Chapter 1: Introduction **1**
1.1 Manifolds 1
1.2 Equations 2
1.3 Parametrizations // several spellings are allowed 2
1.4 Configuration Spaces 2
Chapter 2: Differential Forms on R^n **2**
2.1 Elementary Properties of differential forms on R^n 2
2.2 The Exterior Derivative [20] 3
2.3 Closed and Exact Forms [22] 4
2.4 The Hodge star operator [23] 4
2.5 div, grad and curl [24] 5
Chapter 3: Pulling Back Forms **6**
3.1 Determinants 6
3.2 Pulling back forms [36] 6
A. The notion of "pulling back" a form from y-space to x-space where $y = \varphi(x)$ [36] 6
B. Theorems involving the pullback operator φ^* [38] 11
Chapter 4: Integration of 1-forms **16**
4.1 Definition of the integral and its elementary properties [47] 16
4.2 Integration of exact 1-forms [49] 17
4.3 The global angle function and the winding number [51] 18
Chapter 5: Integration and Stokes' theorem **19**
5.1 Integration of forms over chains 19
5.2 The boundary of a chain 21
5.3 Cycles and Boundaries 23
5.4 Stokes's Theorem 25
Chapter 6: Manifolds **26**
6.1 The definition of a manifold 26
6.2 The regular value theorem 30

Chapter 1: Introduction

As stated, material here is useful preliminary definitions with examples: manifolds, equations to describe them, parameter equations to describe them, vectors in a configuration space like double pendulum, notion of an affine subspace with $\Phi\mathbf{x} = \mathbf{c}$ and of a tangent space $T_{\mathbf{x}}M$

1.1 Manifolds

For now, **manifold** is a *subset* of R^n. 1D manifold is smooth non-self-intersecting curve. Examples are shown on p2. Remove self intersections by adding another dimension p 3.

1.2 Equations

Manifolds are usually described by an equation, such as that for a circle. Generally $\boldsymbol{\varphi}(\mathbf{x}) = \mathbf{c}$ describes a surface in R^n. For linear case have matrix $\Phi\mathbf{x} = \mathbf{c}$ which describes some "flat thing" in R^n, a plane of some dimension less than n. Such a plane is called an **affine subspace** of R^n.

A sphere in R^n is then called S^{n-1}

A **tangent space** is defined at a point on a sphere. It is an affine subspace, but it is usually then translated to include the origin, then it is a genuine linear subspace of R^n. It is of course normal to the surface at the point in question, so local vectors in the affine space are then "tangent" to the surface. The whole planar surface is tangent to the point before it is translated. More to come later on tan space, and no mention here of cotangent space.

tangent space at point **x** on manifold S^{n-1} = $\{\mathbf{y}$ in $R^n \mid \mathbf{y}\bullet\mathbf{x} = 0\}$ = $T_{\mathbf{x}}M$

$T_{\mathbf{x}}M$ means "the tangent space at point **x** in M" // contains origin.

1.3 Parametrizations // several spellings are allowed

You can parameterize a curve using one parameter, and a surface using two, for example. Buck.

1.4 Configuration Spaces

Double pendulum has $\theta_1,\varphi_1, \theta_2,\varphi_2$ which you could regard as a sort of R^4 subset. Point is that a vector in R^n need not be a simple position vector **x**. Example of pump and wheel.

Chapter 2: Differential Forms on R^n

Now we start into the Main Act.

2.1 Elementary Properties of differential forms on R^n

Differential form: $\alpha = \Sigma_I\, f_I(\mathbf{x})\, dx_I$ on R^n // multi-index, see my Buck notes.

First we have a full multi-index I. The "k" of k-form is number of dx_i in dx_I . k is the form's **degree**. Generally one uses a Greek letter lower case to denote a differential form.

Swap Rule: $dx_i dx_j = - dx_j dx_i$ declared by fiat, no explanation. Implies $dx_i dx_i = 0$.

The **wedge symbol** is so far not used. Notion of an ordered multi-index and counts of how many terms.

2.9.16. Sja makes no notational distinction between his "general multiindex sum" and his "increasing multiindex sum", see center page 18. He does mention wedge notation but does not use it.

Exterior Product of two general forms: [can add forms only of same degree]

α = k-form $= \Sigma_I f_I \, dx_I$

β = m-form $= \Sigma_J g_J \, dx_J$

$\alpha\beta = (\Sigma_I f_I \, dx_I)(\Sigma_J g_J \, dx_J) = \Sigma_{I,J} f_I \, g_J \, dx_I \, dx_J$ // **exterior product** of two diff forms

Degree of product is sum of the degrees of the factors.

2.9.16. My tensor wedge doc talks about a product like $dx_I \, dx_J$ which would be the wedge product of two tensors, my notation would be perhaps $\mathcal{T} \wedge \mathcal{S}$ in the dual space. But I never mention diff forms.

Corollary: α^p is a differential form.

Fact: For α defined on R^n, $\alpha^{n+1} = 0$.

Fact 2.1: $\beta\alpha = (-1)^{ab}\alpha\beta$ "almost commuting" "graded commutativity"

Definition: $\Omega^k(U)$ = **collection of all k-forms on U in R^n.**

Fact 2.2: $\alpha^2 = (-1)^k\alpha^2$ for α = k-form.

Corollary : $\alpha^2 = 0$ for any k-form where k is odd (odd degree)

2.2 The Exterior Derivative [20]

This term refers to $d\alpha$ where α is a form, so really "**differential of a form**". Main idea:

$\alpha = \Sigma_I f_I(\mathbf{x}) \, dx_I$

$d\alpha \equiv \Sigma_I (df_I(\mathbf{x})) \, dx_I = \Sigma_I \; \Sigma_j \, \partial_j(f_I) dx_j dx_I$

Note that

(1) $d\alpha$ problem becomes just df_I problem where f_I is a regular function

(2) if α is of degree k, then $d\alpha$ has degree k+1. So d raises degree by 1.

Some interesting exterior derivatives:

1-form: $\alpha = \Sigma_i f_i \, dx_i \;\Rightarrow\; d\alpha = \Sigma_{i<j} [\; (\partial_i f_j) - (\partial_j f_i)] \; dx_i dx_j$ (2.2)

2-form $\alpha = \Sigma_{i<j} \; f_{ij} dx_i dx_j \;\Rightarrow\; d\alpha = \Sigma_{i<j<k} [\; (\partial_k f_{ij}) + (\partial_i f_{jk}) - (\partial_j f_{ik}) \;] \, dx_i dx_j dx_k$ (2.4)

Rule Set 2.5: (i) $d(a\alpha+b\beta) = ad\alpha + bd\beta$ α,β = k-forms p 21

(ii) $d(\alpha\beta) = (d\alpha)\beta + (-1)^k\alpha(d\beta)$ α = k-form β = m-form

The first is a combination of the "sum rule" and "constant multiple rule" for a normal derivative.
The second is a modified Leibniz's Rule (aka "the product rule")

Fact 2.6: $d^2 = 0$ which says $d(d\alpha) = 0$ for any form α.

2.3 Closed and Exact Forms [22]

If $d\alpha = 0$, the form α is **closed**. It is like something being "a constant" in regular calculus.
If $d\beta = \alpha$, the form α is **exact**. It is like an "exact differential" in regular calculus.

Fact 2.7: α exact $\Rightarrow$ α closed

Proof: α exact $\Rightarrow \alpha = d\beta \Rightarrow d\alpha = d^2\beta = 0$ using d^2 rule below, thus $d\alpha = 0$ so is closed.

When is a 1-form α exact? Need to find β such that $d\beta = \alpha$, and we know β must be a 0-form so just call that g. So problem is $d\beta = \alpha$ or $dg = \alpha$ or $\partial_i g dx_i = \alpha = \Sigma_i f_i dx_i$ so need $\partial_i g = f_i$ which is a system of simple PDE's. Sja gives some examples of how you might solve this system.

Fact 2.10. The **angle form** is closed but is not exact. (p 23) Note that $d\alpha = 0$ though $\alpha \neq$ constant.

When is a 2-form α exact? Result is given p 23. You must integrate this PDE system:

$(\partial_i g_j) - (\partial_j g_i) = f_{ij}$ // agrees with S (2.6).

2.4 The Hodge star operator [23]

First, here is what this * operator does to a product of differentials

$(*dx_I) = (\text{sign})\, dx_{Ic}$

Here dx_I is an ordered product, and dx_{Ic} is the ordered product of all the dx_i *not* in dx_I in R^n.
You then get

$dx_I\,(*dx_I) = (\text{sign})\, dx_I\, dx_{Ic} = +\, dV$

The (sign) appearing in $(*dx_I)$ is chosen to get the result +dV on the second line! Sja gives no expression for this sign, but in a separate doc I show that

$(\text{sign}) = S = (-1)^{a+b+..+q}\ (-1)^{k(k+1)/2}$

for

$dx_I = dx_a\, dx_b \,....\, dx_q$ $\quad a < b < ... < q$ $\quad$ a k-form $\quad$ *(2)*

Exercise 2.12: $*(*\alpha) = (-1)^{kn+k}\alpha$ for a k-form on R^n (shown in same separate doc)

Exercise 2.18. Concerns writing **Maxwell's Equations** in terms of differential forms.

2.5 div, grad and curl [24]

Fact: Can write any 0-form as $\alpha = f$, so can associate ∇f with $d\alpha$ since $d\alpha = \nabla f \bullet dx$

Fact: Can write any 1-form as $\alpha = \mathbf{F} \bullet d\mathbf{x}$ so have 1-to-1 association $\alpha \leftrightarrow \mathbf{F}$ (vector)

Fact: Can write any 2-form as $\alpha = \Sigma_{i<j} F_{ij}\, dx_i dx_j$ so associate $\alpha \leftrightarrow F_{ij}$ (rank 2 tensor)

Fact: Can associate with a k-form α a rank-k tensor $F_{ij\ldots k}$ (k indices)

In this section Sja shows all these facts:

$\alpha = f$	0-form in R^n	$\alpha \leftrightarrow f$	PL1
$d\alpha = df$	1-form in R^n	$d\alpha \leftrightarrow \nabla f$	// namely, $d\alpha = \nabla f \bullet d\mathbf{x}$
$\alpha = f$	0-form in R^n	$\alpha \leftrightarrow f$	
$*d[*(d\alpha)] = \partial_i^2 f = \nabla^2 f$	0-form on R^n	$*d[*(d\alpha)] = \nabla^2 f$	PL2
$\alpha = \mathbf{F} \bullet d\mathbf{x}$	1-form on R^n	$\alpha \leftrightarrow \mathbf{F}$	
$*[d(*\alpha)] = (\text{div } \mathbf{F})$	0-form on R^n	$*[d(*\alpha)] = \text{div } \mathbf{F}$	PL3
$\alpha = \mathbf{F} \bullet d\mathbf{x}$	1-form in R^3	$\alpha \leftrightarrow \mathbf{F}$	one-to-one
$*(d\alpha) = [\text{curl } \mathbf{F}] \bullet d\mathbf{x}$	1-form in R^3	$*(d\alpha) \leftrightarrow \text{curl } \mathbf{F}$	PL4

Thus, starting with a differential form α of appropriate degree, and using just the d and * operators, one can construct all the basic differential operators of tensor analysis! Here is a second summary

$\alpha = f$	$\leftrightarrow f$	α = 0-form	R^n
$d\alpha$	$\leftrightarrow \nabla f$		
$*d[*(d\alpha)]$	$\leftrightarrow \nabla^2 f$		
$\alpha = \mathbf{F} \bullet d\mathbf{x}$	$\leftrightarrow \mathbf{F}$	α = 1-form	R^3 only
$*[d(*\alpha)]$	$\leftrightarrow \text{div } \mathbf{F}$		
$*(d\alpha)$	$\leftrightarrow \text{curl } \mathbf{F}$		

The curl **F** one only exists for $R^n = R^3$, however.

