \mathbf{dx} = f(\mathbf{b}) - f(\mathbf{a})\ . \tag{H.2.3}$$

When both sides are converted to regular calculus integrals (two definitions of Section 10.11), one gets

$$\int_C \nabla f \bullet \mathbf{dx} = f(\mathbf{b}) - f(\mathbf{a}) \tag{H.2.4}$$

which we shall call the "**line integral of a gradient theorem**" .

H.3 Laplacian

Start again with a simple 0-form and compute various interesting objects :

$\alpha = f$ // 0-form

$d\alpha = (\partial_i f)\, dx^i$ // (10.3.3) (H.3.1)

$*(d\alpha) = (\partial_i f)\, *dx^i$

$d(*d\alpha) = (\partial_j\partial_i f)\ dx^j \wedge *dx^i$ // (10.3.3)

$= (\partial_j\partial_i f)\ \ \delta_{i,j}\ dV$ // (H.1.19)

$= (\partial^2{}_i f\,)\, dV$

$= (\nabla^2 f)\ dV$

$*(d(*d\alpha)) = \nabla^2 f\ (*dV) = \nabla^2 f\ .$ // (H.1.18) (H.3.2)

One then has the following "Hodge correspondence",

$\alpha = f$	0-form in R^n	$\alpha \leftrightarrow f$	
$*(d(*d\alpha)) = \nabla^2 f$	0-form on R^n	$*(d(*d\alpha)) \leftrightarrow \nabla^2 f\ .$	(H.3.3)

Consider now,

$\beta \equiv f\,\nabla g \bullet (*\mathbf{dx}) = f\,\nabla g \bullet \mathbf{dA} = f\,(\partial_i g)\,(*dx^i)$ // (H.1.15) (H.3.4)

$d\beta = d\,[\,f\,(\partial_i g)\,]\ \ (*dx^i)$

$= \partial_j\,[\,f\,(\partial_i g)\,]\, dx^j \wedge (*dx^i)$ // (10.3.3)

$= \partial_j\,[\,f\,(\partial_i g)\,]\,\delta_{i,j}\, dV$ // (H.1.19)

$= \partial_i\,[\,f\,(\partial_i g)\,]\, dV$

$= [\,f\,(\partial_i{}^2 g) + (\partial_i f)(\partial_i g)\,]\, dV$

$= [\,f\,(\nabla^2 g) + \nabla f \bullet \nabla g\,]\, dV\,.$ (H.3.5)

Apply Stokes's theorem,

$\int_{\mathbf{M}} d\beta = \int_{\partial\mathbf{M}} \beta$

$\int_{\mathbf{M}}\ [\,f\,(\nabla^2 g)\ +\ \nabla f \bullet \nabla g\,]\, dV = \int_{\partial\mathbf{M}} f\,\nabla g \bullet \mathbf{dA}\ \ .$ (H.3.6)

When both sides are converted to regular calculus integrals (two definitions of Section 10.11), one gets

$$\int_V [\, f\,\nabla^2 g + \nabla f \bullet \nabla g\,]\, dV \;=\; \int_S f\,\nabla g \bullet d\mathbf{A} \;=\; \int_S f\,\nabla g \bullet [\hat{\mathbf{n}}\, dA] \;=\; \int_S f\,(\partial_n g) dA \qquad \text{(H.3.7)}$$

which is known as **Green's first identity**. Swapping f↔g and subtracting gives

$$\int_V [\, f\,\nabla^2 g - g\,\nabla^2 f]\, dV \;=\; \int_S [f\,(\partial_n g) - g\,(\partial_n f)]\, dA \qquad \text{(H.3.8)}$$

which is **Green's second identity**.

H.4 Divergence

Start this time with a 1-form and compute various interesting objects :

$$\alpha = F_i\, dx^i \;=\; \mathbf{F} \bullet \mathbf{dx} \qquad // \text{ 1-form} \qquad \text{(H.4.1)}$$

$$*\alpha = \; F_i\, (*dx^i) \;=\; \mathbf{F} \bullet *\mathbf{dx} \;=\; \mathbf{F} \bullet \mathbf{dA} \qquad // \text{ (H.1.15)} \qquad \text{(H.4.2)}$$

$$d(*\alpha) \;=\; \partial_j F_i\, dx^j \wedge (*dx^i) \qquad // \text{ (10.3.3)}$$

$$=\; \partial_j F_i\, \delta_{i,j}\, dV \qquad // \text{ (H.1.19)}$$

$$= (\partial_i F_i)\, dV$$

$$= (\text{div } \mathbf{F})\, dV \qquad \text{(H.4.3)}$$

$$*(d(*\alpha)) \;=\; (\text{div } \mathbf{F})\, *dV \;=\; \text{div } \mathbf{F} \quad . \qquad //\text{(H.1.18)} \qquad \text{(H.4.4)}$$

