Sjamaar Chapter 7 Notes PhL 8.14.15

Cutting first real new ground here after Cape Cod trip several weeks ago.

Chapter 7: Differential Forms on Manifolds.

7.1 First definition

In earlier sections, we learned about differential forms defined only on R^n. The question now is how you might generalize the concept of differential forms for use on a manifold M of dimension n *within* R^N. Earlier we always had things like dxdy where these were R^n coordinates.

But didn't we already do some forms work on surfaces? Stokes' theorem for example involves regions and boundaries either of which could be surfaces. I should review that a bit before continuing. On page 47 (Chap 4) we integrate over a curve c which exists in R^n . Then in Chap 5 we do $\int_c \alpha$ where c is a surface of dimension k within R^n. So to the extend these integration domains were manifolds, we have in fact already done some integration of forms over manifolds. So I will be looking for what is new in this Chapter. [2.14.16: Chap 5 integrated forms over k-cubes and k-chains, not over manifolds.]

Sja states the "problem" here as that if you are at a point on a manifold, the "usual" basis vectors I guess of R^N might take you out of the manifold so those derivatives don't exist within M. But recall from the last chapter that when Sja talked about embeddings and manifolds, a point **x** on the manifold M was always surrounded by an open set V whose intersection with M gives a path on the surface. So if that set V exists surrounding x in all directions, then the derivatives in all directions do make sense. When we play with (Dψ), we are saying that all these derivatives exist. The existence of that cushion of space V surrounding a point **x** on M is of course directly connected to the notion of "embedding" and "manifold".

So Sja wants us to think of an atlas of embeddings whose charts are labeled by index i, such as ψ_i , which cover M, and we can then think within each of these embeddings. We break our domain space into a set of probably overlapping U_i and these go to cushions V_i with ψ_i.

Recall that a 0-form is just a function f where we usually wrote $\alpha = \Sigma_I \; f_I(\mathbf{x})\, dx_I$ and for a 0-form we would write α = f(x). In the pullback chapter we wrote $\alpha = \Sigma\, f_I(y)\, dy_I$ (p 36) with $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ where x was a point in our flat pullback space, picture page 37. But here as in Chap 6 we want $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ where t is in our flat space as in the toroid picture of page 68. I now have to clarify and merge the forms world with the manifolds worlds, and get a consistent notation. On page 81 we have

$\psi_i: U_i \to R^N$ $\quad$ $\mathbf{t} \in U_i$ $\quad$ $\psi_i(U_i)$ = patch on M, part of a cover of M

Let's assume for the moment that we are dealing with $\alpha = \Sigma_I \; f_I(\mathbf{x})\, dx_I$ where **x** lies on M. And for the simple 0-form case we have α = f(**x**) where $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ for our embedding. At a particular patch then we have

$\alpha_i = f(\mathbf{x}_i) = f(\psi_i(\mathbf{t})) \equiv f_i(\mathbf{t})$ // new *functional* form of t (different function of t)

Then $f_i(t)$ is the **local representative** of f in the tiny mapping i. Now recall that we can write the above using the pullback operator ψ_i^* in this way

$f_i(\mathbf{t}) = f(\psi_i(\mathbf{t})) = \psi_i^* f(\mathbf{t}) = (f \circ \psi_i)(\mathbf{t})$.

So we have $f_i = \psi_i^* f$ and we can say that that the 0-form representative f_i is a pullback of f.

So now we are interested in derivatives of $f_i(\mathbf{t})$ in the little ψ_i world only (why?). Right now I think Sja is talking only derivatives with respect to **t** components in the flat space side. Maybe f_i is C^k for all i, and then we can say that f is C^k overall.

What then happens at the **overlap of two charts** i and j? Let **t** in U_i and **u** in U_j and both these are in an overlap region. Then

$\mathbf{x} = "\psi(t)" = \psi_i(\mathbf{t}) = \psi_j(\mathbf{u}) \Rightarrow$ **x** = point on the manifold M

$f(\mathbf{x}) = f(\psi_i(\mathbf{t})) = f(\psi_j(\mathbf{u})) \Rightarrow$

$f(\mathbf{x}) = f_i(\mathbf{t}) = f_j(\mathbf{u})$

So consider: [recall that ψ^{-1} exists for any embedding]

$\psi_i(\mathbf{t}) = \psi_j(\mathbf{u}) \Rightarrow \mathbf{t} = \psi_i^{-1}(\psi_j(\mathbf{u})) = (\psi_i^{-1} \circ \psi_j)(\mathbf{u})$ (*)

$f_i(\mathbf{t}) = f_j(\mathbf{u}) \Rightarrow \mathbf{t} = f^{-1}(f_j(\mathbf{u})) = (f_i^{-1} \circ f_j)(\mathbf{u})$

From the top right result (*) we can write

$f_i(\mathbf{t}) = f_i[(\psi_i^{-1} \circ \psi_j)(\mathbf{u})]$.

But using pullback notation, this says

$f_i(\mathbf{t}) = (\psi_i^{-1} \circ \psi_j)^* f_i(\mathbf{u})$.

But since $f_i(\mathbf{t}) = f_j(\mathbf{u})$ on the left we end up with

$f_j(\mathbf{u}) = (\psi_i^{-1} \circ \psi_j)^* f_i(\mathbf{u})$ for any **u** in U_j

and so we suppress the argument to say

$f_j = (\psi_i^{-1} \circ \psi_j)^* f_i$ which is p 81 A "consistency condition"

This is a consistency condition on representatives of the 0-form f on two regions U_i and U_j which overlap and which have embedding functions ψ_i and ψ_j. Interestingly, you can interpret f_i as a pullback of f_j and vice versa. It seems odd that two objects on a fairly equal footing are pullbacks of each other.

We can alternately write from (*) that

$$\mathbf{u} = \psi_j^{-1}(\psi_i(\mathbf{t})) \;=\; (\psi_j^{-1} \circ \psi_i)(\mathbf{t}) \,.$$

If we let **t** run over all of U_i , we get this set of possible **u** values.

$$\{\mathbf{u}\} = (\psi_j^{-1} \circ \psi_i)(U_i) \qquad \text{which is} \qquad \text{p 81 B}$$

where we are now assuming that **u** and **t** and **x** are all in a region of chart i and j overlap.

In going from region U_i with **t** to region U_j with **u**, we are doing a **change of coordinates**. Switching from map of New York to map of New Jersey.

[82]

With this as introduction, we now want to talk about forms. We just showed that on manifold M we can replace f by its local representative f_i on each U_i region of the flat space, so that is how we dealt with 0-forms. Now for k-forms, we try doing the same idea: represent α locally by a form α_i and mimic the above

$$f_j = (\psi_i^{-1} \circ \psi_j)^* f_i \qquad \Rightarrow \qquad \text{// line below mimics this line}$$

$$\alpha_j = (\psi_i^{-1} \circ \psi_j)^* \alpha_i \qquad \text{which is} \qquad \text{p 82} \quad (7.1)$$

Now remember that the φ^* type operator affects both the function and the derivative part of something like $\alpha = \Sigma_I \; f_I(\mathbf{x})\, dx_I$, so we are generalizing off our *function* result to get a more general *form* result.

Now α_i is the local representative of the k-form α on U_i and for ψ_i.

And now we have α_i being a pullback of α_j and vice versa, where α_i and α_j are now representatives of k-form α on the two regions U_i and U_j which are associated with embedding functions ψ_i and ψ_j.

Also, for the function we had

$$f_i(\mathbf{t}) = f(\psi_i(\mathbf{t})) = \psi_i^* f(\mathbf{t}) = (f \circ \psi_i)(\mathbf{t})$$

which we abbreviate to say

$$\psi_i^* f(\mathbf{t}) = f_i(\mathbf{t}) \;.$$

If we replace f_i by α_i and f by α, then we have

$$\psi_i^* \alpha = \alpha_i$$

The k-form α_i representative on U_i is thus a pullback under ψ_i^* of the full α.

Now item (7.1) stated above is a "consistency condition on the k-form representatives α_i and α_j where regions U_i and U_j overlap. This consistency condition involves the embedding functions for these same two regions going to the V_i and V_j cushions around **x**.