Hint of things to come: We know that curl F and div F and ∇f occur in various integral theorems like the divergence theorem and Stokes's theorem and various others. One wonders what these theorems look like in terms of α and * and d?

Comments: At this point the reader is familiar with much of the differential form machinery. But no contact has been made yet between Chapter 2 on forms and Chapter 1 on manifolds. The reader is thinking: "OK, I know how to manipulate "forms", but what good are they? How do they help me do anything or understand anything?" At this point, that question is unanswered. But the reader does see some dim connection between forms and differential operators. I think we will have to wait until Chapter 4 before any usefulness appears. So far forms are a quaint curiosity.

Chapter 3: Pulling Back Forms

3.1 Determinants

Sjamaar does his little piece on determinants, just as I did my piece in Lagrange Appendix B which I just enhanced in an extra doc on "ways to write dets". His gimmick is to define the determinant in terms of three axioms. He then talks about "elementary column operations" and shows how det(A) behaves under each of them: swap columns get -1, mult col by c get c, add col to another get 1. This is a nice approach, but I like my Appendix B better and my notation there as well. He ends up with page 33 A as a way to write out det(A) and this is just one of my many ways. He does not use the word "group".

He uses a few strange terms:

inversions = swaps
length = the number of swaps of a permutation relative to z_0

His inversion is really the number of pairs of integers in a permutation which are out of standard order. So that is the number of swaps you need to do to get them back in. He does not dwell on this.

At the end on page 35A he gives some standard facts about determinants, all familiar to me. He then pulls the volume change det(S) = vol ratio out of his hat without any proof.

So I am happy with this section, there is nothing here I did not already know.

3.2 Pulling back forms [36]

A. The notion of "pulling back" a form from y-space to x-space where y = φ(x) [36]

I was unclear about some things in the raw notes, here I think I have straightened things out a bit.

Preliminary Questions:

1. Consider in 1D the statement dy = g(x) dx. Is this a "mapping"? You would have to write this somehow as map: {f(x)} x {ds} → {ds}. For each pair {g,dx} you get a unique dy. So you could maybe write

dy = dy(x,dx) = g(x) dx

where you imagine dy is a function then of two variables x and dx. But dx is not a variable, it is the differential of a variable x.

Conclusion: If dy = g(x)dx, you cannot claim that dy is a "function" in either of these senses

dy = dy(x) dy = dy(x,dx)

That is why "forms" are different from "functions".

2. Implication: if we write a differential form $\alpha = g(x)dx$, we cannot treat α as a function α(x). You will note that Sjamaar never writes α(**x**) for such a thing. This notation would imply that for each value of **x**, there is some unique value of α(**x**). It is true that $\alpha \equiv g(x)dx$ is *associated with* x-space because the differential is in x-space and the function g(x) is defined on x-space, but α is not a function of **x**.

3. Implication: It is tempting to make this statement based on $\mathbf{y} = \varphi(\mathbf{x})$

$\alpha(\mathbf{y}) = \Sigma_I \, f_I(\mathbf{y}) \, dy_I(\mathbf{y})$
$\alpha(\varphi(\mathbf{x})) = \Sigma_I \, f_I(\varphi(\mathbf{x})) \, dy_I(\varphi(\mathbf{x}))$.

Then the whole subject of changing variables is done as if α were a *function*. This is just plain wrong, and I have allowed myself to be deceived. As shown below, you *can* write the second line this way

$\varphi^*\alpha = \Sigma_I \, f_I(\varphi(\mathbf{x})) \, d\varphi_I(\mathbf{x})$.

Definition of the Pullback of a differential form:

Let $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ be some vector-valued function of a vector **x**, and $\varphi: R^n \rightarrow R^m$. **φ** is a function.

D1. Define the operator φ* acting on a *function* in this way

$\varphi^* f(\mathbf{y}) \equiv f(\varphi(\mathbf{x})) = g(\mathbf{x})$. // = (f o φ)(**x**)

Notice that $\varphi^*[f_1 + f_2] = \varphi^* f_1 + \varphi^* f_2$ and $\varphi^*[cf_1] = c\varphi^*(f)$, where f, f_1 f_2 are functions!

The operator φ* maps a y-space function into an x-space function.

D2. Define the operator φ* acting on a *differential* in this way (an example)

$\varphi^* (dy_1 dy_2) = d\varphi_1(\mathbf{x}) \, d\varphi_2(\mathbf{x}) = [\partial_i \varphi_1(\mathbf{x}) \, dx_i][\, [\partial_j \varphi_2(\mathbf{x}) \, dx_j] = [(\partial_i \varphi_1(x))(\partial_j \varphi_2(x))] \, dx_i dx_j$.

We *cannot* write this as

$\varphi^* (dy_1 dy_2)(\mathbf{y}) = (dy_1 dy_2)(\boldsymbol{\varphi}(\mathbf{x}))$

but we *can* write this as

$\varphi^* (dy_1 dy_2) = d\varphi_1(\mathbf{x})\, d\varphi_2(\mathbf{x})$. // perhaps write schematically as $d\varphi_I(\mathbf{x})$:

$\varphi^* (dy_I) = d\varphi_I(\mathbf{x})$

The operator φ^* maps a differential in y-space into a differential form in x-space, see fig page 37 and see item D2 above as an example.

D3. Define the operator φ^* acting on the *product of a function and a differential* in this way

$\varphi^*[f(\mathbf{y})\, dy_I] \equiv \varphi^*(f(y))\, \varphi^*(dy_I)$.

D4. What could you say then about this:

$\varphi^*[f_1(\mathbf{y})\, dy_1 dy_2 + f_2(\mathbf{y})\, dy_2 dy_3] = ?$

From what is stated so far, we don't have an answer. We have to add a new part of our φ^* definition:

$\varphi^*[f_1(\mathbf{y})\, dy_1 dy_2 + f_2(\mathbf{y})\, dy_2 dy_3] = \varphi^*[f_1(\mathbf{y})\, dy_1 dy_2] + \varphi^*[\, f_2(\mathbf{y})\, dy_2 dy_3]$.

Thus, φ^* is an operator which acts on differential forms. Based in this above item, we conclude that

D5. $\varphi^*(\alpha+\beta) = \varphi^*(\alpha) + \varphi^*(\beta)$.

D6. Using the above steps, we conclude that

$\varphi^*\alpha = \varphi^* [\Sigma_I\, f_I(\mathbf{y})\, dy_I\,] = \Sigma_I\, (\varphi^* f_I(\mathbf{y}))\, (\varphi^*[dy_I])$

$= \Sigma_I\, f_I(\varphi(\mathbf{x}))\, d\varphi_I(\mathbf{x})$.

The operator φ^* maps a differential form in y-space into a differential form in x-space. Both differential forms have the same degree.

D7. $\varphi^*(c\alpha) = \varphi^* [c(\, [\Sigma_I\, f_I(\mathbf{y})\, dy_I\,]\,)] = \varphi^*[\Sigma_I\, (c f_I(\mathbf{y}))\, dy_I] = [\Sigma_I\, (c f_I(\mathbf{y}(\mathbf{x})))\, d\varphi_I]$

$= c[\Sigma_I\, (f_I(\mathbf{y}(\mathbf{x})))\, d\varphi_I] = c[\varphi^*(\alpha)] = c\, (\varphi^*\alpha)$

Here are some examples:

(a) Consider for a 1-form $\alpha = f(y)\, dy$ with y in R^1 :

$y = \varphi(x) = x^2$ $\qquad dy = d\varphi = \partial_{\mathbf{x}}\varphi\, dx = 2x\, dx$

$\varphi^*\alpha = f(x^2)\, 2x\, dx \equiv \beta$.

Whereas $\alpha = f(y)dy$ is "associated with y-space", the pullback form $\beta \equiv \varphi^*\alpha$ is associated with x-space. But we do not write $\alpha(\mathbf{y})$ or $\beta(\mathbf{x})$!

(b) Consider for a 2-form with $x = x_1, x_2$

$$y_1 = \varphi_1(\mathbf{x}) = x_1^2 x_2 \qquad dy_1 = d\varphi_1 = \partial_j\varphi_1\, dx_i = (\partial_1\varphi_1)dx_1 + (\partial_2\varphi_1)dx_2 = 2x_1x_2dx_1 + x_1^2dx_2$$
$$y_2 = \varphi_2(\mathbf{x}) = x_1x_2^3 \qquad dy_2 = d\varphi_2 = \partial_j\varphi_2\, dx_i = (\partial_1\varphi_2)dx_1 + (\partial_2\varphi_2)dx_2 = x_2^3dx_1 + 2x_1x_2^2dx_2$$

$$dy_I = dy_1dy_2 = d\varphi_1d\varphi_2 = d\varphi_I = (\, 2x_1x_2dx_1 + x_1^2dx_2)(\, x_2^3dx_1 + 2x_1x_2^2dx_2)$$
$$= (4x_1^2x_2^3 - x_1^2x_2^3)\, dx_1dx_2 = 3x_1^2x_2^3dx_1dx_2$$

So in this last example, $dy_I = d\varphi_I(\mathbf{x}) = 3x_1^2x_2^3\, dx_I$.
In this case where number of variables stays the same, a 2-form gives a 2-form. It happens in this example that the starting 2-form is just $\alpha = dy_I = dy_1dy_2$ with no function part. If we now add a function part, we would have

$$\alpha = f(\mathbf{y})dy_I \qquad \varphi^*\alpha = f(\boldsymbol{\varphi}(\mathbf{x}))\, d\varphi_I(\mathbf{x}) = [f(x_1^2x_2,\, x_1x_2^3)]\, [3x_1^2x_2^3\, dx_I] = g_I(\mathbf{x})\, dx_I$$

The new function g_I has two factors as shown.

(c) Consider for a 2-form with $x = x_1$

$$y_1 = \varphi_1(\mathbf{x}) = x_1^2 \qquad dy_1 = d\varphi_1 = \partial_j\varphi_1\, dx_i = (\partial_1\varphi_1)dx_1 + (\partial_2\varphi_1)dx_2 = 2x_1dx_1$$
$$y_2 = \varphi_2(\mathbf{x}) = x_1 \qquad dy_2 = d\varphi_2 = \partial_j\varphi_2\, dx_i = (\partial_1\varphi_2)dx_1 + (\partial_2\varphi_2)dx_2 = 1\, dx_1$$

$$dy_I = dy_1dy_2 = d\varphi_1d\varphi_2 = d\varphi_I = (\, 2x_1dx_1)(\, 1\, dx_1) = 0$$

Here since only one variable, the resulting 2-form vanishes since only one dx_1 to deal with

(d) Consider a 2-form with $x = x_1, x_2, x_3$

$$y_1 = \varphi_1(\mathbf{x}) = x_1^2x_2x_3 \qquad dy_1 = d\varphi_1 = \partial_j\varphi_1\, dx_i = (\partial_1\varphi_1)dx_1 + (\partial_2\varphi_1)dx_2 + (\partial_3\varphi_1)dx_3$$
$$y_2 = \varphi_2(\mathbf{x}) = x_1x_2^3x_3 \qquad dy_2 = d\varphi_2 = \partial_j\varphi_2\, dx_i = (\partial_1\varphi_2)dx_1 + (\partial_2\varphi_2)dx_2 + (\partial_3\varphi_2)dx_3$$

$$dy_1 = 2x_1x_2x_3\, dx_1 + x_1^2x_3\, dx_2 + x_1^2x_2\, dx_3$$
$$dy_2 = x_2^3x_3dx_1 + 2x_1x_2^2x_3dx_2 + x_1x_2^3dx_3$$

$$dy_I = dy_1dy_2 = d\varphi_1d\varphi_2$$
$$= (2x_1x_2x_3\, dx_1 + x_1^2x_3\, dx_2 + x_1^2x_2\, dx_3)(x_2^3x_3dx_1 + 2x_1x_2^2x_3dx_2 + x_1x_2^3dx_3)$$

$$= (\text{stuff})\, dx_1dx_2 + (\text{stuff})\, dx_1dx_3 + (\text{stuff})\, dx_2dx_3 \ .$$

Here we again get a 2-form in x-space, but it has several terms, not just one term. Here we have more variables in the x-world than in the y-world.