One then has the following "Hodge correspondence",

$\alpha = \mathbf{F} \bullet \mathbf{dx}$	1-form on R^n	$\alpha \leftrightarrow \mathbf{F}$	
$*(d(*\alpha)) = \text{div } \mathbf{F}$	0-form on R^n	$*(d(*\alpha)) \leftrightarrow \text{div } \mathbf{F}$.	(H.4.5)

Apply Stokes' Theorem with $\beta = *\alpha$:

$$\int_M d\beta \;=\; \int_{\partial M} \beta$$

$$\int_M d(*\alpha) \;=\; \int_{\partial M} (*\alpha)$$

$$\int_M \text{div } \mathbf{F}\;\, dV \;=\; \int_{\partial M} \mathbf{F} \bullet \mathbf{dA} \quad . \qquad \text{(H.4.6)}$$

When both sides are converted to regular calculus integrals (two definitions of Section 10.11), one gets

$$\int_V \text{div}\, \mathbf{F}\ dV = \int_S \mathbf{F} \bullet d\mathbf{A} \qquad \text{(H.4.7)}$$

which is the **divergence theorem** in n dimensions. For R^3 this is Gauss's Theorem.

H.5 Curl

Start again with a 1-form and compute objects of interest:

$$\alpha = \Sigma_j F_j\, dx^j = \mathbf{F} \bullet \mathbf{dx} \qquad \text{// 1-form} \qquad \text{(H.5.1)}$$

$$d\alpha = \Sigma_{i<j} (\partial_i F_j - \partial_j F_i)\, dx^i \wedge dx^j \qquad \text{// (10.3.24b), } d\alpha \text{ written in standard form} \qquad \text{(H.5.2)}$$

$$*(d\alpha) = \Sigma_{i<j} (\partial_i F_j - \partial_j F_i)\, *(dx^i \wedge dx^j) \qquad \text{// on next line specialize to } R^3 :$$

$$= (\partial_1 F_2 - \partial_2 F_1)\, *(dx^1 \wedge dx^2) + (\partial_1 F_3 - \partial_3 F_1)\, *(dx^1 \wedge dx^3) + (\partial_2 F_3 - \partial_3 F_2)\, *(dx^3 \wedge dx^2)$$

$$= (\partial_1 F_2 - \partial_2 F_1)\, *(dx^1 \wedge dx^2) + (\partial_3 F_1 - \partial_1 F_3)\, *(dx^3 \wedge dx^1) + (\partial_2 F_3 - \partial_3 F_2)\, *(dx^3 \wedge dx^2)$$

$$= (\text{curl}\, \mathbf{F})_3\, dx^3 + (\text{curl}\, \mathbf{F})_2\, dx^2 + (\text{curl}\, \mathbf{F})_1\, dx^1$$

$$= (\text{curl}\, \mathbf{F}) \bullet \mathbf{dx}\ . \qquad \text{(H.5.3)}$$

One then has the following "Hodge correspondence",

$$\begin{array}{lll} \alpha = \mathbf{F} \bullet \mathbf{dx} & \text{1-form in } R^3 & \alpha \leftrightarrow \mathbf{F} \\ *(d\alpha) = [\text{curl}\, \mathbf{F}] \bullet \mathbf{dx} & \text{1-form in } R^3 & *(d\alpha) \leftrightarrow \text{curl}\, \mathbf{F}\, . \end{array} \qquad \text{(H.5.4)}$$

Apply Stokes' Theorem in R^n to get

$$\int_M d\alpha = \int_{\partial M} \alpha$$

$$\int_M \Sigma_{i<j} (\partial_i F_j - \partial_j F_i)\, dx^i \wedge dx^j = \int_{\partial M} \mathbf{F} \bullet \mathbf{dx} \qquad \text{(H.5.5)}$$

• In R^2 there is only one term in the sum on the left and one gets,

$$\int_M (\partial_1 F_2 - \partial_2 F_1)\, dx^1 \wedge dx^2 = \int_{\partial M} [\, F_1 dx^1 + F_2\, dx^2\,]\, . \qquad \text{(H.5.6)}$$

When both sides are converted to regular calculus integrals (two definitions of Section 10.11),

$$\int_A (\partial_1 F_2 - \partial_2 F_1)\, dx^1 dx^2 = \int_C [\, F_1 dx^1 + F_2 dx^2]\, . \qquad \text{(H.5.7)}$$

Setting $x^1 = x$, $x^2 = y$, $F_1 = f$ and $F_2 = g$ one gets

$$\int_A (\partial_x g - \partial_y f)\, dxdy = \int_C [\, fdx + gdy] \qquad \text{(H.5.8)}$$

which is known as **Green's Theorem in the plane**.