As other notation

$\Omega^k(M)$ = collection of k-forms on the manifold M

Now suppose we have these representative forms in a region of overlap between i and j.

$$\alpha_i = \Sigma_I f_I dt_I$$
$$\alpha_j = \Sigma_J g_J dt_J$$

What then does (7.1) shown above say?

$$\alpha_j = (\psi_i^{-1} \circ \psi_j)^* \alpha_i$$

$$\Sigma_J g_J dt_J = (\psi_i^{-1} \circ \psi_j)^* \Sigma_I f_I(t) dt_I$$

Lets define

$$\varphi \equiv \psi_i^{-1} \circ \psi_j$$

Then we have (see page 15 of raw Ch 3 notes or page 42 B)

$$\alpha_j = (\psi_i^{-1} \circ \psi_j)^* \alpha_i = \varphi^* \alpha_i$$

or

$$\Sigma_J g_J dt_J = \varphi^* (\Sigma_I f_I(\mathbf{t}) dt_I) = \Sigma_I f_I(\varphi(\mathbf{t})) d\varphi_I = \Sigma_{I,J} f_I(\varphi(\mathbf{t})) \det(D\varphi)_{I,J} dt_J$$

Comparing the two sides we get

$$g_J = \Sigma_I f_I(\varphi(\mathbf{t})) \det(D\varphi)_{I,J} = \Sigma_I \varphi^* f_I(\mathbf{t}) \det(D\varphi)_{I,J} . \qquad (**)$$

Write this out then as

$$g_J = \Sigma_I (\psi_i^{-1} \circ \psi_j)^* f_I \det[D(\psi_i^{-1} \circ \psi_j)]_{I,J}$$ which is page 82A

So this relates the component functions g_I of the form representative α_j to those f_I of α_i. This is what Sja means by "components". I might call these things coefficient functions of the differentials in the form.

So I arrive at the end of Section 7.1, I am still standing but slightly dazed. What has happened here? We deal with a manifold as being covered by U_i sets in the flatspace [so V_i intersect M in the M space] . In each U_i part of the problem we have a form α_i. Since regions i and j overlap, we have strange consistency conditions for points in such overlap which have the form of pullbacks. The point is that we have a nice embedding ψ_i well defined in our little portion i of the problem. One hopes that eventually there will be some examples.

Comments: According to this "first definition" of "a form on a manifold", you have a general form α on a manifold M and it is *defined* in terms of its pullbacks α_i on each covering patch U_i. As noted above, $\alpha_i = \psi_i^*\alpha$ where ψ_i is the transformation or embedding associated with patch U_i. This certainly seems a reasonable approach to me. We even have an expression above $g_J = \Sigma_I f_I(\varphi(\mathbf{t}))\, \det(D\varphi)_{I,J}$ showing how the form coefficients of α_j are related to those of α_i, where $\varphi \equiv \psi_i^{-1} \circ \psi_j$.

7.2 Second definition

2.15.16: In this "second definition" of "a form on a Manifold", a form is *identified with* an element of the dual wedge space which I call Λ^k in wedge doc. This of course leads to a world of alternating functions, but I am not clear how this association "helps me do something useful".

This is a **Grassman** theory section and I just know it is going to be brutal. The wedge product is going to appear, along with alternating multi-linear functions, that whole ball of wax is coming down the pike!

[page 83]

2.15.16: Thus stuff was totally strange to me on first reading. Now that I have written wedge doc, it is all I think pretty well understood.

The opening half page is rather shocking. It seems to be an alternate forms universe as presented by Grassmann. The key new idea is this

$dx_i \equiv e_i^T = (0,0,...1,..0)$ basic form dx_i = a unit vector

This is a *totally mysterious identification* of a differential distance with a vector, which is really an identification of dx^i with λ^i which is a basis vector of the dual wedge space $\Lambda^1 = L^1$. The reader has to forget that dx_i in some other World is a differential. Here it is just a *symbol* identified with e_i^T.

Then we write a 1-form as

$\alpha = \Sigma_i f_i\, dx_i = \Sigma_i f_i\, e_i^T = (f_1,f_2....f_n)$ = a row vector of functions (expanded on the dual basis)

So now a 1-form is identified with a vector of that dual space! The following gray was written before wedge doc and I have corrected it where I can changing $\otimes$ to ^ .

I will just *guess* (this guess was not quite right!) for a 2-form something like this

$$\alpha = \Sigma_{ij}\, f_{ij}\, dx_i\, dx_j = \Sigma_i\, f_{ij}\, [e_i^T \wedge e_i^T] = \text{a rank-2 tensor of functions, such as in my tensor doc}$$

Actually, I will refine this a bit. Recall from tensor doc that one useful tensor expansion of tensor A is this

$$A = \Sigma_{ijk}\, A^{ijk}\, (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k) \quad . \qquad \text{(E.2.6)}$$

where A^{ijk} are contravariant tensor components. But in Cartesian space up and down don't matter, so they are also covariant components. So maybe the 2-form will be

$$\alpha = \Sigma_i\, f_{ij}\, dx_i\, dx_j = \Sigma_{ij}\, f_{ij}\, \mathbf{e}_i \wedge \mathbf{e}_j$$

where f_{ij} are covariant rank-2 tensor components, and these e_i are those of Sjamaar, not the tangent base vectors of tensor doc. This is just a conjecture right now, and I don't know about the transpose things.

Any way, Sja is reporting out f = $(f_1, f_2 \ldots f_n)^T$ as a covariant vector = **covector**. I think in general, any vector in any dual space is called a covector.

So in this completely new approach, a 1-form is a covariant vector, or covector for short. I suppose this is not too different from the "association" we earlier discussed between 1-forms α and vector fields **F**.

Another definition I think I have seen somewhere: (yet another meaning of the * symbol!)

$(R_n)^*$ = space of constant row vectors = **the dual** of R^n

Side Note: If **g** is a vector function we can write

$$d\mathbf{g} = \Sigma_i(\partial_i \mathbf{g})\, dx_i = (\mathbf{Dg})\, d\mathbf{x} = (\partial_1 \mathbf{g}, \partial_2 \mathbf{g} \ldots . \partial_n \mathbf{g}) = \text{the usual Jacobian}$$

where we have used the new fact that $dx_i \equiv e_i^T$.

The Dual Space Concept

I think this is something I already know a little about from tensor doc. Recall there from page 75:

Notes on Reciprocity (Duality)

1. Suppose some set of vectors $\mathbf{b}_n$ forms a complete basis for x-space. Can one find a set of vectors $\mathbf{B}_n$ that have the property

$\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$? // duality relation; $\mathbf{B}_a$ and $\mathbf{b}_a$ are reciprocal (6.2.8)

As shown below, the answer is normally "yes", and the vectors $\mathbf{B}_n$ are uniquely determined by the $\mathbf{b}_n$. One says that the set $\{\mathbf{B}_n\}$ is "dual to" the set $\{\mathbf{b}_n\}$ and vice versa. If we regard $\mathbf{b}_n$ as a basis, then $\mathbf{B}_n$ is the "dual basis", and $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is the "duality relation". Another terminology is that the vectors $\mathbf{B}_n$ are "reciprocal to" the vectors $\mathbf{b}_n$ and vice versa.

3. One can solve for the $\mathbf{B}_n$ in terms of the $\mathbf{b}_n$. Each $\mathbf{B}_n$ has N components, so there are N^2 unknowns. The duality relation $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is a set of N^2 equations. This is basically a Cramer's Rule problem in N^2 variables. Since the $\mathbf{b}_n$ form a complete basis, one can expand $\mathbf{B}_m$ on the $\mathbf{b}_n$ with some coefficients we will call w'_{mn} (at this point w'_{mn} is unknown),

$$\mathbf{B}_m = w'_{mn}\mathbf{b}_n \; . \tag{6.2.9}$$

Then from (6.2.8),

$$\delta_{m,k} = \mathbf{B}_m \bullet \mathbf{b}_k = w'_{mn}\, \mathbf{b}_n \bullet \mathbf{b}_k \; . \qquad // \quad \mathbf{b}_n \bullet \mathbf{b}_k = \bar{g}_{ij}(b_n)_i(b_k)_j \; . \tag{6.2.10}$$

Define matrix W' by,

$$W'_{nk} \equiv \mathbf{b}_n \bullet \mathbf{b}_k \tag{6.2.11}$$

and note that W'_{nk} is symmetric. Then (6.2.10) says

$$\delta_{m,k} = w'_{mn}W'_{nk} \quad \text{or} \quad w'W' = 1 \quad \text{or} \quad w' = W'^{-1} \; . \tag{6.2.12}$$

Assuming for the moment that $\det W' \neq 0$, the solution is given by $w' = W'^{-1}$. The (5.6.4) "Digression" showed that $(A^{-1})^T = (A^T)^{-1}$ for invertible A, so $w'^T = (W'^{-1})^T = (W'^T)^{-1} = W'^{-1} = w'$ and therefore w' is symmetric as well. Since W' is known from (6.2.11), $w' = W'^{-1}$ and the $\mathbf{B}_m = w'_{mn}\mathbf{b}_n$ of (6.2.9) have been found. Finally,

$$\mathbf{B}_m \bullet \mathbf{B}_n = w'_{mi}\mathbf{b}_i \bullet w'_{nj}\mathbf{b}_j = w'_{mi}w'_{nj}\, \mathbf{b}_i \bullet \mathbf{b}_j = w'_{mi}w'_{nj}W'_{ji} = w'_{mi}\delta_{ni} = w'_{nm} \; . \tag{6.2.13}$$

So in this tensor doc context, I can regard the $\mathbf{b}_i$ = Sja $\mathbf{v}_i$ as some basis vectors in the vector space spanned by the $\mathbf{b}_i$, and then I can regard the $\mathbf{B}_i$ = Sja $\boldsymbol{\lambda}_i$ as a set of row vectors that span the "row space".