(e) Consider this specific 1-form with a specific $\mathbf{y} = \varphi(\mathbf{x})$:

$$\alpha = f_1(y_1)dy_1 + f_2(y_2)dy_2 = [-y_2/(y_1^2+y_2^2)]dy_1 + [+y_1/(y_1^2+y_2^2)]dy_2$$

$y_1 = \varphi_1(\mathbf{x}) = \cos(x_1)$ // x-space has only 1 variable so $\varphi: R^2 \rightarrow R^1$
$y_2 = \varphi_2(\mathbf{x}) = \sin(x_1)$

$dy_1 = d\varphi_1(x) = -\sin(x_1)dx_1$ $\quad$ $f_1(y_1) = [-y_2/(y_1^2+y_2^2)] = -\sin(x_1)/1$
$dy_2 = d\varphi_2(x) = \cox(x_1)dx_1$ $\quad$ $f_2(y_2) = [+y_1/(y_1^2+y_2^2)] = \cos(x_1)/1$

$$\varphi^*\alpha = \mathbf{f}(\varphi(\mathbf{x}))\bullet d\boldsymbol{\varphi}(\mathbf{x}) = f_1(\varphi(x))d\varphi_1(x) + f_2(\varphi(x))d\varphi_2(x)$$

$$= [-\sin(x_1)][-\sin(x_1)]\,dx_1 + [\cos(x_1)][\cox(x_1)dx_1]\,dx_1 = dx_1$$

So for this special form and function $\varphi(x)$ where y-space has 2 variables and x-space only 1, we find that the complicated form α shown above is "pulled back" to the very simple form dx_1 in x-space.

So go back to the general case:

$$\alpha = \Sigma_I f_I(y)dy_I$$

$$\varphi^*\alpha = \Sigma_I\, f_I(\varphi(\mathbf{x}))\, d\varphi_I(\mathbf{x}) .$$

The last factor is

$$dy_I = dy_a dy_b ... dy_q = d\varphi_a d\varphi_b ... d\varphi_q = d\varphi_I$$

$$= \Sigma_{AB..Q}\, (\partial_A\varphi_a)dx_A\, (\partial_B\varphi_b)dx_B \,.... (\partial_Q\varphi_q)dx_Q$$

$$= \Sigma_{AB..Q}\, (\partial_A\varphi_a)(\partial_B\varphi_b)... (\partial_Q\varphi_q)dx_A dx_B ....dx_Q$$

$$= \Sigma_{AB..Q}\, [(\partial_A\varphi_a)(\partial_B\varphi_b)... (\partial_Q\varphi_q)]\, dx_I$$

If there are k-factors of dy_i (α was a k-form), then there are k-factors of dx_i (still a k-form).

So we then can say

$$\varphi^*\alpha = \Sigma_I\, f_I(\varphi(\mathbf{x}))\, \{\Sigma_{AB..Q}\, [(\partial_A\varphi_a)(\partial_B\varphi_b)... (\partial_Q\varphi_q)]\, dx_I \}$$

$$= \Sigma_I\, \{\ f_I(\varphi(\mathbf{x}))\Sigma_{AB..Q}\, [(\partial_A\varphi_a)(\partial_B\varphi_b)... (\partial_Q\varphi_q)]\ \}\, dx_I$$

$= \Sigma_I \; f_I(\varphi(\mathbf{x}))\, s_I(\mathbf{x})\, dx_I \qquad s_I(\mathbf{x}) \equiv \Sigma_{AB..Q} [(\partial_A \varphi_a)(\partial_B \varphi_b)... (\partial_Q \varphi_q)]$

Fact: No matter what $\mathbf{y} = \varphi(\mathbf{x})$ looks like (same or different # variables, non-linear, not 1-1) the above mechanical process takes the form dy_I into a form [stuff]dx_I in x-space. So that last factor is then not so mysterious after all.

Question: Can you write

$$\alpha \circ \varphi = \Sigma_I \, (f_I \circ \varphi) \, (dy_I \circ \varphi) \; ?$$

I think not! The object $dy_I \circ \varphi$ implies that dy_I is a function, and we have just said that it is not a function! Also $\alpha \circ \varphi$ implies α is a function, but it too is not a function. So at best you can say

$$\varphi^*\alpha = \Sigma_I \, (f_I \circ \varphi)(\mathbf{y}) \, \varphi^*(dy_I) = \Sigma_I \, (f_I \circ \varphi)(\mathbf{x}) \, d\varphi_I(\mathbf{x}) = \Sigma_I \, f_I(\varphi(\mathbf{x}))d\varphi_I(\mathbf{x}) \; .$$

Here then is the big idea in terms of mappings:

$\varphi: X \to Y$	$\varphi(\mathbf{x}) = \mathbf{y}$	$\mathbf{x} \to \mathbf{y}$
$\varphi^*: \Omega^k(Y) \to \Omega^k(X)$	$\varphi^*\alpha = \Sigma_I \, (\varphi^* f_I)(\mathbf{x}) \, d\varphi_I(\mathbf{x}) \equiv \beta$	$\alpha \to \beta$

The mapping φ takes a point $\mathbf{x}$ into $\mathbf{y}$ and think of this as the **forward** mapping direction X to Y space (left to right). The mapping φ^* "**pulls back** the form α from Y-space into X-space" (right to left) . Whereas α is associated with y-space, $\beta = \varphi^*\alpha$ is associated with x-space. Again we do not write $\alpha(\mathbf{y})$ or $\beta(\mathbf{x})$ because these are not functions. We are pulling α back "along φ". See drawing page 37.

In regular calculus you could imagine this

$\varphi: X \to Y$	$\varphi(\mathbf{x}) = \mathbf{y}$	$\mathbf{x} \to \mathbf{y}$
$\varphi^*: F(Y) \to F(X)$	$\varphi^* f = (\varphi^* f)(\mathbf{x}) = f(\varphi(\mathbf{x})) = F(\mathbf{x})$	$f(\mathbf{y}) \to F(\mathbf{x})$

Here our operator φ^* moves a function defined on y-space to one defined on x-space. So in this case you could say the function $f(\mathbf{y})$ is pulled back along φ to become the function $F(\mathbf{x})$.

B. Theorems involving the pullback operator φ^* [38]

I will restate these a bit, based on 3.10 on page 38. But first,

PL1: Adding two forms of the same degree is well-defined:

$$\alpha + \beta = \Sigma_I f_I dx_I + \Sigma_J g_J dx_J$$

$$= \Sigma_I f_I dx_I + \Sigma_I g_I dx_I \qquad \text{// where both } f_I \text{ and } g_I \text{ are "extended" over all } dx_I$$

$= \Sigma_I(f_I+g_I)dx_I = \Sigma_I h_I dx_I = \gamma$ (some new form of degree k)

Extended means lots of the f_I and g_I functions will be 0 for terms that don't exist in their respective sums. Notice that you cannot add forms of different degree such as dx + dydz because the result would not be a form of a well-defined degree and an official form must have a degree. Note that the external product of two forms αβ exists where α and β can have the same or different degree.

Example: $\alpha = dx \quad \beta = dy \qquad f_1 = 1 \quad f_2 = 0 \qquad g_1 = 0 \quad g_2 = 1$

$f_1 + g_1 = 1 \qquad f_2 + g_2 = 1 \qquad \alpha+\beta = dx + dy$

PL2. Multiplying a form by a constant is well-defined

$c(\alpha) = c(\Sigma_I f_I dx_I) = \Sigma_I (cf_I)dx_I = \Sigma_I F_I dx_I$ = a new form , call it $\beta = (c\alpha)$

PL3. Multiplying two forms is well-defined

$(\alpha\beta) = (\alpha)(\beta) = (\Sigma_I f_I dx_I)(\Sigma_J g_J dx_J)$ = exterior product discussed on page 19 B

Now come proofs of the three page 38 theorems:

Here then is a proof of **page 38 (i): linear rule** $\varphi^*(a\alpha+b\beta) = a(\varphi^*\alpha) + b(\varphi^*\beta)$

Proof: $\varphi^*(a\alpha+b\beta) = \varphi^*(a\alpha) + \varphi^*(b\beta)$ // using D5 above

$= a(\varphi^*\alpha) + b(\varphi^*\beta)$ // using D7 above (twice) **QED**

Here then is a proof of **page 38 (ii): product rule** $\varphi^*(\alpha\beta) = \varphi^*(\alpha)\ \varphi^*(\beta)$

Proof: $\varphi^*(\alpha\beta) = \varphi^*[(\Sigma_I f_I dx_I)(\Sigma_J g_J dx_J)] = \varphi^*[\Sigma_{I,J} f_I g_J dx_I dx_J] = \Sigma_{I,J}\varphi^*(f_I g_J)\varphi^*(dx_I dx_J)$

$\Sigma_{I,J}(\varphi^*f_I)(\varphi^*g_J)\, d\varphi_I d\varphi_J = (\Sigma_I(\varphi^*f_I)d\varphi_I)(\Sigma_J(\varphi^*g_J)d\varphi_J) = \varphi^*(\alpha)\ \varphi^*(\beta)$ **QED**.

Here then is a proof of **page 38 (iii): natural rule** $\varphi^*(\psi^*\alpha) = (\psi \circ \varphi)^*\alpha$

This proof takes some *work* which I do again here. It all boils down to a vector field chain rule in the end.

Here we have two mappings or variable changes,

$\mathbf{y} = \varphi(\mathbf{x})$ as earlier
$\mathbf{z} = \psi(\mathbf{y})$ new

The new feature here is a double variable change involving ψ and φ (functions, not forms)

Let: $\mathbf{z} = \psi(\mathbf{y}) = \psi(\varphi(\mathbf{x})) \equiv \theta(\mathbf{x}) = (\psi \circ \varphi)(\mathbf{x}) \qquad \theta = \psi \circ \varphi$

If α were just a function, this item would be trivial to prove:

$\varphi^*(\psi^*\alpha(\mathbf{z})) = \varphi^*(\alpha(\psi(\mathbf{y})) = \alpha(\psi(\boldsymbol{\varphi}(\mathbf{x}))$

$(\psi \circ \varphi)^*\alpha(\mathbf{z}) = \theta^*\alpha(\mathbf{z}) = \alpha(\theta(\mathbf{x})) = \alpha(\psi(\boldsymbol{\varphi}(\mathbf{x})))$

But the whole point is that α is NOT a function, it is a differential form! As a result, the proof is much harder, and the theorem correspondingly has more power.