• In R^3 , we can write out the three terms on the left side of (H.5.2)

$$d\alpha = (\partial_1 F_2 - \partial_2 F_1)\, dx^1 \wedge dx^2 + (\partial_1 F_3 - \partial_3 F_1)\, dx^1 \wedge dx^3 + (\partial_2 F_3 - \partial_3 F_2)\, dx^2 \wedge dx^3$$

$$= (\partial_1 F_2 - \partial_2 F_1)\, dA^3 + (\partial_1 F_3 - \partial_3 F_1)\, [-dA^2\,] + (\partial_2 F_3 - \partial_3 F_2)\, dA^1$$

$$= (\partial_2 F_3 - \partial_3 F_2)\, dA^1 + (\partial_3 F_1 - \partial_1 F_3)\, dA^2 + (\partial_1 F_2 - \partial_2 F_1)\, dA^3$$

$$= (\text{curl } \mathbf{F})_1\, dA^1 + (\text{curl } \mathbf{F})_2\, dA^2 + (\text{curl } \mathbf{F})_3\, dA^3$$

$$= (\text{curl } \mathbf{F}) \bullet d\mathbf{A} \,. \qquad \text{(H.5.9)}$$

Then Stokes' Theorem says

$$\int_M d\alpha = \int_{\partial M} \alpha$$

$$\int_M (\text{curl } \mathbf{F}) \bullet d\mathbf{A} = \int_{\partial M} \mathbf{F} \bullet d\mathbf{x} \,. \qquad \text{(H.5.10)}$$

When both sides are converted to regular calculus integrals (two definitions of Section 10.11), one gets

$$\int_A (\text{curl } \mathbf{F}) \bullet d\mathbf{A} = \int_C \mathbf{F} \bullet d\mathbf{x} = \oint_C \mathbf{F} \bullet d\mathbf{x} \qquad \text{(H.5.11)}$$

which is the **traditional Stokes' Theorem** in R^3 where C is the boundary of the area A. Note that the boundary C and its enclosed area A can be non-planar.

<u>An exercise using the ε tensor</u>

Consider the following area 2-form dA_k in R^3 (implied sums on all repeated indices),

$$dA_k \equiv (1/2)\varepsilon_{ijk}\, dx^i \wedge dx^j \qquad \Rightarrow \qquad dx^i \wedge dx^j = \varepsilon_{ijm}\, dA_m \,. \qquad \text{(H.5.12)}$$

This can be verified by applying $\Sigma_{ij}(1/2)\varepsilon_{ijk}$ to the equation on the right,

$$(1/2)\varepsilon_{ijk}[\, \varepsilon_{ijm} q_m] = (1/2)\ \{\varepsilon_{ijk}\, \varepsilon_{ijm}\} dA_m = (1/2)\ \{2\, \delta_{k,m}\} dA_m = dA_k \,.$$

Then consider a general 0-form,

$\alpha = f$ // 0-form

$d\alpha = (\partial_j f)\, dx^j$

$d^2\alpha = \partial_i(\partial_j f)\, dx^i \wedge dx^j$

$= (\partial_{ij} f)\, \varepsilon_{ijm}\, dA_m$ // (H.5.12), = 0 by symmetry on the i,j indices

$= dA_m\, [\, \varepsilon_{mij}\, \partial_i(\partial_j f)]$ // cyclic rule $\varepsilon_{abc} = \varepsilon_{bca} = \varepsilon_{cab}$

$= d\mathbf{A} \bullet [\text{curl grad f}]\ .$ (H.5.13)

Thus one can associate $d^2\alpha = 0$ for a 0-form f with the fact that curl grad f $= \nabla x(\nabla f) = 0$.