You can think of each row vector application as a *linear functional* (which is a special case of a linear mapping)

$\mathbf{B}_m \bullet \mathbf{x}$ = number $\qquad B_m : \; R^n \rightarrow R \qquad \mathbf{x}$ in R^n

$\boldsymbol{\lambda}_m \bullet \mathbf{x}$ = number $\qquad \lambda_m : \; R^n \rightarrow R$

In fact any row vector acts as a functional in this sense, not just the basis vectors. Fine.

$\mathbf{y} \bullet \mathbf{x}$ = number y : $R^n \rightarrow R$ $\mathbf{y}$ in R^{n*} $\mathbf{x}$ in R^n

Sja wants to refer to the vector space of column vectors $\mathbf{x}$ as V = subset of R^n spanned by the $\mathbf{b}_j = \mathbf{e}_j$.

Sja wants to refer to the vector space of row vectors $\mathbf{y}$ as V* = subset of R^{n*} spanned by the $\mathbf{b}_j = \mathbf{e}_j$.

Then the space V* is "dual to" V, and V* is "the dual space". This all seems quite innocent.

[page 84] 2.15.16 review OK to here

Example 7.2 . Here V is a space of real functions (not just a piece of R^n) in the Stakgold sense defined on [a,b]. A functional of interest is $\mu[f] = \int f\,dx$ which gives you a real number. Somehow I*f = integral value. Sja is only here giving an example of a linear functional, he says nothing any "dual space" to V. [but below I show that this particular functional is just one of a infinite number of such in V*]

2.15.16: This little example does make an association between a dual space functional μ and something involving dx. That is a Big Issue I am trying to understand right now: why you are allowed to associate the basis function λ^i with the differential dx^i. You could write $\mu[f] = \int \Sigma_i\, f_i dx^i = \int \alpha$ and then your functional is the integral of α. But Sja is identifying α itself with a functional somehow!

Example 7.3. In this example, the functional is $\mu[\mathbf{x}] = \mathbf{v} \bullet \mathbf{x}$ for some fixed $\mathbf{v}$ where now we have the "standard inner product" used in R^n. I have been using this notation all along, while Sja has been avoiding it.

Comment: Recall that scalar products exist in Hilbert Spaces which are much more specific that a generic vector space. Sja wants to think in terms of the more general vector space idea, so Sja is not going to use any inner products.

Now, suppose $\mathbf{v}_i$ are the basis vectors of V of dimension n. ($\mathbf{b}_i$ in tensor doc). Sja expands a general vector $\mathbf{v}$ as [Sja often uses $\mathbf{v}_i$ as a basis function instead of my preferred $\mathbf{e}_i$]

$$\mathbf{v} = \Sigma_j\, c_j\, \mathbf{v}_j$$

where $\mathbf{v}_i$ is a supposed basis for V. Now *define* a covector $\lambda_j(\mathbf{v})$ in the dual space according to

$$\lambda_j(\mathbf{v}) \equiv c_j$$

This λ_j is a linear functional on V (different functional for each j value). The functional is defined so it equals the coefficient that appears for $\mathbf{v}_j$ in the expansion of general V vector $\mathbf{v}$. So then

$\mathbf{v} = \Sigma_j \lambda_j(\mathbf{v})\, \mathbf{v}_j \qquad\qquad c_j = \lambda_j(\mathbf{v})$

Notice that no dot product was used.

So this object $\boldsymbol{\lambda}(\mathbf{v})$ is a row vector $\mathbf{c}^T$ of constants c_j which when linearly combined with the $\mathbf{v}_j$ gives some vector $\mathbf{v}$. Using dot products I would say then that

$\lambda_j(\mathbf{v}) = \mathbf{v} \bullet \mathbf{v}_j$ and then $\mathbf{v} = \Sigma_j\, (\mathbf{v} \bullet \mathbf{v}_j)\, \mathbf{v}_j$.

But his definition does not require the existence of an inner product, it just requires a vector space. Working again with his definition, it seems clear that

$\lambda_j(\mathbf{v}_j) = 1$
$\lambda_j(\mathbf{v}_i) = 0 \qquad i \neq j$

and you arrive at this just looking at the linear combination. For example,

$\mathbf{v}_i = \Sigma_j \lambda_j(\mathbf{v}_i)\, \mathbf{v}_j \quad \Rightarrow \quad \lambda_j(\mathbf{v}_i) = \delta_{j,i}$

So we then have this set of **coordinate functions** λ_i and each of these is a linear functional mapping $R^n \rightarrow R$ such as $\lambda_j(\mathbf{v}) = c_j$. [Below I show the these λ_i really are linear.]

OK, so I think the idea is to try to avoid talking about dot products to be general to any vector space, not just a Hilbert space.

2.15.16: The above Sja notes seem confusing. Now I would write $\lambda_j(\mathbf{v}) = v_j$ as the action of λ on a vector, it projects out the jth coordinate. But above I write $\lambda_j(\mathbf{v}) \equiv c_j$. I think this was a confusion on my part and Sja probably never says what I write above.

Digression: Wiki and Stakgold on Dual Spaces

I looked at wiki at this point. It says that the space V* is the space of all linear functionals on V. So let's look again at the above examples. When $V = R^n$ which is a finite dimensional space, I think this is the most general linear functional you can write:

$\mathbf{v} = \Sigma_{j=1}^n c_j\, \mathbf{v}_j = \Sigma_j \lambda_j(\mathbf{v})\, \mathbf{v}_j \qquad\qquad c_j = \lambda_j(\mathbf{v})$

Now exactly what is the linear functional here? I would say that *each* of the λ_j functions is a linear functional. I show linearity below. So for each j we have $\lambda_j: R^n \rightarrow R$. Thus, we are aware right away of a set of n different functionals on V. Below we show that these λ_j functionals form a basis for V* and therefore V* has dimension n. Wiki would say that the elements of V* are called covectors, so therefore each of the λ_j functions would be a covector. Wiki allows that $\lambda_j(\mathbf{v}) = [\lambda_j, \mathbf{v}] = \langle \lambda_j, \mathbf{v} \rangle$ are other common notations.

Wiki says for finite n V, you can take $\mathbf{e}_i$ as the V basis and then $\mathbf{e}^i$ is the V* basis where $e^i e_j = \delta_{ij}$, so this is the same as my tensor doc idea. The dual space is the space of row vectors. Wiki goes on to talk about V being ∞ dimensional, but does not really address any space of functions. But:

Suppose V instead of being R^n is some infinite dimensional function space. Consider $\mu[f] = \int f\,dx$ where f lies in V. Here μ is a linear functional and we have $\mu: V \rightarrow R$. So what is the dual space V* in this case? I have displayed only one linear functional, but V* is the space of *all* such linear functionals. So I don't have an answer here. Notice that $\mu[f] = \int f^\alpha\,dx$ for $\alpha \neq 1$ would not be a *linear* functional. But the form $\mu[f] = \int f\,dx\,s(x)$ might also be a linear functional where s(x) is a weight function. Stak talks about this subject. [Below we have Riesz who says each linear functional is associated with some element of V. So my linear functional just stated is associated with s(x) of V. You would say $\mathbf{f} \bullet \mathbf{s} = \int f(x)\ s(x)\,dx$.]

In Stakgold, a functional is called T[**x**] where **x** lies in V. When V is a Hilbert Space, you can state the famous **Riesz Representation Theorem** Stak chap 2.

" Any continuous linear functional which acts on any S in any Hilbert space V can be "represented" as T[x] = <x,f> (for all x in V) where f is some unique element of V "

Here we have **x** being some vector in V, and the theorem claims that any linear functional can be expressed as the inner product of **x** with some other unique element of V which above is called f. Thus, the number of linear functionals would be the same as the number vectors **x** in V. For $V = R^n$ both these numbers are infinite, and you can think of f as being the row vectors. For a function space where **x** is a function, I might restate this theorem as

" Any continuous linear functional which acts on any S in any function Hilbert space V can be "represented" as T[f] = <f,F> (for all f in V) where F is some unique function of V " . Then

$$T[f] = \int f(x)\,F(x)\,dx = \langle f,F \rangle$$

Then the example above happens to use the function F(x) = 1. So then my conjecture of weight functions would be correct.

Note that Riesz is in the context of a Hilbert Space, not a simple vector space.

Now back to Sjamaar:

Lemma 7.4. The λ_i form a basis for V*.

Now what does λ_i mean without an argument? It is a *function* so can only exist in a space of functions.