Process each side of the claimed theorem $\varphi^*(\psi^*\alpha) = (\psi \circ \varphi)^*\alpha$:

LHS = $\varphi^*(\psi^*\alpha) = \varphi^*[\psi^*(\Sigma_I f_I(\mathbf{z})dz_I)] = \varphi^*[(\Sigma_I f_I(\psi(\mathbf{y}))d\psi_I)] = \varphi^*[\Sigma_I f_I(\psi(\mathbf{y}))d\psi_a d\psi_b \ldots d\psi_q]$

$= \varphi^*[\Sigma_I f_I(\psi(\mathbf{y}))[\Sigma_A(\partial^{(y)}{}_A\psi_a)dy_A][\Sigma_B(\partial^{(y)}{}_B\psi_b)dy_B]....[\Sigma_Q(\partial^{(y)}{}_Q\psi_q)dy_Q]$

$= \varphi^*[\Sigma_I \{ f_I(\psi(\mathbf{y})) \Sigma_{A,B..Q} (\partial^{(y)}{}_A\psi_a)(\partial^{(y)}{}_B\psi_b)..(\partial^{(y)}{}_Q\psi_q) \} dy_A dy_B....dy_Q]$

$= \varphi^*[\Sigma_I \{ \Sigma_{A,B..Q} f_I(\psi(\mathbf{y})) J_{ab..q,AB..Q}(\mathbf{y}) \} dy_A dy_B....dy_Q]$

$= [\Sigma_I \{ \Sigma_{A,B..Q} f_I(\psi(\boldsymbol{\varphi}(\mathbf{x}))) J_{ab..q,AB..Q}(\boldsymbol{\varphi}(\mathbf{x})) \} d\varphi_A d\varphi_B....d\varphi_Q]$

$= [\Sigma_I \{ \Sigma_{A,B..Q} f_I(\psi(\boldsymbol{\varphi}(\mathbf{x}))) J_{ab..q,AB..Q}(\boldsymbol{\varphi}(\mathbf{x})) \} (\partial_{a'}\varphi_A)dx_{a'} (\partial_{b'}\varphi_B)dx_{b'}...(\partial_{q'}\varphi_Q)dx_{q'}]$

$= \Sigma_I \{ \Sigma_{A,B..Q} \Sigma_{a'b'..q'} f_I(\psi(\boldsymbol{\varphi}(\mathbf{x}))) * (\partial^{(y)}{}_A\psi_a)(\varphi(\mathbf{x}))(\partial^{(y)}{}_B\psi_b)(\varphi(\mathbf{x}))..(\partial^{(y)}{}_Q\psi_q)(\varphi(\mathbf{x}))$
$* (\partial_{a'}\varphi_A)(\mathbf{x})(\partial_{b'}\varphi_B)(\mathbf{x}) .....(\partial_{q'}\varphi_Q)$
$* dx_{a'}dx_{b'}.....dx_{q'}$

But now regroup factors pairwise and hold: Note that $(\partial^{(y)}f)(\varphi(\mathbf{x}))$ means $(\partial^{(y)}f(y))|^{y=\varphi(x)}$.

$= \Sigma_I \{ f_I(\psi(\boldsymbol{\varphi}(\mathbf{x}))) * \Sigma_{A,B..Q} \Sigma_{a'b'..q'}$
$[(\partial^{(y)}{}_A\psi_a)(\varphi(\mathbf{x})) (\partial_{a'}\varphi_A)(\mathbf{x})] [(\partial^{(y)}{}_B\psi_b)(\varphi(\mathbf{x})) (\partial_{b'}\varphi_B)(\mathbf{x})] .....[..] dx_{a'}dx_{b'}.....dx_{q'}$

Meanwhile

RHS = $(\psi \circ \varphi)^*\alpha = (\psi \circ \varphi)^*(\Sigma_I f_I(\mathbf{z})dz_I) = \theta^*(\Sigma_I f_I(\mathbf{z})dz_I) = \Sigma_I f_I(\theta(\mathbf{x})) d\theta_I$

$= \Sigma_I f_I(\theta(\mathbf{x})) d\theta_a d\theta_b....d\theta_q$

$= \Sigma_I f_I(\theta(\mathbf{x})) \Sigma_{a'b'..q'}(\partial_{a'}\theta_a)(\partial_{b'}\theta_b) ....(\partial_{q'}\theta_q).dx_{a'}dx_{b'}.....dx_{q'}$

So we have a proof if we can show that

$\Sigma_{A,B..Q} [(\partial^{(y)}{}_A\psi_a)(\varphi(\mathbf{x})) (\partial_{a'}\varphi_A)(\mathbf{x})] [(\partial^{(y)}{}_B\psi_b)(\varphi(\mathbf{x})) (\partial_{b'}\varphi_B)(\mathbf{x})] .....[..]$

$= (\partial_a, \theta_a)(\partial_b, \theta_b) \ldots (\partial_q, \theta_q)$.

And this would be in turn be proved if we could just show

$$\Sigma_A [(\partial^{(y)}{}_A \psi_a)(\varphi(\mathbf{x})) (\partial_a, \varphi_A)(\mathbf{x})] = (\partial_a, \theta_a)(\mathbf{x}) .$$

Write this out in clearer notation

$$\Sigma_A \frac{\partial \psi_a(\mathbf{y})}{\partial y_A} |^{\varphi(x)} \frac{\varphi_A(\mathbf{x})}{\partial x_a,} = \frac{\partial \theta_a(\mathbf{x})}{\partial x'_a} .$$

Well we know that

$$\theta(\mathbf{x}) = \psi(\varphi(\mathbf{x}))$$

$$\theta_a(\mathbf{x}) = \psi_a(\boldsymbol{\varphi}(\mathbf{x}))$$

$$\frac{\partial \theta_a(\mathbf{x})}{\partial x'_a} = \Sigma_A \left[\frac{\partial \psi_a(\mathbf{y})}{\partial y_A} |^{y=\varphi(x)} \right] \frac{\varphi_A(\mathbf{x})}{\partial x_a,}$$ "chain rule" **QED**

So yes, this matches the above and our proof is complete. I think my proof is perhaps better than his in that it does a full attack all at once. This kind of fancy chain rule appears as (B.6) but only for a special case.

Summary of the last three theorems:

$\varphi^*(a\alpha+b\beta) = a(\varphi^*\alpha) + b(\varphi^*\beta)$	linear rule
$\varphi^*(\alpha\beta) = \varphi^*(\alpha)\, \varphi^*(\beta)$	product rule
$\varphi^*(\psi^*\alpha) = (\psi \circ \varphi)^*\alpha$	natural rule

Chain Rule for Vector Fields // probably his (B.6)

Suppose $\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{h}(\mathbf{x}))$ where all bolded things are vectors. Then a component is

$$f_i(\mathbf{x}) = g_i(\mathbf{h}(\mathbf{x}))$$

Then

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j} = \frac{\partial}{\partial x_j} [g_i(\mathbf{h}(\mathbf{x}))] = \Sigma_k \frac{\partial g_i(\mathbf{h})}{\partial h_k} \frac{\partial h_k(\mathbf{x})}{\partial x_j}$$ // vector chain rule

Theorem 3.11. $\varphi^*(d\alpha) = d(\varphi^*\alpha)$ [39]

This theorem is rather astounding, saying that the φ^* and d operators exactly commute!

Proof:

$$\text{LHS} = \varphi^*(d\alpha) = \varphi^*(d[\Sigma_I\, f_I(\mathbf{y})\, dy_I)] = \varphi^*(\Sigma_I\, [df_I(\mathbf{y})]\, dy_I) = \varphi^*(\Sigma_I\, [(\partial^{(y)}{}_k f_I) dy_k dy_I]$$

$$= \Sigma_I\, [(\partial^{(y)}{}_k f_I)(\varphi(\mathbf{x}))\;\; d\varphi_k d\varphi_I \quad]$$

$$\text{RHS} = \; d(\varphi^*\alpha) = \; d(\varphi^*[\Sigma_I\, f_I(\mathbf{y})\, dy_I]) \; = d(\; \Sigma_I\, f_I(\boldsymbol{\varphi}(\mathbf{x}))\, d\varphi_I\;) = \; \Sigma_I\, d[\; f_I(\boldsymbol{\varphi}(\mathbf{x}))\, d\varphi_I(\mathbf{x})]$$

But we know that $d(d\varphi_i(x)) = d[\; (\partial_k \varphi_i) dx_k \;] = d[\; (\partial_k \varphi_i)\;] dx_k = \partial_{jk}\varphi_i dx_j dx_k = 0$ by symmetry!
We also know this since $d(d\varphi_i(x)) = d^2\varphi_i = 0$. Thus we have

$$\text{RHS} = \; d(\varphi^*\alpha) = \Sigma_I\, d[\; f_I(\boldsymbol{\varphi}(\mathbf{x}))]\, d\varphi_I \; = \; \Sigma_I\, [(\partial^{(y)}{}_k f_I)(\varphi(\mathbf{x}))\, d\varphi_k\, d\varphi_I\;] \; .$$

But this is the same as LHS above, QED. Here I avoid grinding both sides down to dx_I. I think my proof is simpler, but I suppose I could have missed some major thing.

Show that $d\varphi_I = J\, dx_I$ where J is a certain determinant. [40-41]

This is quite interesting to me. This is not the usual Jacobian rule but one special case of it does give that Jacobian rule (p 42C) (apart from wedges!) . But in general, we get a certain determinant as follows:

$$d\varphi_a d\varphi_b d\varphi_c \ldots d\varphi_q \;=\; \Sigma_{A<B<C\ldots<Q}\; \det(R_{abc\ldots q,ABC\ldots Q})\; dx_A dx_B dx_C \ldots dx_Q$$
k factors — k factors

and I derive this in the raw notes at the end of Chap 3.

Pause for emphasis: In $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ we have maybe $\varphi: R^n \to R^m$ with $n \neq m$. So the number of y_i is not the same as the number of x_i. *But*, for each $y_i = d\varphi_i$ in $d\varphi_I$ we get $dy_i = \; d\varphi_i = \partial_j \varphi_i dx_j$. Thus, for each of the k factors dy_i we get exactly one factor of dx_i. Thus, if we have k factors on the left above, we also have k factors on the right. The most k can be is dim(y) = m, but dim(x) = n and we would have m > n or m < n or m = n.

The determinant is that of a kxk submatrix of the full R matrix. The rows of this submatrix have the labels a,b,c...q while the columns have labels A,B,C...Q. The determinant only appears when you reorganize the sum into the ordered sum shown! This is a very major point!

Without this ordering you get instead

$$d\varphi_a d\varphi_b d\varphi_c \ldots d\varphi_q \;=\; \Sigma_{A,B,C\ldots Q}\; R_{abc\ldots q,ABC\ldots Q}\; dx_A dx_B dx_C \ldots dx_Q \quad .$$
k factors — k factors

Remember that all these things are really wedge products! Note that

$$R_{abc\ldots q,ABC\ldots Q} \equiv R_{aA} R_{bB} \ldots R_{aQ} \qquad \text{where} \qquad R_{aA} = \frac{\partial \varphi_a}{\partial x_A} \; = \partial_A \varphi_a \equiv (D\varphi)_{aA} \quad .$$

In a more compact notation we can write the above as

$d\varphi_I = \Sigma_J R_{I,J}\, dx_J = \Sigma_J (D\varphi)_{I,J}\, dx_J$ // non-ordered sum on J

$d\varphi_I = \Sigma_J \det(R_{I,J})\, dx_J = \Sigma_J [\det[(D\varphi)_{I,J}]\, dx_J$ // ordered sum on J

Note that the full R = Dφ matrix need not be square, but in either sum we are using just a square piece of that matrix which is k x k if we are doing k-forms. Bucks would refer to R = Dφ as the differential of the transform $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$.

The Final Act is then this:

$\alpha = \Sigma_I f_I(\mathbf{y})\, dy_I$

$\varphi^*\alpha = \Sigma_I f_I(\boldsymbol{\varphi}(\mathbf{x}))\ d\varphi_I$

$= \Sigma_I f_I(\boldsymbol{\varphi}(\mathbf{x}))\ [\ \Sigma_J \det[(D\varphi)_{I,J}]\, dx_J$

$= \Sigma_J \{\ \Sigma_I f_I(\boldsymbol{\varphi}(\mathbf{x}))\ \det[(D\varphi)_{I,J}]\ \}\ dx_J$

$= \Sigma_J \{g_J(\mathbf{x})\ \}\ dx_J$

$= \Sigma_J\, g_J(\mathbf{x})\, dx_J \qquad g_J(\mathbf{x}) \equiv \Sigma_I f_I(\boldsymbol{\varphi}(\mathbf{x}))\ \det[(D\varphi)_{I,J}]$ // p 42 A and B

$\equiv \beta$

Thus we have a "general formula" for the pullback β of an arbitrary k-form α. The pullback β is associated with x-space, while the original α is associated with y-space.