Similarly, consider the following 3-form in R^3 (implied sums on all repeated indices),

$dV \equiv (1/6)\, \varepsilon_{ijk}\, dx^i \wedge dx^j \wedge dx^k \qquad \Rightarrow \qquad dx^i \wedge dx^j \wedge dx^k = \varepsilon_{ijk}\, dV\ .$ (H.5.14)

This can be verified by applying $\Sigma_{ijk}(1/6)\, \varepsilon_{ijk}$ to the equation on the right,

$(1/6)\, \{\, \varepsilon_{ijk}\varepsilon_{ijk}\}\, dV = (1/6)\, \{\, 3!\}\, dV = dV\ .$

Then consider a general 1-form in R^3,

$\alpha = f_k\, dx^k = \mathbf{f} \bullet d\mathbf{x}$

$d\alpha = (\partial_j f_k)\, dx^j \wedge dx^k$

$d^2\alpha = \partial_i(\partial_j f_k)\, dx^i \wedge dx^j \wedge dx^k$

$= \partial_i(\partial_j f_k)[\, \varepsilon_{ijk}\, dV]$ // (H.5.14), = 0 by symmetry on the i,j indices

$= dV\, \partial_i[\varepsilon_{ijk}(\partial_j f_k)]$

$= dV\ \partial_i\, [\text{curl } \mathbf{f}\,]_i$

$= dV \text{ div curl } \mathbf{f}\ .$ (H.5.15)

Thus one can associate $d^2\alpha = 0$ for a 1-form $\alpha = \mathbf{f} \bullet d\mathbf{x}$ with the fact that div curl $\mathbf{f} = \nabla\bullet(\nabla x\mathbf{f}) = 0$.

H.6 Exercise: Maxwell's Equations in Differential Forms

This section is based on Sjamaar p 30 Exercise 2.23, but we use SI units instead of cgs units.

Maxwell's equations in SI units are,

$\text{curl}\ \mathbf{H} = \partial_t\mathbf{D} + \mathbf{J}$ Maxwell curl H equation

$\text{curl}\ \mathbf{E} = -\ \partial_t\mathbf{B}$ Maxwell curl E equation

$\text{div}\ \mathbf{D} = \rho$ Maxwell div D equation

$\text{div}\ \mathbf{B} = 0\ .$ Maxwell div B equation (H.6.1)

Write these equations in components and think of time $cdt = dx^4$ with $c = 1$, so we are working here in spacetime R^4. The metric tensor is $\pm\text{diag}(1,1,1,-1)$ but this fact has no effect on the presentation below.

$(\partial_iH_j - \partial_jH_i) - \varepsilon_{ijk}(\partial_4D_k) = \varepsilon_{ijk}\ J_k$ // for example, $(\partial_1H_2 - \partial_2H_1) - \partial_4D_3 = J_3$

$(\partial_iE_j - \partial_jE_i) + \varepsilon_{ijk}(\partial_4B_k) = 0$

$\partial_iD_i = \rho$

$\partial_iB_i = 0\ .$ (H.6.2)

In the above, indices i,j,k range from 1 to 3 and all implied sums have this range.

Define two differential 2-forms α and β as follows,

$\alpha \equiv (\mathbf{E} \bullet \boldsymbol{dx}) \wedge dx^4 + \mathbf{B} \bullet \boldsymbol{dA}$ // $\boldsymbol{dA} = {}^*\boldsymbol{dx}$

$\beta \equiv -\ (\mathbf{H} \bullet \boldsymbol{dx}) \wedge dx^4 + \mathbf{D} \bullet \boldsymbol{dA}$ (H.6.3)

where $\boldsymbol{dx}$ and $\boldsymbol{dA}$ and dV refer to R^3 objects as used in earlier sections above.

Start with α written in components and compute dα. Again, all *implied* sums are summed 1 to 3. Then,

$\alpha = E_j\ dx^j \wedge dx^4 + B_j\ {}^*dx^j$

$d\alpha = \Sigma_{i=1}^4\ (\partial_iE_j)\ dx^i \wedge dx^j \wedge dx^4 + \Sigma_{i=1}^4\ (\partial_iB_j)\ dx^i \wedge {}^*dx^j$ // (10.3.6)

$= (\partial_iE_j)\ dx^i \wedge dx^j \wedge dx^4 + (\partial_iB_j)\ dx^i \wedge {}^*dx^j$
$+ (\partial_4E_j)\ dx^4 \wedge dx^j \wedge dx^4 + (\partial_4B_j)\ dx^4 \wedge {}^*dx^j\ .$