[I now refer to this as a functional.] He would have then to show that if λ were an arbitrary function, you could write $\lambda = \Sigma\, d_i\, \lambda_i$ where d_i are some constant coefficients. If we could construct d_i, then maybe we have our proof. So evaluate this arbitrary function λ for one of the V-space basis vectors

$$\lambda(\mathbf{v}_j) = \Sigma\, d_i\, \lambda_i(\mathbf{v}_j) = \Sigma\, d_i\, \delta_{i,j} = d_j \ .$$

So there you have d_j constructed. We then have

$$\lambda = \Sigma\, \lambda(\mathbf{v}_j)\, \lambda_i$$

and we have then shown that λ_i is a basis since it can generate any function λ by linear combination. So I guess this looks good, and again we have avoided mention of any inner product. If you had an inner product, you could write $\lambda_j(\mathbf{v}) = \mathbf{v} \bullet \mathbf{v}_j$ but we don't want to assume an inner product, so we just have these functions λ_j. Fine by me.

Because the set of n functions λ_i spans V*, we must have dim(V*) = n. So I have just shown that:

Fact : The V* basis of *functions* $\{\lambda_i\}$ is dual to the basis of *vectors* $\{\mathbf{v}_j\}$ of V.

Note that λ_i is a scalar function of a vector just as $\lambda_j(\mathbf{v}) = \mathbf{v} \bullet \mathbf{v}_j$ would suggest were there an IP. Note that the functions λ_i are in fact *functionals* since they map to R. [correct]

Are the λ_i *linear* functionals? Let's investigate:

$$\mathbf{v} = \Sigma_j\, \lambda_j(\mathbf{v})\, \mathbf{v}_j$$

$$(\alpha\mathbf{v}) = \Sigma_j\, \lambda_j(\alpha\mathbf{v})\, \mathbf{v}_j = \alpha\, [\, \Sigma_j\, \lambda_j(\mathbf{v})\, \mathbf{v}_j\,] = \Sigma_j\, (\alpha\, \lambda_j(\mathbf{v}))\, \mathbf{v}_j = \Sigma_j\, (\lambda_j(\alpha\mathbf{v}))\, \mathbf{v}_j$$

Since the coefficients are unique, we have

$$\lambda_j(\alpha\mathbf{v}) = \alpha\, \lambda_j(\mathbf{v}) \qquad \text{first part of showing linear.}$$

Next,

$$\mathbf{v} + \mathbf{v}' = \Sigma_j\, [\lambda_j(\mathbf{v}) + \lambda_j(\mathbf{v}')]\, \mathbf{v}_j = \Sigma_j \lambda_j(\mathbf{v}+\mathbf{v}')\, \mathbf{v}_j \qquad \Rightarrow$$

$$\lambda_j(\mathbf{v}+\mathbf{v}') = \lambda_j(\mathbf{v}) + \lambda_j(\mathbf{v}')$$

So YES, the functions λ_i are linear functions and are linear functionals. No dot product needed.

Example 7.5. Let $\mathbf{e}_i$ be the basis vectors of V, as usual. According to our fiat statement above, we know that $dx_i = \mathbf{e}_i{}^{\mathbf{T}}$ which lies in the dual space R^{n*} . It seems pretty reasonable that the $\{\mathbf{e}_i{}^{\mathbf{T}}\}$ form a basis of R^{n*} and in fact

$\mathbf{e}_i^T \mathbf{e}_j = e_i \bullet e_j = \delta_{i,j}$

so then we can say: " the set $\{dx_i\}$ form a basis in R^{n*}". I guess one could write

$\mathbf{v} = \Sigma\, v_i \mathbf{e}_i$ general vector in R^n
$v^T = \Sigma\, v_i \mathbf{e}_i^T$ general vector in R^{n*}

So in this example you can really associate the dual space with simple row vectors, not so much abstract functions. We have just shown that

Fact: The set $\{dx_i\}$ form a basis in R^{n*} where R^n is the usual space of vectors $\mathbf{v}$.

Comment: At this point we are claiming that the differentials dx_i form a basis for the space V* which is dual to the space V which lies in R^n. In this world, we have never used the fact that dx_i is a differential, it is just a symbol so far! I agree, it is interesting all in itself to come of with some kind of vector space in which the dx_i are "basis vectors". Then are really this:

$dx_i(\mathbf{v})$ = linear functional on R^n where $dx_i(\mathbf{v}) = e_i^T v = \mathbf{e}_i \bullet \mathbf{v} = v_i$

Here I define $\bullet$ as shown by the last equal sign, forget Hilbert Space. So we end up with

$dx_i(\mathbf{v}) = v_i$ linear functional dx_i

but Sja does not write down this fact anywhere.

Comment: Notice that so far, there is no meaning to something like $dx_i dx_j$ other than it is the product of two linear functionals, and you could then say $dx_i dx_j$ means $dx_i(\mathbf{v})dx_j(\mathbf{v}) = v_i v_j$. This is NOT the meaning for $dx_i dx_j$ that will appear below as the wedge product. For example, if we use the meaning of this paragraph, we would get $dx_i(\mathbf{v})dx_i(\mathbf{v}) = v_i v_i$ whereas we know that $dx_i dx_i = 0$ as a wedge.

Final random item on page 84:

Suppose we have $L: V \to W$ where V has basis $\mathbf{v}_i$ and W has basis $\mathbf{w}_i$. Consider

Lv_j = some vector in W $= \Sigma_i\, \ell_{i,j}\, \mathbf{w}_i$ defines some numbers $\ell_{i,j}$

Again, no inner product yet.

Lemma 7.6. Imagine this situation for $L: V \to W$

$\mathbf{v}_i$ = basis in V
λ_i = basis for V*

$\mathbf{w}_i$ = basis for W

μ_i = basis for W*

Then compute

$$\mu_i(Lv_j) = \mu_i(\Sigma_k \ell_{k,j} w_k) = \Sigma_k \ell_{k,j} \mu_i(w_k) = \Sigma_k \ell_{k,j} \delta_{i,k} = \ell_{i,j}$$

Not sure why this is interesting, but it is correct.

[page 85] 2.15.16 review OK to here

Definition: Sja defines the simple property of a **k-multilinear function** $\lambda(\mathbf{v_1},\mathbf{v_2}.....\mathbf{v_k})$ This is $\lambda : V^k \rightarrow R$, so there are k distinct vector arguments. If there are two arguments, it is bilinear = 2-multilinear. Some simple examples are given. [these are in fact multilinear *functionals* since they map to R]

Comment: I think this is the obvious generalization of a linear functional of one vector variable to many vector variables in the same space V. Of course $V^k = V \otimes V \otimes .. \otimes V$ which is a direct product space as in my tensor doc discussion.

More: A multilinear function is linear on each argument separately. In my own notation here is an example where we consider four arguments so we have $\lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4})$. Below I show linearity just for the third argument, but the same equations would apply an each argument separately.

$$\lambda(\mathbf{v_1},\mathbf{v_2}, \alpha\mathbf{v_3}, \mathbf{v_4}) = \alpha\, \lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}) \qquad \text{// scalar rule}$$

$$\lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{u} + \mathbf{w}, \mathbf{v_4}) = \lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{u}, \mathbf{v_4}) + \lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{w}, \mathbf{v_4}) \qquad \text{// addition rule}$$

Sja likes to combine these into a single rule like this

$$\lambda(\mathbf{v_1},\mathbf{v_2}, \alpha\mathbf{u} + \beta\mathbf{w}, \mathbf{v_4}) = \alpha\lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{u}, \mathbf{v_4}) + \beta\lambda(\mathbf{v_1},\mathbf{v_2}, \mathbf{w}, \mathbf{v_4}) \qquad \text{page 85A}$$

I find his notation more confusing than mine.

PL Exercise. What could you say about

$$\lambda(M\mathbf{v_1},M\mathbf{v_2}, M\mathbf{v_3}, M\mathbf{v_4})$$

where M is a matrix? Perhaps write

$$M\mathbf{v_1} = M\,[\Sigma_i(\mathbf{v_1})_i \mathbf{e}_i] = \Sigma_i(\mathbf{v_1})_i\, M\mathbf{e}_i$$

Now $M\mathbf{e}_i$ is a vector, so I have written $M\mathbf{v_1}$ as a linear combination of vectors. Then we can say

$$\lambda(M\mathbf{v_1},M\mathbf{v_2}, M\mathbf{v_3}, M\mathbf{v_4}) = \Sigma_i(\mathbf{v_1})_i\, \lambda(M\mathbf{e}_i,M\mathbf{v_2}, M\mathbf{v_3}, M\mathbf{v_4})$$

Then you could do this with each argument to get

$$\lambda(M\mathbf{v}_1, M\mathbf{v}_2, M\mathbf{v}_3, M\mathbf{v}_4) = \Sigma_{i_1 i_2 i_3 i_4} (\mathbf{v}_1)_{i_1}(\mathbf{v}_2)_{i_2}(\mathbf{v}_3)_{i_3}(\mathbf{v}_4)_{i_4} \lambda(M\mathbf{e}_{i_1}, M\mathbf{e}_{i_2}, M\mathbf{e}_{i_3}, M\mathbf{e}_{i_4})$$

There is not much more one can do with this structure.