Chapter 4: Integration of 1-forms

4.1 Definition of the integral and its elementary properties [47]

I spend some time in the raw notes showing how the notation used here is an application of the general notation used in Chapter 3.

Chapter 3	Chapter 4	
x_i in x-space = R^n and U (p37)	t in t-space = R^1 and [a,b] (p47)	left
y_i in y-space = R^m and V (p37)	$\mathbf{x}$ in R^m and U	right
$\mathbf{y} = \boldsymbol{\varphi}(x)$	$\mathbf{x} = \mathbf{c}(t)$ curve [Buck $\mathbf{c}(t) = \boldsymbol{\gamma}(t)$]	
α	α	
φ*α	c*α [Buck c*α = α* = ω*]	

The main result is this, given that $\alpha = \Sigma_i f_i(\mathbf{x}) dx_i$ is a 1-form:

$$\int_{\mathbf{c}} \alpha \equiv \int_{[a,b]} c^*\alpha = \int_a^b g(t)\,dt = \int_a^b \Sigma_i f_i(\mathbf{c}(t))(\partial_t c_i)\,dt$$

In other words, the integral over the curve c in R^m of the form α is *defined* as the integral of the pullback c*α in R^1, the space on the left. This certainly is an integral we know how to do.

Example 4.1: For α = the angle form g(t) = 1 so c*α = g(t) dt = dt. then

$$\int_{\mathbf{c}} \alpha = \int_a^b g(t)\,dt = \int_0^{2\pi} 1\,dt = 2\pi \ .$$

Reparametrizing a Curve (I call it respeeding a curve).

Here is the main conclusion, where t = p(s) is the respeeding fuction with p' either + or -.

$$\int_{\mathbf{c}\,\circ\,\mathbf{p}} \alpha = \text{sign}(p'(s)) \int_{\mathbf{c}} \alpha \qquad \text{Theorem 4.3 [48]}$$

Apart from a sign, reparametrizing a curve does not change the 1-form integral! I prove it in raw notes.

[48] Interpretation of $\int_{\mathbf{c}}\alpha$ in $\mathbf{R}^3$. Recall α = **F**•d**x** is associated (one to one!) with vector **F**. If d**x** is a physical distance in physical space, and if **F** is the force on a particle, then α is the work done on that particle as it moves d**x**. For example if **F** and d**x** are aligned, you are doing positive work on the particle, you are accelerating it or maybe increasing its potential energy. So our integral is then specifically "the total work done on the particle" as it traverses the curve in the presence of this force field **F** .

Question: Where does the Buck arc-length type integral fit into this picture? Buck p 320 in our Sja notation reads

$$L(\mathbf{c}) = \int_a^b |\,\mathbf{c}'(t)\,|\,dt = \int_a^b |\,\partial_t \mathbf{c}(t)\,|\,dt = \int_a^b \sqrt{\Sigma_i\,[\partial_t c_i]^2}\,dt$$

Then later on page 367 we see the integral of a function over a curve,

$$\int_{\mathbf{c}} f = \int_a^b f(\mathbf{c}(t))\,|\,\partial_t \mathbf{c}(t)\,|\,dt = \int_a^b f(\mathbf{c}(t))\sqrt{\Sigma_i\,[\partial_t c_i]^2}\,dt \qquad \text{(7-1) Buck}$$

Neither of these integral types fit into the current Sjamaar framework. I guess these integrals are not integrals of 1-forms. Maybe Sja will have something to say about this soon. // Yes, it is coming, but far in the future! Maybe chapter 7 and chapter 9.

4.2 Integration of exact 1-forms [49]

Theorem 4.4: If α is exact (meaning α = dg) then define h(t) = (**c***g)(t) to show that

$$\int_{\mathbf{c}} \alpha \equiv \int_{[a,b]} c^*\alpha = \int_{[a,b]} dh = h(b) - h(a) = g(\mathbf{c}(b)) - g(\mathbf{c}(a))$$

This says that the integral of an exact 1-form is a function only of the coordinates of the endpoints of the curve which are **c**(**a**) and **c**(**b**). The curve can take any path you want between these endpoints, and the resulting integral is the same. That is rather dramatic, really. Like complex integration contour. The closest "integral theorem" I have is this

$$\int_{\mathbf{A}}^{\mathbf{B}} \nabla\varphi \bullet d\boldsymbol{\ell} = \varphi(B) - \varphi(A) \qquad (9) \qquad \text{MM 4-36 p 156}$$

and you can see the analogy. In my theorem you integrate $\nabla\varphi$ and φ appears on the right. In Theorem 4.4 you integrate dg and g appears on the right.

Theorem 4.5. Three ways to say the same thing, with fairly long proofs.

(i) α is exact, meaning there exists g such that $\alpha = dg$

(ii) $\int_c \alpha = 0$ for any closed curve c [seems obvious given the above Thm 4.4 which is (iii) below]

(iii) $\int_{\mathbf{c}} \alpha$ depends only on the endpoints of the curve c.

Theorem of Exercise 2.6 Any closed 1-form on (*all of*) R^n is exact. I went back and proved this and I concluded as well that the theorem is true for U = an open cube centered at the origin in R^n. The proof requires that $\partial_j f_i$ be defined in all of U. If you knock out the origin, for example, then $\partial_j f_i$ is not defined there, so proof fails. This then explains why the angle form which is a 1-form is closed but not exact.

Pause to do some comparisons between Sjamaar and Buck.

Sjamaar	Buck	
α	ω	a generic differential form
c	γ	a curve
$\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$	$\mathbf{x} = \boldsymbol{\varphi}(\mathbf{u})$	a transformation
$\varphi^*\alpha$	ω^*	a pullback

4.3 The global angle function and the winding number [51]

It is shown that you can extend the angle θ which ranges 0 to 2π in this manner, where α is the angle form. That is, once in pull-back form, no problem with large θ range, and use that as definition.

$$\theta(t) = \theta(0) + \int_{\mathbf{0}}^{\mathbf{t}} c^*\alpha = \text{the pull back of } \alpha \text{ relative to c'}$$

If a closed curve winds around the origin k times, you will find that $\Delta\theta = k(2\pi)$ and k = winding number. Again this seems to have a connection to complex variable theory which also has a winding number.

Chapter 5: Integration and Stokes' theorem

In Chapter 4 we only did integrals of 1-forms. Here we have the general theory for integrating k-forms.

Overview added 2.14.16. Below we are going to have only cubes in the domain (or chains of cubes), so we have a very limited domain of interest. Under the mapping $\mathbf{x} = \mathbf{c}(\mathbf{t})$ a true cube in t-space maps into a warped cube in x-space. This warped cube has the same dimensionality k as the domain cube (generally speaking), and the warped cube can be imagined to be embedded in some large R^N space which we don't talk about much. For example, if k = 2, then we are mapping the interior of a square in t-space into a floating 2D patch in R^N which I call a "warped square" as pictured below which happens to show everything in the plane of paper (the warped square on the right is not generally in plane of paper). So the surfaces in R^N of dimension k are in a very special class in this chapter: they are surfaces that are obtained starting with a cubic domain in t-space and mapping through $\mathbf{x} = \mathbf{c}(\mathbf{t})$. We are going to talk about integrating over the warped patch in x-space (think patch on a toroid) and this integral is defined as the pullback integral over the cube in t-space. So yes, we are integrating k-forms here over a certain selected set of surfaces in R^N which have dimensionality k and which are the images of perfect k-cubes. Integrating a k-form over *anything* is a step forward in our learning, but we are NOT yet integrating over a manifold M, because a general manifold M is probably not going to be the image of a perfect cube or even of a chain of such cubes. Below we will eventually prove the Ch 5 Stokes' Theorem which is this,

$$\int_{\mathbf{c}} d\alpha = \int_{\partial \mathbf{c}} \alpha$$

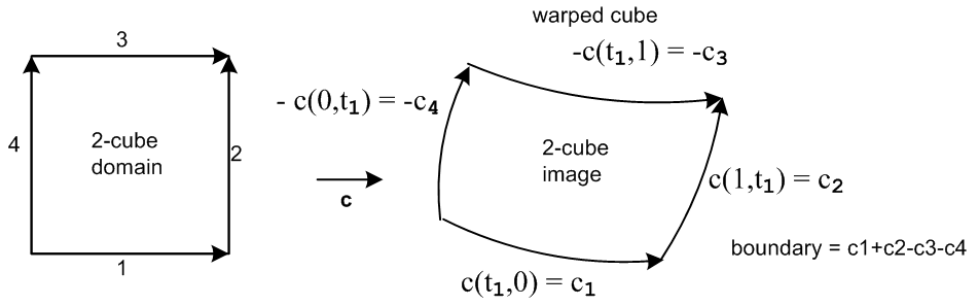
$d\alpha$, c = k-form, k-chain

α, ∂c = (k-1)-form, (k-1)-chain

On each side of this equation we are integrating some q-form over a q-cube or q-chain of q-cubes, where it happens that q = k on the left side and q = k-1 on the right side. This theorem requires the notion of a boundary ∂c of a q-cube c, and lots of the text below deals with this boundary notion. When we are done, this version of Stokes' Theorem only applies to q-cubes and q-chains. We do NOT have Stoke's Theorem on a Manifold. We don't even know yet how to integrate anything on a Manifold. We haven't even defined a Manifold at this point in Sjamaar. That is done in Chap 6, and then we will learn how to integrate a form over a manifold in Chap 7,8,9 so we have a long way to go after Chapter 5.

5.1 Integration of forms over chains

In the raw notes I spend a lot of time converting the pullback notation of Chapter 3 to the notation used in this chapter. The old mapping $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ is now $\mathbf{x} = \mathbf{c}(\mathbf{t})$. In Ch 3 we had $\boldsymbol{\varphi}: R^n \to R^m$, but in Ch 5 we do $\mathbf{c}: R^k \to R^n$ where $n \geq k$. Moreover, the domain for our mapping rather than being some general shape will be the unit cube in R^k which we write as $[0,1]^k$. Sja starts out with R ≡ arbitrary block in R^k, but later when we learn about reparametrization, we might as well take that block to be $[0,1]^k$. So we are then doing c: $[0,1]^k \rightarrow$ k-dimensional image within R^n. I refer to this image as a "warped cube" just to have something easy to visualize. For k = 2 for example



As the t_i run over the source cube, the image runs over a warped cube. In the above, a square maps into a 2D patch which is our warped 2-cube. Wee are interested not just in 1-forms, but in k-forms, so we have

dc = det[(**Dc**)] dt **c**: **t** → **x** k factors on each side where:

$\alpha = f(\mathbf{x})\, dx$ $dc \equiv \Pi_{i=1}^{k} dc_k$ $dx \equiv \Pi_{i=1}^{k} dx_k$ $dt \equiv \Pi_{i=1}^{k} dt_k$

We want to "pull back" our form α from x space to t space, so

$\alpha = f(\mathbf{x})\, dx$
$c^*\alpha = f(\mathbf{c}(\mathbf{t}))\, dc = f(\mathbf{c}(\mathbf{t}))\, \det(Dc)\, dt \equiv g(\mathbf{t})\, dt$

Our integral of interest will be this

$\int_c \alpha$ = integral of a k-form over the k-dimensional image region c of the unit k-cube (in R^n)

$= \int_{[0,1]^k} c^*\alpha = \int_{[0,1]^k} f(\mathbf{c}(\mathbf{t}))\, \det(Dc)\, dt = \int_{[0,1]^k} g(\mathbf{t})\, dt$

The term **k-cube** really refers to the *mapping* $c:R^k \to R^n$ but you can think of it also as the image of that mapping, like the 2D patch shown above on the right. [Similarly, the mapping $c:R^1 \to R^k$ is a curve, but we usually think of the curve as the image of the mapping.]