The third term vanishes since there are two dx^4 vectors present. In the second term use

$$dx^i \wedge *dx^j = \delta_{i,j}\, dx^i \wedge *dx^i = \delta_{i,j}\, dV \qquad // \text{ (H.1.19)}$$

so

$$(\partial_i B_j)\, dx^i \wedge *dx^j = (\partial_i B_j)\delta_{i,j}\, dV = (\partial_i B_i)\, dV .$$

Then

$$d\alpha = (\partial_i E_j)\, dx^i \wedge dx^j \wedge dx^4 + (\partial_i B_i)\, dV + 0 + (\partial_4 B_k)\, dx^4 \wedge *dx^k$$

$$= (\partial_i E_j)\, dx^i \wedge dx^j \wedge dx^4 + (\partial_4 B_k)\, dx^4 \wedge *dx^k + (\partial_i B_i)\, dV . \qquad \text{(H.6.4)}$$

Recall from (H.1.17) that

$$*dx^k = dA^k = (1/2)\, \varepsilon_{kij}\, dx^i \wedge dx^j = (1/2)\, \varepsilon_{ijk}\, dx^i \wedge dx^j . \qquad \textit{(H.1.17)}$$

Using this fact, and writing the first term in dα as two terms, we find

$$d\alpha = (1/2)\, (\partial_i E_j - \partial_j E_i)\, dx^i \wedge dx^j \wedge dx^4 + (\partial_4 B_k)\, (1/2)\, \varepsilon_{ijk}\, dx^4 \wedge dx^i \wedge dx^j + (\partial_i B_i)\, dV$$

$$= (1/2)\, [\, (\partial_i E_j - \partial_j E_i) + \varepsilon_{ijk}(\partial_4 B_k)\,]\, dx^i \wedge dx^j \wedge dx^4 + (\partial_i B_i)\, dV . \qquad \text{(H.6.5)}$$

According to Maxwell's equations (H.6.2) each of these terms vanishes so the result is simply

$$d\alpha = 0 . \qquad \text{(H.6.6)}$$

The form β in (H.6.3) is the same as α with replacements: **E** → -**H** and **B** → **D**. We can then convert result (H.6.5) to get

$$d\beta = (1/2)\, [- (\partial_i H_j - \partial_j H_i) + \varepsilon_{ijk}(\partial_4 D_k)\,]\, dx^i \wedge dx^j \wedge dx^4 + (\partial_i D_i)\, dV . \qquad \text{(H.6.7)}$$

According to Maxwell's equations (H.6.2) we then get (writing the result many ways),

$$d\beta = (1/2)\, [\, - \varepsilon_{ijk}\, J_k\,]\, dx^i \wedge dx^j \wedge dx^4 + \rho\, dV$$

$$= - J_k\, [\, (1/2)\, \varepsilon_{ijk}\, dx^i \wedge dx^j] \wedge dx^4 + \rho\, dV$$

$$= - J_k\, dA^k \wedge dx^4 + \rho\, dV = - (\mathbf{J} \bullet d\mathbf{A}) \wedge dx^4 + \rho\, dV$$

$$= - J_k\, *dx^k \wedge dx^4 + \rho\, dV = - (\mathbf{J} \bullet *d\mathbf{x}) \wedge dx^4 + \rho\, dV . \qquad \text{(H.6.8)}$$

Now compute

$$d(d\beta) = d\, [\, - J_k\, *dx^k \wedge dx^4 + \rho\, dV\,]$$

$$= - \Sigma_{j=1}{}^4\, (\partial_j J_k)\, dx^j \wedge *dx^k \wedge dx^4 + \Sigma_{j=1}{}^4\, (\partial_j \rho)\, dx^j \wedge dV$$

$$= - (\partial_j J_k)\, dx^j \wedge *dx^k \wedge dx^4 + (\partial_j \rho)\, dx^j \wedge dV$$
$$- (\partial_4 J_k)\, dx^4 \wedge *dx^k \wedge dx^4 + (\partial_4 \rho)\, dx^4 \wedge dV$$

$$= - (\partial_j J_k)\, \delta_{j,k}\, dV \wedge dx^4 + 0 \qquad // \text{(H.1.19)}$$
$$- 0 - (\partial_4 \rho)\, dV \wedge dx^4$$

$$= - [(\partial_j J_j) + (\partial_4 \rho)]\, dV \wedge dx^4$$

$$= - [\text{div}\, \mathbf{J} + (\partial_t \rho)]\, dV \wedge dx^4 \ . \tag{H.6.9}$$