Definition: If you add the property that swapping any two arguments negates λ, then you have an **alternating or asymmetric** k-multilinear function.

The big example here is $\det(\mathbf{c}_1, \mathbf{c}_2 .... \mathbf{c}_k)$ is a well-known alternating k-multilinear function where $\mathbf{c}_i$ are the column vectors of the matrix whose determinant we are expressing. Sja has discussed this animal earlier, and there I broke it down into two properties, each of which determinants have. He gives a second simpler alternating example that is bilinear.

Now Sja does a very weird thing. He defines a single function which has the following name:

"$\lambda_1 \lambda_2 .... \lambda_k$"

The arguments of this function are going to be $\mathbf{v}_1, \mathbf{v}_2 ... \mathbf{v}_k$ which are the basis vectors of V. And here is how the function is defined:

$$\lambda_1 \lambda_2 .... \lambda_k(\mathbf{v}_1, \mathbf{v}_2 ... \mathbf{v}_k) \equiv \det[\lambda_i(\mathbf{v}_j)]$$

The matrix here is this:

$$\begin{matrix} \lambda_1(\mathbf{v}_1) & \lambda_1(\mathbf{v}_2) & \lambda_1(\mathbf{v}_3) & ..... & \lambda_1(\mathbf{v}_k) \\ \lambda_2(\mathbf{v}_1) & \lambda_2(\mathbf{v}_2) & \lambda_2(\mathbf{v}_3) & ..... & \lambda_2(\mathbf{v}_k) \\ ... & & & & \\ \lambda_k(\mathbf{v}_1) & \lambda_k(\mathbf{v}_2) & \lambda_k(\mathbf{v}_3) & ..... & \lambda_k(\mathbf{v}_k) \end{matrix}$$

This fancy function is a k-multilinear function because that is the way determinants work. It is also alternating since that is *also* how they work.

This function $\lambda_1 \lambda_2 .... \lambda_k$ is called "the **wedge product** of the λ's" function and one writes then

$$[\lambda_1 \wedge \lambda_2 \wedge .... \wedge \lambda_k](\mathbf{v}_1, \mathbf{v}_2 ... \mathbf{v}_k) \equiv \det[\lambda_i(\mathbf{v}_j)] \qquad (*)$$

I see the mechanical definitions, but I have *no idea* how this relates to anything I know about forms.
Make this clearer:

$$\lambda_1 \wedge \lambda_2 \wedge .... \wedge \lambda_k = \text{"}\lambda_1 \lambda_2 .... \lambda_k\text{"} = \text{the name of a certain alternating k-multilinear functional}$$

Question: Can we identify $\lambda_i(\mathbf{v}_j) = \delta_{i,j}$ inside the determinant shown in (*)?
Answer: Well, in the current context, the $\mathbf{v}_i$ are arbitrary arguments of the multilinear function and are general vectors in V. They could be, but are in general *not*, the set of basis vectors. Only if we happened to evaluate the wedge thing at the basis vectors would be make such a claim.

Now:

A^kV = the complete set of alternating k-multilinear functions. [= my space Λ^k_f or just Λ^k]

This A^kV is in fact a vector space, and as such, it should have some basis.

Comment: Recall that V* is the collection of linear functionals on V -- the dual space. For finite V with basis *vectors* $\mathbf{v}_i$ we found that this dual space is spanned by the set of *functionals* λ_i. Each of these functionals is a function of **v** where **v** lies in V.

Now we generalize so that V is replaced by $V^k = V \otimes V \otimes .. \otimes V$. [nope] What is the corresponding dual space? Perhaps it is the set of *multilinear* functionals on V^k. I don't know how many there are right now, but write *one* of these as $\mu(\mathbf{v}_1, \mathbf{v}_2.....\mathbf{v}_k)$. I presume then that the set of all possible such multilinear functionals would be the dual space $(V^k)^*$. That would seem to be the logical generalization.

For reasons not yet clear, Sja wants us to consider only multilinear functionals μ which have the *alternating property*, and instead of $(V^k)^*$ he refers to this restricted space as AV^k. So before the reasons for being interested in this sort of dual space are given, Sja wants to develop some properties of this space, and only then can he explain the motivation.

Sja is not referring to either $(V^k)^*$ or AV^k as a "dual space", but either seems a reasonable generalization of the dual space concept.

Naively you might think that

$(V^k)^* = V^* \otimes V^* \otimes .. \otimes V^*$

and then maybe the basis vectors of $(V^k)^*$ would be

$\lambda_i \otimes \lambda_j \otimes .....\otimes \lambda_k$ i = 1..k, j = 1..k etc

I will come back to this idea later. This approach however does not work in that alternating property. But maybe this would be the normal dual space of $(V^k)^*$. I think it is.

2.15.16: In these original notes above I am confusing the tensor and tensor dual space with the wedge and wedge dual space. I now understand how all this works.

Properties of the space AV^k.

1. It is a vector space. Well, $(V^k)^*$ is a vector space since it is a dual space, and if we restrict to multilinear functionals which have the alternating property, we still get a vector space. Key fact is that if you add two alternators, the sum will be alternating.

2. As a vector space, AV^k must have some basis vectors and must have some dimension.

[page 86] 2.15.16 review OK to here

Define I as our usual increasing multi-index.

Example 7.11. Here $V = R^3$ with the usual $\mathbf{e}_i$ basis vectors and the usual e_i^T dual space basis vectors. Recall from page 83 that we identify $e_i^T = dx_i$. So I guess dx_i is a basis vector in V*. In general, a dual space basis vector was the function λ_i so I guess we can identify

$\lambda_i = e_i^T = dx_i$ for i = 1..3. // my Big Mystery, why is this allowed?

Now I is an ordered multi-index with k components. The space is $V = R^n$ so k is some integer $\leq n$.

Suppose I = (1,2) so we are working with k = 2. Then we can say,

dx_I = "$dx_1 dx_2$" = $dx_1 \wedge dx_2$ = the 2-multilinear functional as shown in page 85 C

We are supposed to regard "$dx_1 dx_2$" as the function name in the sense of page 85 C. Then we can evaluate this functional $dx_I(\mathbf{v}_1, \mathbf{v}_2)$ for the following specific arguments:

$$dx_I(\mathbf{e}_1, \mathbf{e}_2) = \det(\lambda_i(e_j) = \begin{vmatrix} \lambda_1(\mathbf{e}_1) & \lambda_1(\mathbf{e}_2) \\ \lambda_2(\mathbf{e}_1) & \lambda_2(\mathbf{e}_2) \end{vmatrix} = \begin{vmatrix} dx_1(\mathbf{e}_1) & dx_1(\mathbf{e}_2) \\ dx_2(\mathbf{e}_1) & dx_2(\mathbf{e}_2) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Note that $dx_i(\mathbf{e}_j) = \delta_{ij}$ as we know from earlier, hence the rightmost two = expressions above.

Now Sja has defined a new object here without saying so. Namely

$\mathbf{e}_I \equiv (\mathbf{e}_1, \mathbf{e}_2)$ I = (1,2) in our example.

Then we have $dx_I(\mathbf{e}_I) = dx_I(\mathbf{e}_1, \mathbf{e}_2)$ so the above line becomes

$$dx_I(\mathbf{e}_I) = \det(\lambda_i(e_j) = \begin{vmatrix} \lambda_1(\mathbf{e}_1) & \lambda_1(\mathbf{e}_2) \\ \lambda_2(\mathbf{e}_1) & \lambda_2(\mathbf{e}_2) \end{vmatrix} = \begin{vmatrix} dx_1(\mathbf{e}_1) & dx_1(\mathbf{e}_2) \\ dx_2(\mathbf{e}_1) & dx_2(\mathbf{e}_2) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

This is what was confusing me but now seems just fine.

Now he says let's take another multi-index J = (2,3) so then

$$dx_I(\mathbf{e}_J) = dx_I(\mathbf{e}_2, \mathbf{e}_3) = \begin{vmatrix} dx_1(\mathbf{e}_2) & dx_1(\mathbf{e}_3) \\ dx_2(\mathbf{e}_2) & dx_2(\mathbf{e}_3) \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = 0$$

Notice how you insert the vectors in the argument list of dx_I into the determinant based on position.

In this example, notice how we found this result

$dx_I(\mathbf{e}_J) = \delta_{I,J}$ = 1 only if the two k=2 multi-indices are the same.