The reason we are working on this particular *kind* of $\int_c \alpha$ is because that is the kind of integral that will be appearing in Stokes's theorem when we get to it! But more generally, we want to know abstractly how one integrates a k-form.

Reparametrization. Sja shows (and I proved) that if you reparametrize **t** = **p**(**s**), you find that

$\int_{c\,o\,p} \alpha = \text{sign}[\det(Dp)] \int_c \alpha = \pm \int_c \alpha$

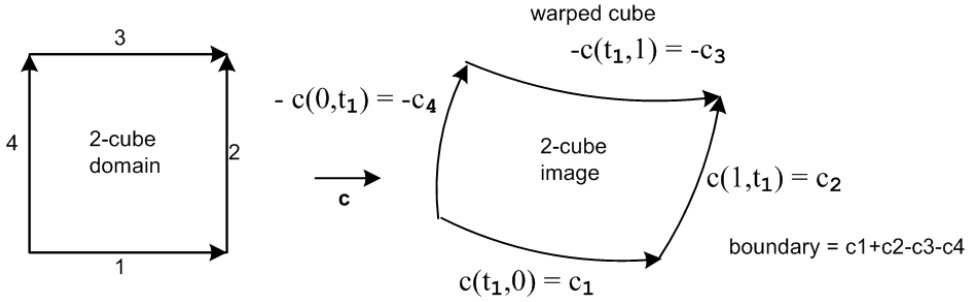
so at most you get a sign difference depending on the "orientation" of your mapping (does it flip the k-cube inside out, right handed to left handed, whatever).

A **k-chain** is simply any linear combination of k-cubes, a no brainer. In proofs, you generally show that something is true for a k-cube, then it is usually easy to see that the result then applies to any k-chain.

5.2 The boundary of a chain

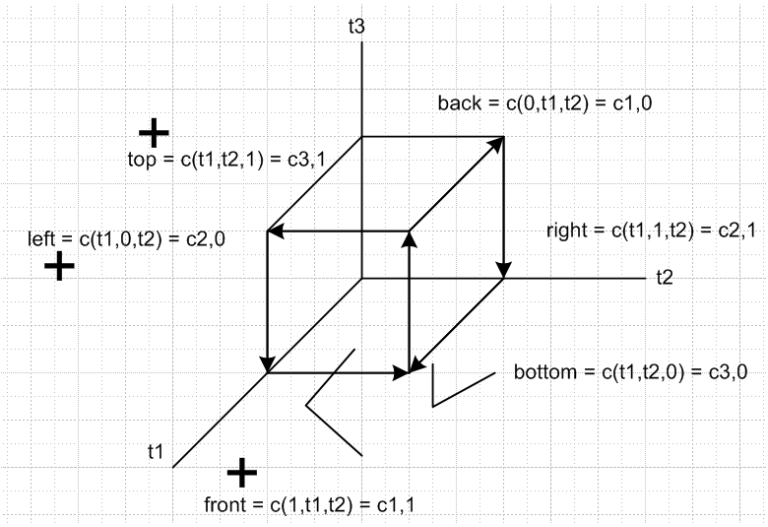
Very important because boundary is going to appear also in Stokes's theorem as ∂c. We start with k = 1, and two endpoints which are the boundary. The tail endpoint gets a minus sign, very important!

Now move to k =2 so square goes to patch as shown above and repeated below:



The mapping of each face (edge) of the cube is a mapping $R^2 \rightarrow R^2$ so now we are allowed to have planar curves for edges of the warped image cube. The arrows on the left show direction of the t_i variable increase, and those arrows appear on the right. But we want a boundary on the right that goes around in a nice CCW manner, so that is why we get ∂c = c1+c2-c3-c4 so two of the edges have minus signs because we have to reverse them. The opposite side faces (edges here) are pairs and in any pair, one edge must have a minus sign. [Note: we are taking the union of the four edges (sets) to make a boundary, but I guess since these are functions, union is the same as addition.]

What happens with k = 3 and a normal 3D cube? You *cannot* arrange that all 6 faces have a clean oriented boundary, as this picture shows, If you orient the front and right face, the top is no go.



The boundary here is defined in a tricky way so that the top left and front are +, while the other three are -. Again, for each opposite side pair of faces, one is + and one is -, just as in the 2D case. The definition seems arbitrary with regard to the signs relative to pair faces.

We pause now to talk about **double-subscript face notation**:

$$c_{j,\rho} = \quad [c(t_1,t_2,t_3..\ t_{j-1},t_j,t_{j+1}..t_k)]_{j,\rho} \rightarrow c(t_1,t_2,t_3..\ t_{j-1},\rho,t_{j+1}..t_k) \rightarrow\ c(t_1,t_2,t_3..\ t_{j-1},\rho,t_j..t_{k-1})$$

so that

$$c_{j,\rho} \equiv\ [c(t_1,t_2,t_3..\ t_{j-1},t_j,t_{j+1}..t_k)]_{j,\rho}\ =\ c(t_1,t_2,t_3..\ t_{j-1},\rho,t_j..t_{k-1})$$

You first replace the t_j coordinate with $\rho = 0$ or 1 (meaning you are talking either of two opposite faces of the face pair in the t_j direction). You then lower all the indices on the t_i which occur *after* the place where the ρ inserted. Here are some examples for the 3D cube:

$$[c(t_1,t_2,t_3)]_{1,\rho} = c(\rho,t_1,t_2) \qquad \mathbf{t} = t_1,t_2 \qquad \rho = 0 \text{ or } 1$$
$$[c(t_1,t_2,t_3)]_{2,\rho} = c(t_1,\rho,t_2)$$
$$[c(t_1,t_2,t_3)]_{3,\rho} = c(t_1,t_2,\rho) \qquad (1.1)$$

and also for the 2D cube shown above earlier.

$$[c(t_1,t_2)]_{1,\rho} = c(\rho,t_1) \qquad \text{// 2 faces of 2-cube (two sides of a squarish patch)}$$
$$[c(t_1,t_2)]_{2,\rho} = c(t_1,\rho)\ . \qquad \text{// 2 faces of 2-cube (two sides of a squarish patch)}$$

The reason for the second step is that you need 1 fewer variables to describe a face.

Here then is the postulated general expression for the boundary of a k-cube

$$\partial c = \ \Sigma_{i=1}^{k}\ \Sigma_{\rho=0,1}\ (-1)^{i+\rho} c_{i,\rho} \qquad \text{p 60 D}$$

The only mystery here is how the faces are assigned signs! The sum makes sense: there are k pairs of opposite faces labeled by ρ=0 and ρ=1. The formula works for k = 1 and k = 2, so somehow we assume it is correct for higher k. I guess it is really a definition for $k \geq 3$. We are sure of the faces, and here we assign signs to each face in a certain simple manner.

[60] Theorem 5.4 $\partial(\partial c) = 0$ "the boundary of any boundary is always 0"

I go through the proof in detail [in my "detail work... doc"], it follows simply from the above ∂c definition. The proof necessarily involves the recursive idea of $[c_{i,\rho}]_{j,\sigma}$ which is a face of a face of a k-cube. In other words $c_{i,\rho}$ is the face of a k-cube which has dimension k-1, and then $[c_{i,\rho}]_{j,\sigma}$ is a face of that (k-1)-cube which is a (k-2)-cube.

It is curious that we have both $\partial^2 = 0$ for boundaries, and $d^2 = 0$ for forms.

5.3 Cycles and Boundaries

If a k-cube image does not depend on one of its t_i variables, then two opposite faces are identical, and the k-cube is really a deflated (k-1)-cube of some kind. It has had the air sucked out and has been crushed into a lower dimensional object. So from our $\int_c \alpha$ perspective stated above, there will be no contribution from such a k-cube integration. This seems intuitively clear to me (but is proven by Sja). Such a k-cube is **degenerate**. I think it is degenerate if 1 *or more* coordinates are crushed out.

A k-chain is degenerate if *all* its k-cubes are degenerate.

Lemma 5.5. The integral $\int_c \alpha = 0$ if c is a *degenerate* k-chain and α is a k-form.

I prove this in the raw notes by showing that the Jacobian object in the pullback form vanishes because the R matrix has a column of all zeros.

Lemma 5.7: The boundary of a degenerate k-cube is a degenerate (k-1)-cube. [I prove this in raw notes]

Definition: If the *boundary* of a k-chain (which includes a k-cube as a special case) is *degenerate*, then one says that the k-chain is **closed** or is a **cycle**.

Meanings of the word "closed": earlier we had sets being closed$_1$ or open, in Stak we have operators being closed$_2$, then we had differential forms being closed$_3$ $d\alpha = 0$, and *now* we have a k-chain being closed$_4$. A curve having no endpoint is also called closed$_5$. We also have the notion of a closed$_6$ surface which contains all its limit points, similar to a closed$_1$ set.

Definition: If you have $c = \partial b + c'$ where b is a k-chain, and where c' is a *degenerate* (k-1)-chain, then we can just ignore c' from a k-form integration point of view (lemma 5.5) and this is as if $c = \partial b$ and we then say that c is the **boundary** of b, despite the existence of the extra degenerate piece c'.

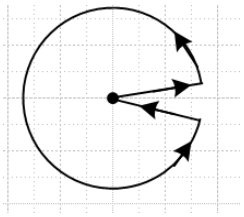
Example 1: Combine **two curves** so they match at the ends to make a closed$_5$ loop. In this case there is no boundary at all, and we include this in the meaning of a degenerate boundary. Thus, this closed$_5$ curve (no boundary) is regarded as closed$_4$ or a cycle, where we now have two meanings for the word closed. So this is Sjamaar's first example of a k-chain that is closed: it happens to be a 1-chain so k = 1.

Example 2: Consider a **disc** (even a warped one) having a boundary which is a closed$_5$ curve. The disk is a 2-chain, and that boundary is a 1-chain, it has the right dimension, it is not degenerate, so such a disk is then not "closed$_4$ or a cycle". It is not closed$_4$ because its boundary is not degenerate. However, such a disk is closed$_1$ as a set.

Example 3: Consider a **sphere** with a normal spherical shell being the boundary. This is a 3-cube with a non-degenerate and normal boundary, so such a 3-cube is not closed$_4$. But as a set, the surface is closed$_6$ because every limit of a sequence on that surface lies on the surface, so there are no set type boundary points of that surface. The sphere including the surface is also a closed$_1$ set. So nomenclature is rather messy here.

Example 4: Consider a sphere and a piece of great circle. This composite object has a normal boundary (the 2D shell) and this extra 1D great circle which would contribute nothing to a 2D integration over the surface. So this piece of great circle would be an example of c' above.

Example 5: (Pac-Man) Suppose your warped 2-cube patch is a disk. Then your two parameters could be r and θ, and you end up with this situation where I keep the wedge slightly open so it can be seen.



There are two pairs of edges, just as for the regular patch. In the limit as you extend the domain to its entirety in r-θ space, the two ray edges cancel each other out in the range (image). The two other edges are the circle and the point at the origin which is a circle compressed to a point. So the boundary of this object is the outer circle plus a degenerate boundary c' at the center. We can ignore that point, and we can then say that $c = \partial b + c' \approx \partial b$ so the circle is still the boundary of the disk by the official definition of boundary just stated.

Lemma 5.7. If c' is a degenerate k-chain, so is ∂c'. Simple proof.

Theorem 5.8: Every boundary is a "closed$_4$ or a cycle".

This does not say that any k-chain is a cycle. It says that the *boundary* of any k-chain is a cycle, with boundary as defined above. Thus, for example, the circle which is the boundary of a disk is "closed$_4$ or a cycle". In fact this circle has no boundary at all, so the circle itself as a 1-chain is "closed$_4$ or a cycle". If the Pac-Man central point is included, it makes no difference by definition of "boundary".