But $d^2\beta = 0$ from (10.3.10) so we conclude that

$$\text{div}\, \mathbf{J} + (\partial_t \rho) = 0 \tag{H.6.10}$$

which is the well-known equation of continuity stating that charge is conserved,

$$- \partial_t[\int_V \rho\, dV] = \int_S \mathbf{J} \bullet d\mathbf{S} \ . \tag{H.6.11}$$

"charge enclosed in V decreases at a rate equal to the current flowing out through boundary S"

Here then is a summary of our results :

$$\alpha \equiv (\mathbf{E} \bullet d\mathbf{x}) \wedge dx^4 + \mathbf{B} \bullet d\mathbf{A} \qquad // d\mathbf{A} = *d\mathbf{x}$$

$$d\alpha = 0 \quad \Leftrightarrow \quad \text{curl}\, \mathbf{E} = - \partial_t \mathbf{B} \text{ and div}\mathbf{B} = 0$$

$$\beta \equiv - (\mathbf{H} \bullet d\mathbf{x}) \wedge dx^4 + \mathbf{D} \bullet d\mathbf{A}$$

$$d\beta = - (\mathbf{J} \bullet d\mathbf{A}) \wedge dx^4 + \rho\, dV \quad \Leftrightarrow \quad \text{curl}\, \mathbf{H} = \partial_t \mathbf{D} + \mathbf{J} \text{ and div}\mathbf{D} = \rho$$

$$d^2\beta = 0 \quad \Leftrightarrow \quad \text{div}\, \mathbf{J} + (\partial_t \rho) = 0 \ . \qquad // \partial_\mu J^\mu = 0 \tag{H.6.12}$$

In free space where $\mathbf{J} = \rho = 0$, Maxwell's Equations take this impressively simple form,

$$d\alpha = 0$$
$$d\beta = 0 \ . \tag{H.6.13}$$

References

Listed in alphabetical order by last name of first author. A quick web search on article title can usually locate documents with broken links. Those below were last checked June 4, 2016.

I.M. Benn and R.W. Tucker, *An Introduction to Spinors and Geometry with Applications in Physics* (Adam Hilger/IOP Publishing, Bristol, 1987).

G. Birkhoff and S. Mac Lane, *A Survey of Modern Algebra, 4th Ed*. (McMillan, New York, 1977).

R.C. Buck with E.F. Buck, *Advanced Calculus, 2nd Ed.* (McGraw-Hill, New York, 1965). Our references are to this 2nd edition, but there is a third edition (McGraw-Hill, New York, 1978), reprinted by (Waveland Press, Long Grove IL, 2003). This excellent and now classic book was first published in 1956.

K. Conrad, *Exterior Powers* (2013, find at http://www.math.uconn.edu/~kconrad/blurbs/).

J. Denker, *Introduction to Clifford Algebra*, (2003, https://www.av8n.com/physics/clifford-intro.pdf).

P.A.M. Dirac, *The Principles of Quantum Mechanics, 3rd Ed.* (Oxford University Press, London, 1947).

S. Lang, *Fundamentals of Differential Geometry* (Springer-Verlag, New York, 1999).

S. Lang, *Algebra, 3rd Ed*. (Springer, New York, 2002). A soft-cover edition appeared in 2012.

P. Lucht, *Tensor Analysis and Curvilinear Coordinates* (2016, http://user.xmission.com/~rimrock).

S. Mac Lane and G. Birkhoff, *Algebra*, *3rd Ed* (Amer. Math. Soc. / Chelsea Pub, New York, 1988).

A. Messiah, *Quantum Mechanics* (John Wiley, New York, 1958), two volumes.

S. Roman, *Advanced Linear Algebra, 3rd Ed.* (Springer, New York, 2008).

R. Shankar, *Principles of Quantum Mechanics*, *2nd Ed.* (Plenum Press, New York, 1994).

R. Sjamaar, *Manifolds and Differential Forms* (2015, http://www.math.cornell.edu/~sjamaar/manifolds).

M. Spivak, *Calculus on Manifolds* (Addison-Wesley, New York, 1965).

M. Spivak, *A Comprehensive Introduction to Differential Geometry*, *Volume 1, 3rd Ed*. Volume 1 (Publish or Perish Inc, Houston, 1999), part of a well-regarded five volume set.

J. Suter. *Geometric Algebra Primer* (2003, www.jaapsuter.com/geometric-algebra.pdf).

L.W. Tu, *An Introduction to Manifolds, 2nd Ed*. (Springer, New York, 2011).