He will now generalize this result:

Lemma 7.12. The claim here is in fact that [2.15.16: old hat now when written using λ^I]

$dx_I(\mathbf{e}_J) = e_I^T(\mathbf{e}_J) = \lambda_I(\mathbf{e}_J) = \delta_{I,J}$ for any two ordered multi-indices I and J of order k

Proof: I agree with 86 A as a displayed determinant. Remember that all the i_r indices are different, and all the j_s indices are different, since part of ordered multi-indices. Thus, for example, j_2 can match only one index in I (at most), and vice versa. I think this means that each row in the matrix can have only a single 1 with the other row entries 0. When I = J, all those 1's are on the diagonal and det = 1. It is not obvious what happens if I ≠J so Sja discusses that in the text. I did not track the proof exactly, but he is showing that if I≠J, then either a whole row or a whole column vanishes, so det = 0.

Lemma 7.13. If $\lambda(v_I)$ vanishes for all ordered multi-indices I of degree k, then $\lambda \equiv 0$.

Question: What is the meaning of $\lambda(v_I)$ here? This is Sja's first use of isolated symbol λ. I guess this is meant to be some arbitrary general element of the space AV^k, some $\lambda(\mathbf{v}_1, \mathbf{v}_2....\mathbf{v}_k)$ k-multilinear functional of generic arguments. But,

$\lambda(v_I) = \lambda(\mathbf{v}_{i1}, \mathbf{v}_{i2}....\mathbf{v}_{ik})$ = λ evaluated at a set of *basis vectors* which are in increasing order

So the claim is that if your generic λ vanishes for all v_I of degree k, then $\lambda \equiv 0$.

Proof: Let $\mathbf{w}_i$ be an *arbitrary* set of vector arguments. If we can show that $\lambda(\mathbf{w}_1, \mathbf{w}_2....\mathbf{w}_k) = 0$ for any such set of vectors, then surely $\lambda \equiv 0$. Sja expands each $\mathbf{w}_i$ on the basis vectors $\mathbf{v}_i$ as shown top p 87, but really he means us to do this:

$\mathbf{w}_j = \Sigma_{i_j} a_{ji_j}\mathbf{v}_{i_j}$ j = 1,2....k

The summation index used for each w_j is different and we denote it by i_j. Then we get

$\lambda(\mathbf{w}_1, \mathbf{w}_2....\mathbf{w}_k) = \lambda(\Sigma_{i_1} a_{1i_1}\mathbf{v}_{i_1}, \Sigma_{i_2} a_{2i_2}\mathbf{v}_{i_2}.... \Sigma_{i_k} a_{ki_k}\mathbf{v}_{i_k})$

$= \Sigma_{i_1}\Sigma_{i_2}....\Sigma_{i_k} a_{1i_1}a_{2i_2}....a_{ki_k} \lambda(\mathbf{v}_{i_1}, \mathbf{v}_{i_2} ....\mathbf{v}_{i_k})$ = page 87 B

Now our premise of Lemma 7.13 is that $\lambda(v_I) = 0$ for all increasing multi-index choices of the arguments. But swapping any pair of vector arguments negates λ, so then we really have $\lambda(\{\mathbf{v}\}) = 0$ for any set of basis vector arguments. If two arguments are the same, we also get 0 due to the negation fact. Thus, the premise implies that $\lambda(\mathbf{v}_{i_1}, \mathbf{v}_{i_2} ....\mathbf{v}_{i_k})$ for any values of the subscripts. Thus $\lambda(\mathbf{w}_1, \mathbf{w}_2....\mathbf{w}_k) = 0$ and then we conclude that $\lambda \equiv 0$.

[**page 87**] 2.15.16 review OK to here

Theorem 7.14. Assume that

V has basis $\{\mathbf{v}_1, \mathbf{v}_2....\mathbf{v}_n\}$ and thus dimension n [the $\mathbf{v}_i$ are of course vectors in V]

V* has basis $\{\lambda_1, \lambda_2....\lambda_n\}$ and thus dimension n [the λ_i functionals on V]

Then the set of functions λ_I where we enumerates all multi-indices I forms a basis for A^kV. [fine]

Note that: $\lambda_I \equiv$ "$\lambda_{i_1}\lambda_{i_2}.....\lambda_{i_k}$" = a function name

My proof: Want to show that you can write $\lambda = \Sigma_I\, c_I\, \lambda_I$ for any λ, so then λ_I is a basis. Evaluate this equation at argument v_J where J is another multi-index. Then we have

$\lambda(\mathbf{v}_J) = \Sigma_I\, c_I\, \lambda_I(\mathbf{v}_J) = \Sigma_I\, c_I\, \delta_{I,J}$ // by Lemma 7.12 !!

$= c_J$

Thus for any λ we can do the expansion on c_I shown and we have constructed the solution coefficients to be $c_I = \lambda(\mathbf{v}_I)$ and easy to show this is a unique solution. So yes, this proof is analogous to 7.4 where we showed that any simple λ could be expanded on the λ_i so λ_i was a compete basis for V*.

PL Example. Suppose V is R^n. We have this fancy space A^kV which is the space of alternating multilinear functionals on V^k. The basis of V* is $\{\lambda_i\} = \{e_i^T\} = \{dx_i\}$. We have just shown that the basis for A^kV is then $\lambda_I = e_I^T = dx_I$ where I is an ordered index having k elements, and where these are shorthands for function *names* like $\lambda_1\lambda_2...$ $= e_1^T e_2^T.... = dx_1dx_2...$ You can write these instead (for example) as

$\lambda_1\lambda_2..\lambda_k. = e_1^T e_2^T....e_k^T = dx_1dx_2...dx_k$

.

$\lambda_1\wedge\lambda_2\wedge...\wedge\lambda_k = e_1^T\wedge e_2^T\wedge....\wedge e_k^T = dx_1\wedge dx_2\wedge...\wedge dx_k$

The most general element of A^kV can then be written

λ = a general element of $A^kV = \Sigma_I\, a_I\, dx_I$ // $= \Sigma_I\, a_I\, \lambda^I$

where a_I is a coefficient that goes with basis element dx_I. We know that $a_I = \lambda(\mathbf{e}_I)$. [How do we know this? Well, look at λ just above and close to get $\lambda(\mathbf{e}_J) = \Sigma_I\, a_I\, dx_I(\mathbf{e}_J) = \Sigma_I\, a_I\, \lambda^I(\mathbf{e}_J) = \Sigma_I\, a_I\, \delta^I{}_J = a_J$. For some reason Sjammar is referring to the k-form normally called α by the name λ. Fine. Now I would write the above as

α = a general element of $\Lambda^k = \Sigma_I\, f_I\, dx_I = \Sigma_I\, f_I\, \lambda^I$ = a general k-form $f_I = \alpha(\mathbf{e}_I)$

In this notation, we are closing the k-form α or λ with a V^k vector e_I to make a tensor function.

Let's rewrite the above showing all arguments:

$\lambda = \Sigma_I \lambda(\mathbf{e}_I)\, dx_I$ ok

$\lambda(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \lambda(\mathbf{e}_I)\, dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$ ok

It would take a lot of effort to write this out using proper subscripts like i_1 and j_1.

Now replace the letter λ by the letter α in a hope to make contact with earlier book sections:

$\alpha = \Sigma_I \alpha(\mathbf{e}_I)\, dx_I$ // $\alpha(\mathbf{e}_I) = f_I$ in notation of Chapter 2 page 17 ok

$\alpha(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \alpha(\mathbf{e}_I)\, dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$ ok

Here all the $\mathbf{v}_i$ and the factors of $\mathbf{e}_I$ like $\mathbf{e}_3$ are elements of V. There is no **x** sitting in this thing yet!

Pause: The above equations "look like" the differential forms we discussed in Chapter 2. But in that Chapter dx_1 was a differential distance, whereas here dx_1 is a basis vector e_1^T in the dual space to R^n. So somehow we have these parallel developments and I suspect there is going to be an isomorphism between the two Worlds. [still looking for this Big Mystery connection on 2.15.16]

Now in page 87 A he simply "tacks on" the **x** label like this:

$\alpha(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \alpha(\mathbf{e}_I)\, dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$

$\alpha_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \alpha_{\mathbf{x}}(\mathbf{e}_I)\, dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$ (*)

$\alpha_{\mathbf{x}} = \Sigma_I \alpha_{\mathbf{x}}(\mathbf{e}_I)\, dx_I = \Sigma_I f_I(\mathbf{x})\, dx_I$ = a k-form of earlier chapters.

This connect $f_I(\mathbf{x}) = \alpha_{\mathbf{x}}(\mathbf{e}_I)$ is explicitly stated by Sja as page 88 B.

Comment: In Chapter 2 we had $\alpha = \Sigma_I f_I\, dx^I$ and we always there regarded f_I as $f_I(x)$, see any example there for example. So what was α in Chapter 2 has become $\alpha_{\mathbf{x}}$ here in Chapter 7. Here **x** is I think a point on the manifold M where we have a tangent space $T_{\mathbf{x}}M$ and perhaps things like e_I are specific to that tangent space and should perhaps be $e_I(\mathbf{x})$. In Chapter 2, **x** was a point in some simple R^n I think.