Proof: Consider $c = \partial b + c'$ where c is boundary of b. Apply ∂ to get $\partial c = \partial c'$. Since c' is degenerate, so is ∂c' by the above lemma 5.7. Therefore ∂c is degenerate. By the definition above, c is then a cycle (it is closed).

Comment: So we find that every boundary of a k-chain is closed$_4$ or a cycle. The circle around a disk or the shell around a sphere are both closed$_4$ or a cycle (in their dimension). You might think that you could concoct some k object whose boundary was not closed/cycle, but you cannot.

Example: Consider a 3-cube whose boundary is 6 faces which make up a 2-chain. What is the boundary of this boundary? This 2-chain contains all its limit points and is a closed$_1$ = closed$_6$ set, just like the circle or the shell for the disk or sphere. Thus, this 2-chain has no boundary whatsoever, and this is the same idea as the 2-chain which is the boundary of a sphere having itself no boundary. So this is just another example of a boundary being "closed or a cycle". You cannot get to an edge of such a boundary. In fact, you cannot get to the edge of any boundary. A boundary is always "continuous " in some sense at any point and in all allowed directions. This presentation could use improvement.

5.4 Stokes's Theorem

Here is the main point of this chapter.

$$\int_c d\alpha \;=\; \int_{\partial c} \alpha \qquad\qquad d\alpha,\ c = \text{k-form, k-chain}$$
$$\alpha,\ \partial c = \text{(k-1)-form, (k-1)-chain}$$

On the left, the form $d\alpha$ is assumed to be a k-form and it is integrated over c which is a k-chain. Up to now we have thought of $\int_c \alpha$ where α was a k-form, here we have a *different* α. So with this slight notation change in the meaning of α, the left integral $\int_c d\alpha$ is the integral of a k-form over a k-chain, and is exactly the kind of integral studied in the previous sections (with $\alpha \to d\alpha$).

Now if $d\alpha$ is a k-form, we know that α must be a (k-1)-form. And we know that since c is a k-chain, the boundary ∂c is a (k-1)-chain. Thus, the integral on the right is also in the same class of integral as we have been studying.

The difference is that the left integral is k over k, whereas the right integral is k-1 over k-1.

Here is a *summary* of my interpretation of the Sja proof of the above Stokes' theorem: [see raw for deets]

$$\int_c d\alpha \;=\; \int_{\partial c} \alpha \qquad\qquad d\alpha,\ c = \text{k-form, k-chain}$$
$$\alpha,\ \partial c = \text{(k-1)-form, (k-1)-chain}$$

Step 1: $\int_c d\alpha = \int_{[0,1]^k} c^*(d\alpha) = \int_{[0,1]^k} d(c^*\alpha)$ k over k

pullback form $\varphi^* d = d\varphi^*$ p 39

Step 2: $c^*\alpha = \text{(k-1)-form} = \Sigma_i\, g_i(\mathbf{t})\, dt_1 \wedge dt_2 \wedge \ldots dt_i \ldots \wedge dt_k$

$$d(c^*\alpha) = \Sigma_i\, (dg_i)\, dt_1 \wedge dt_2 \wedge \ldots dt_i \ldots \wedge dt_k = \Sigma_i \Sigma_j\, [(\partial_j g_i) dt_j]\; dt_1 \wedge dt_2 \wedge \ldots dt_i \ldots \wedge dt_k$$

$$= \Sigma_i\, (-1)^{i+1}\, (\partial_i g_i)\, dt_1 \wedge dt_2 \ldots \wedge dt_k \qquad \text{// slide } dt_i \text{ into place, j forced = i}$$

Step 3: $\int_c d\alpha = \int_{[0,1]^k} d(c^*\alpha) = \Sigma_i\, (-1)^{i+1} \int_{[0,1]^k} (\partial_i g_i)\, dt_1 \wedge dt_2 \wedge \ldots \wedge dt_k$

Step 4: $\int_c d\alpha = \int_{[0,1]^k} d(c^*\alpha) = \Sigma_i\, (-1)^{i+1} \int_{[0,1]^k} (\partial_i g_i)\, dt_1 dt_2 \ldots dt_k$ // no ^

Step 5: $\int_c d\alpha = \Sigma_i\, (-1)^{i+1} \int_{[0,1]^{k-1}} \left[\int_{[0,1]} (\partial_i g_i)\, dt_i \right] dt_1 dt_2 \ldots dt_i \ldots dt_k$

Step 6: $\int_c d\alpha = \Sigma_i (-1)^{i+1} \int_{[0,1]^{k-1}} [[g_i]_{i,1} - [g_i]_{i,0}] dt_1 dt_2 \ldots dt_i \ldots dt_k$

$= \Sigma_i \Sigma_\rho (-1)^{i+\rho} \Sigma_\rho \int_{[0,1]^{k-1}} [g_i]_{i,\rho} dt_1 dt_2 \ldots dt_i \ldots dt_k$

$= \Sigma_i \Sigma_\rho (-1)^{i+\rho} \Sigma_\rho \int_{[0,1]^{k-1}} [g_i]_{i,\rho} dt_1 \wedge dt_2 \wedge \ldots dt_i \ldots \wedge dt_k$ // add $\wedge$

Step 7: $[\mathbf{c}_{i,\rho}]^* \alpha = \Sigma_i [g_i]_{i,\rho} dt_1 \wedge dt_2 \wedge \ldots dt_i \ldots \wedge dt_k$

Step 8: $\int_c d\alpha = \Sigma_i \Sigma_\rho (-1)^{i+\rho} \Sigma_\rho \int_{[0,1]^{k-1}} [\mathbf{c}_{i,\rho}]^* \alpha$

$= \Sigma_i \Sigma_\rho (-1)^{i+\rho} \Sigma_\rho \int_{c(i,\rho)} \alpha$

$= \int_{\partial \mathbf{c}} \alpha$ // using page 60 D for ∂c expression

I am not totally thrilled with this proof, since the sign delicately rests on when you add or remove the wedges which Sja does not even talk about (yet). But it is a pretty good proof still.

See raw notes for closing comment on circle case.

I then do **Exercise 5.6** to show now the general Stokes' theorem generates all the famous integral theorems:

$\int_c \nabla f \bullet d\mathbf{x} = f(\mathbf{c}(1)) - f(\mathbf{c}(0))$ "other theorem"

$\int_c (\partial_x g - \partial_y f)\, dxdy = \int_{\partial c} (fdx + gdy)$ ∂c = a curve "Green's theorem"

$\int_c [\text{div } \mathbf{F}]\, dV = \int_{\partial \mathbf{c}} \mathbf{F} \bullet [*d\mathbf{x}]$ divergence theorem (Gauss's theorem)

$\int_c \text{curl } \mathbf{F} \bullet (*d\mathbf{x}) = \int_{\partial \mathbf{c}} \mathbf{F} \bullet d\mathbf{x}$ Stokes's theorem

Chapter 6: Manifolds

6.1 The definition of a manifold

The transformation of interest is written as $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ with ψ: $R^n \rightarrow R^N$ with $N \geq n$. Variable $\mathbf{t}$ runs over set U in R_n and in doing so creates a surface of dimension n in R^N, and it is this "surface" which may or may not be an embedding or a manifold. [elsewhere we have n being m, but here it is n], but I call this surface M regardless of what it is.

I digressed a lot studying the matrix ($D\psi$) which is my R matrix. In this matrix, the functions ψ_i go down while the variables t_i go across. Thus, there are N rows and n columns, so R matrix is "tall", different from Lagrange doc. It is an N x n matrix.

In support doc "the meaning of" I show that the matrix equation $\mathbf{v} = M\mathbf{u}$ can be solved for $\mathbf{u}$ even when matrix M is non-square iff the matrix M has a left inverse. For an m x n matrix M (m rows and n columns), a left inverse exists only if n < m (tall) and the matrix has full rank n (I quote this from wiki without examination; my matrix notes binder really deals only with square matrices). Thus, our R matrix $R = (D\psi)$ has a left inverse if $n < N$ which is in fact the case. Since you can then write $\mathbf{u} = M^{-1}\mathbf{v}$, you know that the matrix is 1-to-1, and thus our matrix ($D\psi$) is one-to-one if it is full rank. This sense of ($D\psi$) being 1-to-1 is part of the definition of an embedding and a manifold below.

I show that the columns $\mathbf{c}_i$ of matrix ($D\psi$) are the unnormalized tangent vectors at a point $\mathbf{x}$ on the surface $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$. I sometimes call these vectors ξ_i but not often. If ($D\psi$) has full rank, then the columns are linearly independent, and thus so are the basis vectors of the tangent space at $\mathbf{x}$ on the surface. Since this is desirable, one of the conditions for an embedding is that ($D\psi$) by 1-to-1 which means it has full rank as just shown above. Saying again: When ($D\psi$) by 1-to-1 the column vectors of ($D\psi$) span T_xM.

In notes added to the raw notes I finally got happy with certain Sjamaar equations. For example page 67 B says that

$T_x[\psi(U)] = [(D\psi)(\mathbf{t})](R^n)$ $\qquad$ $M = \psi(U)$ = n-dimensional surface of interest

This equation is a **challenge for me** to understand, each time I have to pause and write out the meaning, and I will do that again right here.

The starting point is to regard $(\mathbf{t},\mathbf{x})$ as the reference point pair so $\mathbf{x}$ is some point on M. There is a particular tangent space that goes with this point pair, and it is called T_xM.

The object $[(D\psi)(\mathbf{t})]$ is a *matrix* of derivatives of the ψ_i with respect to the coordinates t_i. The $(\mathbf{t})$ part of this notation does *not* mean the matrix $D\psi$ is acting on a vector $\mathbf{t}$. It means that the elements of the matrix ($D\psi$) are functions of the reference point location $\mathbf{t}$.

There are N of the ψ_i since $x_i = \psi_i(\mathbf{t})$ defines the surface M within R^N (of dimension n). There are n of the t_i making up $\mathbf{t}$. The matrix has the ψ_i going down and the t_i going across, so it has N rows and n columns where $n \leq N$, so this is a "tall" matrix.

Now keeping reference point $\mathbf{t}$ fixed, think of U space as having points $\mathbf{t}' = \Sigma_i t'_i \mathbf{e}_i$ where $\mathbf{e}_i$ are basis vectors for U space. Each such point $\mathbf{t}'$ gets mapped to some point in the tangent space T_xM (not to a point on M!) by the matrix $[(D\psi)(\mathbf{t})]$. Thus it makes sense to write $\mathbf{q} = [(D\psi)(\mathbf{t})](\mathbf{t}')$ where $\mathbf{q}$ is then some point in the tangent space T_xM (probably not a point on M). It then it also makes sense to write

$[(D\psi)(\mathbf{t})](U)$ = a *set* of points $\mathbf{q}$ = a continuous *portion of* the tangent space T_xM

Now if we were to take U to be all of the domain space R^n , then the above equation would *exhaust* the entire tangent space and we would write the above as

$[(D\psi)(\mathbf{t})](R^n)$ = a *set* of points **q** = all of the tangent space T_xM

Thus, it makes complete sense to say

$[(D\psi)(\mathbf{t})](R^n) = T_xM$.

But of course $M = \psi(U)$: Here U is the domain in R^n, and $\psi(U)$ is then our surface M of dimension n lying in R^N. So

$[(D\psi)(\mathbf{t})](R^n) = T_xM = T_x\psi(U)$

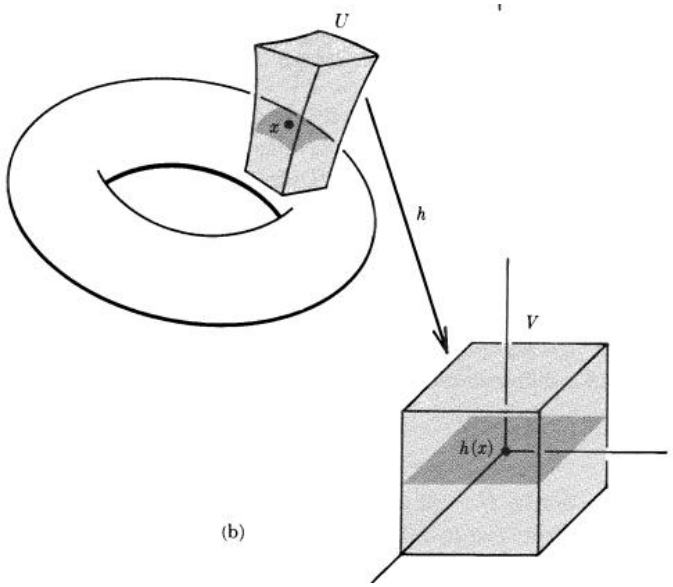
and this is exactly what p 67 B says!!! There is only one reference point **t** (or **x**), so this is NOT some kind of fiber bundle for just one **t**. For a particular fixed **t**, each side of this equation is a set of points in R^N. That set of points exhausts a "flat" surface T_xM which is tangent to M at the point **x** on M.

This form in fact appears much later in Chapter 8 as page 96 item D

Definition of an Embedding: One embeds U into R^N if the ψ map is smooth (C^∞), if its inverse is C^1 at least, if ψ is 1-to-1, and if $(D\psi)$ is 1-to-1. If ψ is not 1-to-1, ψ describes a self-intersecting surface, and then the embedding is reduced to being only an **immersion**.

Example 6.3. Here we study a special case where $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t}))$ where f is C^∞. This case for obvious reasons is called the **graph** of f. I show that this thing does in fact give an embedding regardless of whether or not f is 1-to-1. I get to use my Mini-Columns Theorem from Lagrange doc.

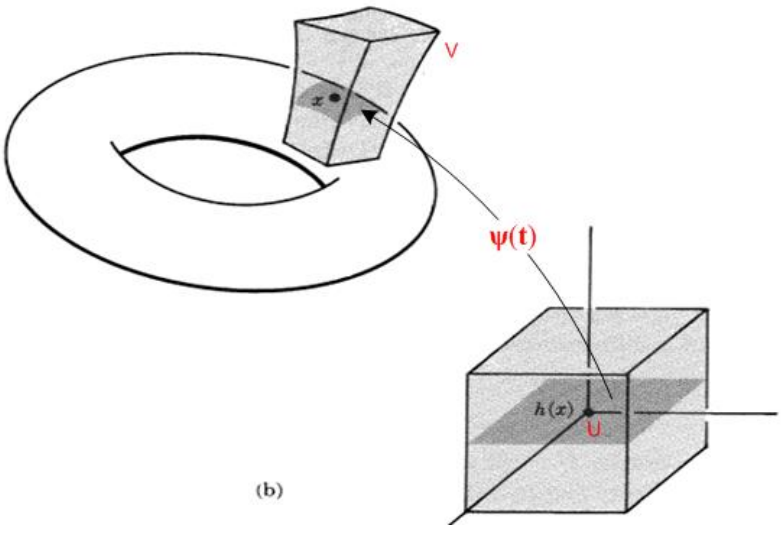
My notes digress at this point into the **Spivak** book. Spivak first defines a **diffeomorphism** between two sets in R^n . This is a mapping between those two sets which is super-smooth C^∞ in both directions. Spivak then uses this concept to define a manifold (his first of two definitions) and provides this excellent picture:



Here the manifold is the torus surface and h is the diffeomorphism. The idea is that the dark gray areas on each end are surrounded by "elbow room" so you can talk about derivatives in all directions.

Spivak goes on to give an "alternative" definition of a manifold which is basically the same as that given by Sjamaar (coming soon). But the two authors use different notations so things are a little confusing. In this second definition, one does not think of the dark square on the right above as being part of a larger space, you just think of it all by itself. Then the dark gray area is $U \subset R^n$ in Sjamaar, and the torus is embedded in R^N. (Note that Spivak uses symbol U to denote a different object in the picture above.)

Here I first adjust the above Spivak picture to make it agree with Sjamaar notation:



Definition of a Manifold: You consider a tiny open domain region U in the R^n space (a square say like the dark square above), and you consider a tiny open region V in R^N surrounding the point **x** on the surface $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$. In the above picture, the dark gray square on the right would be U and the light warped

cube on the top would be V, which surrounds **x** in $R^{\mathbf{N}}$. If this mapping is an embedding for *any* tiny U you might select in a larger domain of interest, then overall the surface **x** = **ψ**(**t**) is then a manifold. The patch on the manifold surface M can be written as ψ(U) = V∩M . Here is direct from Sjamaar:

6.4. DEFINITION. An *n-dimensional manifold*[1] (or *n-manifold* for short) in $\mathbf{R}^N$ is a subset M of $\mathbf{R}^N$ such that for all $\mathbf{x} \in M$ there exist

- an open subset $V \subseteq \mathbf{R}^N$ containing $\mathbf{x}$,
- an open subset $U \subseteq \mathbf{R}^n$,
- and an embedding $\psi\colon U \to \mathbf{R}^N$ satisfying $\psi(U) = V \cap M$.

The Basic Idea is this: if the surface M is *locally* an embedding at all points, then it is globally a manifold. None of the local test points can have a self-intersection of M, for example, so if global M has a self-intersection, this embedding test would fail near that location, and M would not be a manifold. The embedding idea means things are smooth (asymmetrically since C^{∞} forward and $C^{\mathbf{1}}$ reverse) and at any point on M you have a non-degenerate set of tangent vectors, so the tangent space is always nice. You can imagine that if you want to integrate over M, you might need the tangent space to be clean everywhere on M.

Although you are always allowed to "cover" your manifold with as many little patches as you want, in practice some small total number of embeddings works. I show that for a circle as a manifold, you could have only two embeddings. Each of these embeddings is called a **chart**, and the set of them an **atlas**. The model here is that the earth surface is a manifold, and we know we can cover it with a finite set of charts. The charts usually overlap. Chart is also called a **coordinate map.**

The number N-n is called the **codimension** of the manifold, and is equal to 3-2 = 1 for the torus example.

Sjamaar then gives us lots of examples of manifolds. One example is the surface of a sphere $S^{\mathbf{n-1}}$ embedded in $R^{\mathbf{n}}$.

I think the reason that you need the "elbow room" region V around **x** (see second picture above) is that the tangent space $T_{\mathbf{x}}M$ close to **x** is part of this region (it goes off M), and you might have to do derivatives or some related activities in the region V.

6.2 The regular value theorem

We now consider a *totally different* way to think of (and define) a manifold. In the above we had **x** = **ψ**(**t**) map out the manifold M surface as **t** runs around in domain U. Thus M = ψ(U). Now, instead of using **x** = **ψ**(**t**) , this section studies the vector equation **c** = **φ**(**x**) where **c** is a constant vector and the idea is that for a fixed vector **c**, this equation describes some surface M in $R^{\mathbf{N}}$ which might be a manifold M. In **c** = **φ**(**x**) there are m equations and variable **x** has N components.

We can define a new matrix (Dφ) which has m rows and N columns, and we are interested in N > m, so this matrix R = (Dφ) is "wide", unlike (Dψ) in the previous section, but like in Lagrange doc. Now we have the $\varphi_{\mathbf{i}}$ going down (so # rows = m) and the $x_{\mathbf{i}}$ going across (so # cols = N).

We know that $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ defines a surface within $R^{\mathbf{N}}$ of dimension N-m ≡ n. The question is this: when is this implied surface a manifold? Sjamaar sometimes writes this surface using the strange notation $M = \varphi^{-1}(\mathbf{c})$ where M would then be the locus of **x**, our candidate manifold M. He notes that φ normally does not have an inverse in the function sense, so this is just a special notation. We have a different surface in $R^{\mathbf{N}}$ for each value of **c**. This topological structure (for a set of **c** values) is a level set of surfaces, sometimes called a fiber (perhaps a fiber bundle), but Sja does not dwell on that terminology.

Now we study this matrix (Dφ). The maximum possible rank is m since it is an m x N matrix. Since this is really (Dφ)(**x**) [like my R(**x**) in Lagrange doc], the rank can vary with **x**. [Compare to (Dψ)(**t**)] This is the same idea as in Buck for matrices. If for some **c** (Dφ)(**x**) has full rank for all **x** on the surface $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$, then **c** is a **regular value** of c. If on the other hand the rank falls below full rank at some value of **x**, then **c** is a **singular value**. Recall that for Bucks, if a matrix is a function of some coordinate, a point where the rank falls below full rank is called a critical point.

Regular Value Theorem (Theorem 6.10). This deals with the question of whether or not the surface implied by $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ is or is not a manifold. The theorem states:

(1) for any *regular value* of **c**, $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ is in fact a manifold
(2) The tangent space at point **x** is the nullspace of (Dφ).

The proof is a bit involved, and makes use of several ideas developed elsewhere. The equation $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ is first written as $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{u},\mathbf{v})$ where **u** has dimension n and **v** has dimension m = N-n. One asks whether this equation defines an "implicit function" $\mathbf{v} = \mathbf{f}(\mathbf{u})$? Using a result from the Implicit Function Theorem presented in Appendix B (B.4 p 132), one can construct $\mathbf{f}(\mathbf{u})$ in a certain manner and one then has an expression for (Df) from that appendix. This leads to the construction of a candidate embedding which has the form of a "graph" as presented in example 6.3 above. Since a graph is always an embedding, this leads to the proof of part (1) above. There are many details.

The proof of (2) is a little easier and involves the vector chain rule. The proof begins in this manner. We first write $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$, but then we use $\mathbf{x} = \boldsymbol{\psi}(\mathbf{u})$ to make connection to the embedding **ψ** which we have already shown is in fact a manifold. Then we have $\mathbf{c} = \boldsymbol{\varphi}(\boldsymbol{\psi}(\mathbf{u})) \equiv \boldsymbol{\varphi}'(\mathbf{u})$, a new functional form. Clearly then we can write $(\mathbf{D}\boldsymbol{\varphi}') = \mathbf{Dc} = 0$ and then the chain rule makes this say $(\mathbf{D}\boldsymbol{\varphi})(\mathbf{x})\ (\mathbf{D}\boldsymbol{\psi})(\mathbf{u}) = \mathbf{0}$ and then some fiddling around lets you conclude that in fact $(\mathbf{D}\boldsymbol{\varphi})\mathbf{v} = 0$ for any tangent vector **v**. Thus we end up with the tangent space being the nullspace (or kernel) of the matrix (Dφ) at some point **x**.

Sjamaar then presents a series of **examples** of this regular value theorem. Examples 6.12, 6.13 and 6.14 involve only a single function φ, so for these examples m = 1, See raw notes for discussion of these examples. The examples are $\varphi(x,y) = xy$ and then $\varphi(x,y) = x^3+y^3-3xy$ which is similar to a Buck book example page 356. Then 6.14 deals with $\varphi = \mathbf{x}\bullet\mathbf{x}$ which of course is a hypersphere manifold.

Example 6.15 has two φ functions and **x** has 4 components.

The last Example 6.16 is a bit odd, since it involves a manifold of rotation matrices A rather than a manifold of points in space. It was good for Sja to add that because we often hear of manifolds in that group context, but his presentation is not very satisfying at least to me. He was just trying to wedge this in without lots of details. He also has some errors here which I will report.

As of 2.14.16 I am fairly happy with this section 6.2 and its completely alternate way to characterize a manifold. So at this point, the reader at least has some idea of what is a Manifold!