Comment: In order to make contact with Chapter 2, he has to "tack on" the **x** label as shown. That is to say, in the dual space development, we have f_I above being $\alpha(\mathbf{e}_I)$ and there is no "x label" anywhere. But if I were to write $e_I(\mathbf{x})$, then I would have $f_I = \alpha(\mathbf{e}_I(x)) = \langle\alpha \mid \mathbf{e}_I(x)\rangle = f_I(\mathbf{x})$ and then we have the Chapter 2 x argument. I think I am onto something here.

Bell starting to ring here. This space $A^k V$ certainly looks like the space of k-forms as treated in earlier chapters. We obtain the alternating wedge product rule (which was of course intentionally inserted at the start to obtain that result). We obtain the same multi-index general form for a k-form.

Comment: The LHS general function $\alpha_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$ has k-vector arguments. The coefficient $\alpha_{\mathbf{x}}(\mathbf{e}_I)$ is this same function evaluated at a particular set of k vector arguments.

no confusions to this point, but yes confusion below this point!!

Comment 2.15.16. Perhaps this is the gist of Sja's argument: If you blindly associate dx^i with λ^i, and thus blindly associate a form with an element of Λ^k (the dual wedge space), you arrive at something that looks like the form of Chapter 2. But I still don't see how making this Λ^k connection "buys me anything". How does it help me?

[**page 88**] 2.15.16 review OK to here

Page 88 A is extremely mysterious and will take me several hours to interpret.

I spent two *days* or so on the first half of page 88 and was unable to make full sense of it. See my detailed attempts in doc "Attempt to analyze section Pullbacks re-examined on page 88". My conclusion there is this:

Summary: If I am willing to simply *define* $g_J \equiv \varphi^*\alpha(\mathbf{e}_J)$ as a page 88 definition of g_J, I am able to show that this g_J agrees exactly with the Chapter 3 result for g_J which was $g_J(\mathbf{x}) \equiv \Sigma_I \; f_I[\varphi(x)] \det(D\varphi)_{I,J}$. But the whole business is very cloudy since in Chapter 3 things are functions of $\mathbf{x}$, whereas in Chapter 7 page 88 things are functions of those k multilinear arguments. I am really unable to make a clean connection between the two Worlds. Maybe this will get clarified later. I gave it my best shot today. I could go off and read some other detailed source like Spivak, but that would take many more days I suspect. Maybe Sjamaar will clarify things in upcoming chapters.

Here is how one can start off:

Start with page 87 A and page 88 B which combined say

$$\alpha_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \; \alpha_{\mathbf{x}}(\mathbf{e}_I)\, dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$$

But in the spirit of the **x** and **y** use in the pullback discussion of Chapter 2, replace the above with

$$\alpha_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \; \alpha_{\mathbf{y}}(\mathbf{e}_I)\, dy_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) \qquad (**)$$

or

$$\alpha_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \; f_I(\mathbf{y})\, dy_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) \qquad f_I(\mathbf{y}) = \alpha_{\mathbf{y}}(\mathbf{e}_I)$$

Now we apply a pullback mimicking Chapter 3

$$\varphi^*\alpha_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I \; \varphi^*[\alpha_{\mathbf{y}}(\mathbf{e}_I)]\, [\varphi^* dy_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)] \; . \qquad (*)$$

Now make these assumptions:

(***)

$\varphi^*[\alpha_{\mathbf{y}}(\mathbf{e}_I)] = \alpha_{\boldsymbol{\varphi}(\mathbf{x})}(\mathbf{e}_I)$ // action on the coefficient

$[\varphi^* dy_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)] = dy_I((D\varphi)\mathbf{v}_1, (D\varphi)\mathbf{v}_2, .,,,, (D\varphi)\mathbf{v}_k)$ // action on the differential

Inserting these assumptions into (*) we get

$\varphi^*\alpha_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I\ \varphi^*[\alpha_{\mathbf{y}}(\mathbf{e}_I)]\ [\varphi^* dy_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)]$.

$= \Sigma_I\ \alpha_{\boldsymbol{\varphi}(\mathbf{x})}(\mathbf{e}_I)\ dy_I((D\varphi)\mathbf{v}_1, (D\varphi)\mathbf{v}_2, .,,,, (D\varphi)\mathbf{v}_k)$

Now look back at the template (**) above and make substitutions to write the template as,

$\alpha_{\boldsymbol{\varphi}(\mathbf{x})}((D\varphi)\mathbf{v}_1,(D\varphi)\mathbf{v}_2.....(D\varphi)\mathbf{v}_k) = \Sigma_I\ \alpha_{\boldsymbol{\varphi}(\mathbf{x})}(\mathbf{e}_I)\ dy_I((D\varphi)\mathbf{v}_1,(D\varphi)\mathbf{v}_2.....(D\varphi)\mathbf{v}_k)$

Comparing the last two equations which have the same RHS, we obtain a new equation,

$\varphi^*\alpha_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \alpha_{\boldsymbol{\varphi}(\mathbf{x})}((D\varphi)\mathbf{v}_1,(D\varphi)\mathbf{v}_2.....(D\varphi)\mathbf{v}_k)$

I can next make this definition:

$(\varphi^*\alpha)_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) \equiv \varphi^*\alpha_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$ (****)

and then the previous equation becomes

$(\varphi^*\alpha)_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \alpha_{\boldsymbol{\varphi}(\mathbf{x})}((D\varphi)\mathbf{v}_1,(D\varphi)\mathbf{v}_2.....(D\varphi)\mathbf{v}_k)$ // page 88A

Thus, with the assumptions (***) and (****) made above, I have successfully derived page 88A. It purports to tell you "how to do a pullback" in our new world of multilinear functions.

At this point, without specializing the argument set $(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$, there is nothing I can do with any of the above objects having those $(D\varphi)\mathbf{v}_i$ type arguments.

If we do specialize to the case $(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = (\mathbf{e}_{j1},\mathbf{e}_{j2}...\ \mathbf{e}_{jk}) \equiv \mathbf{e}_J$ where J is an ordered index, THEN I am able in the mentioned doc to show that

$(\varphi^*\alpha)_{\mathbf{x}}(\mathbf{e}_J) = \Sigma_I\ f_I(\boldsymbol{\varphi}(\mathbf{x}))\ \det(D\varphi)_{I,J}$ $\qquad f_I(\mathbf{y}) = \alpha_{\mathbf{y}}(\mathbf{e}_I)$

This is the g_J object which appeared in Chapter 3. So in Chapter 7 if we *define*

$g_J \equiv (\varphi^*\alpha)_{\mathbf{x}}(\mathbf{e}_J)$

then the g_J so defined in Chapter 7 agrees with the g_J defined differently in Chapter 3. So I guess Sja is just trying to maintain some connection between the two different Worlds.

[88] Forms on Manifolds 2.15.16 review OK to here

Based on the above discussion, we associate each function in AV^k with a k-form. Recall that the value of x was "tacked on" so we had.

$$\alpha_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I\ \alpha_{\mathbf{x}}(\mathbf{e}_I)\ dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I\ f_I(\mathbf{x})\ dx_I(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k)$$

In this tack-on process, we create label called **x**, and this label is supposed to lie on manifold M. It is true that we can construct a vector space $T_{\mathbf{x}}M$ at this point **x**, whose basis vectors lie on M for very small d**x**. But I don't see why Sja says:

"map $\alpha_{\mathbf{x}}$ on the vector space $T_{\mathbf{x}}M$ " "one for each **x** on M"

I agree that you have an $\alpha_{\mathbf{x}}$ for each such x, and that there is a $T_{\mathbf{x}}M$ at each x, but why "on"? It is really defined on M, not on $T_{\mathbf{x}}M$.

Then he seems to say that if there is a single $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ [fig on page 68 of torus], then the pullback ψ^* acting on a form is given by our result above with general φ replaced with this embedding function ψ, so

$$(\psi^*\alpha)_{\mathbf{t}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \alpha_{\boldsymbol{\psi}(\mathbf{t})}((D\psi)\mathbf{v}_1,(D\psi)\mathbf{v}_2.....(D\psi)\mathbf{v}_k) \qquad \mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) \qquad \text{// p 88 F}$$

But this is really consistent with

$$(\varphi^*\alpha)_{\mathbf{y}}(\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k) = \Sigma_I\ \alpha_{\boldsymbol{\varphi}(\mathbf{x})}(\mathbf{e}_I)\ dy_I(\ (D\varphi)\mathbf{v}_1,\ (D\varphi)\mathbf{v}_2,\ .,,,,,\ (D\varphi)\mathbf{v}_k) \qquad \mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$$

which makes me think he may have mislabeled the earlier thing with **x** instead of **y** or **t**.

But I know what he means. When you do a pullback on α, the above is one way to write the result. In page 88F we have t in U.

It seems odd now that he goes out of his way to call $\psi^*\alpha$ (as shown just above) a "local representative" where we have no local index like the i or j we had before. Maybe this is "the only one". Ah, but I now see that he did throw in the word "local", so he is just omitting labels like i or j.

His BT book quote (Bott and Tu) makes no sense to me because I don't see how our k-form α "eats k-tuples of tangent vectors". Is he saying that the arbitrary arguments $\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k$ are "tangent vectors"? Well, it is true that we have $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ in the M context and that there are tangent vectors at the point **x** for space $T_{\mathbf{x}}M$. The number of such tangent vectors is the dimensionality of the surface M and I don't think this is the same as k necessarily. So the quote is lost on me.

2.15.16: Everything plays out at some point **x** on manifold M which has dimension k. The tangent space $T_{\mathbf{x}}M$ at point **x** is of dimension k (k = 2 on the torus for example). I think you are supposed to regard the vector arguments $\mathbf{v}_1,\mathbf{v}_2.....\mathbf{v}_k$ as being vectors in $T_{\mathbf{x}}M$ (and not just generic arguments) and I would say that the $\mathbf{e}_i$ are basis vectors for $T_{\mathbf{x}}M$ and are thus x-dependent. So at point **x** when you create a tensor

function by closing a form onto a set of vector arguments $\mathbf{v_1,v_2}.....\mathbf{v_k}$, you are in fact "eating (closing onto) k-tuples of tangent vectors" and you are "spitting out numbers", in that $<\alpha \mid \mathbf{v_1,v_2}.....\mathbf{v_k}>$ = scalar. The k-form α is defined at each point x, hence $\alpha_{\mathbf{x}}$, and so "inhabits M" in that sense. In wedge doc these vector arguments are generic, there is no manifold or tangent space. It helps to read Sja's "forms on manifolds" text on page 88. I confirm all this below in the original notes.

Example 7.17. M = a curve in $R^{\mathbf{N}}$. Dimensionality(M) = 1 so at a point **x** on M there is only one tangent vector. In this example, he assumes **v** lies in $T_{\mathbf{x}}M$ where v is the argument of $\alpha_{\mathbf{x}}(\mathbf{v})$.

Conjecture: Maybe you must have k = m, where m is the dimensionality of M. [correct] Then you could take those arbitrary vector arguments v_i to be the basis vectors of $T_{\mathbf{x}}M$. Then the BT comment makes more sense. Then on a 2D surface, you would always be talking a 2-form and not any other kind of form. This is consistent with he earlier comment that we have

$$(\psi^*\alpha)_{\mathbf{t}}(\mathbf{v_1,v_2}.....\mathbf{v_k}) = \alpha_{\boldsymbol{\psi}(\mathbf{t})}((D\psi)\mathbf{v_1},(D\psi)\mathbf{v_2}.....(D\psi)\mathbf{v_k}) \qquad \mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) \qquad \text{// p 88 F}$$

and α is a multilinear map "on the vector space $T_{\mathbf{x}}M$ ". OK, it is not till the lower half of page 88 that Sja even mentions M in this chapter, and so I guess we have something new here:

(1) For a manifold M of dimension n, you set k = n, and you take those general arguments above $(\mathbf{v_1,v_2}.....\mathbf{v_k})$ to be vectors in the tangent space at x, which is $T_{\mathbf{x}}M$. So you have a little "local $R^{\mathbf{n}}$" at the point **x**. In Chapter 2 we talked about forms on $R^{\mathbf{n}}$ as a full Euclidian space, but now to make forms "work" on a manifold M, we have to talk about the *local* $R^{\mathbf{n}} = R^{\mathbf{k}}$ associated with those tangent vectors. At each different **x**, we have a different local $R^{\mathbf{n}}$. So now $(\mathbf{v_1,v_2}.....\mathbf{v_k})$ are arbitrary vectors, but they are vectors in the local $R^{\mathbf{n}}$ for point **x**. I think that is why he is trying to say.

This does sound like the fibre bundle idea again where at point x we have a vector space $T_{\mathbf{x}}M$.

Now Definition 7.16 makes more sense:

A k-form $\alpha_{\mathbf{x}}$ on M is a chosen element of $AV^{\mathbf{k}}$ (an alternating multilinear map) whose arguments $\mathbf{v_i}$ are vectors in the space $T_{\mathbf{x}}M$ at the point **x** . Thus $\alpha_{\mathbf{x}}$ is defined on $T_{\mathbf{x}}M$ for the point **x**. Now let $\dim(T_{\mathbf{x}}M) = \dim(M) = n$. Maybe you are allowed to talk about any $k \leq n$ just as we did in Chapter 2, so that then is why k and n here are not identified. [I think only k-form with dim M = k makes sense? Well, on a 2D surface you could talk about an area or a curve, so I guess less than a k form also works.]

Example 7.17 Revisited. M = a curve in $R^{\mathbf{N}}$. Dimensionality(M) = 1 so at a point **x** on M there is only one tangent vector **v**. In this example, he assumes **v** lies in $T_{\mathbf{x}}M$ where **v** is the argument of $\alpha_{\mathbf{x}}(\mathbf{v})$. Since n = 1 for M, the only interesting value of k is k = 1, so we would have a 1-form of some sort. There is only one argument of $\alpha_{\mathbf{x}}$, so write $\alpha_{\mathbf{x}}(\mathbf{v})$. Now let us choose to have

$$\alpha_{\mathbf{x}}(\mathbf{v}) = \pm \|\mathbf{v}\| \quad = \text{a real number.} \qquad \text{// this is a tensor function}$$

Notice that this is a form, not a linear functional. It is not linear. This 1-form happens to be the speed of the only tangent vector, and we have the usual direction issue. I don't think this is the ONLY $\alpha_{\mathbf{x}}$ you could invent, but this is one of interest. Somehow this form is going to be related to arc length, but in the Sja notes world it is not yet clear why that would be the case. In Buck p 320 we see $L(\gamma) = \int dt\ \|v\|$, so we are integrating a 1-form $\alpha = \|\mathbf{v}\|\ dt$ where t is the parameter. Although Sja Chap 4 was dedicated to the integration of 1-forms, arc length never made an appearance there. But the pullback integration variable was in fact called t and we had $\int_c \alpha = \int g(t)dt$ as the pullback.

In the final section, Sja now talks about two local regions i and j which overlap. He writes

$W_i = \psi_i(U_i)$
$W_j = \psi_j(U_j)$

where there is some overlap region $W_i \cap W_j$ on the manifold. For a point **x** in the overlap region we would have

$\mathbf{x} = \psi_i(\mathbf{t})$ $\quad$ **t** in U_i
$\mathbf{x} = \psi_j(\mathbf{u})$ $\quad$ **u** in U_i

Then as in Section 7.2 we find results like

$$\psi_i(\mathbf{t}) = \psi_j(\mathbf{u})$$

$$\mathbf{t} = \psi_i^{-1}[\psi_j(\mathbf{u})] = (\psi_i^{-1} \text{ o } \psi_j)(\mathbf{u})$$

and he considers

$\alpha_i = \psi_i^* \alpha$
$\alpha_j = \psi_j^* \alpha$

I guess we can then say

$\alpha = [\psi_i^*]^{-1} \alpha_i$
$\alpha = [\psi_j^*]^{-1} \alpha_j$

so then

$$[\psi_i^*]^{-1} \alpha_i = [\psi_j^*]^{-1} \alpha_j$$

and then

$$\alpha_j = \psi_j^* [\psi_i^*]^{-1} \alpha_i$$

Maybe can write this as

$$\alpha_j = \psi_j^* ([\psi_i^{-1}]^* \alpha_i)$$

Then we look at page 38 regarding the "natural rule" which is

$$\varphi^*(\psi^*\alpha) = (\psi \circ \varphi)^*\alpha$$

Applying this rule, we find that

$$\psi_j^* ([\psi_i^{-1}]^* \alpha_i) = (\psi_i^{-1} \circ \psi_j)^* (\alpha_i)$$

and we then get the result

$$\alpha_j = (\psi_i^{-1} \circ \psi_j)^* (\alpha_i)$$ // page 89 B

In Section 7.1 we just conjectured this to be true. Now that in Section 7.2 we have created a theory of forms on M, we have obtained this result as part of the theory. I admit that my proof here is not very good, but good enough for me now.

I probably should do the exercises, but I will postpone them instead. He shows how you can "antisymmetrize" any k-multilinear function to get one in AV^k. This is what happens in Spivak I recall.

2.15.16 Review. I finally reached this point at 10:30 AM, so this concludes my "review" of the hugely complicated chapter 7. This Chapter has exactly two sections which are the "two definitions":

7.1 forms on manifolds using $\alpha_i = \psi_i^*\alpha$ for each patch on the manifold, α_i = pullback
7.2 forms on manifolds using the monstrous dual space machinery with $\lambda^i = dx^i$ as postulate.