Sjamaar Chapter 6 notes PhL 7.17.15

6.1 The definition of a manifold 1
6.2 The regular value theorem 14

6.1 The definition of a manifold

Questions. Look at the picture on page 68 middle.
Why are the two vectors shown tangent to the surface?
What is the normal to the surface at the point on the surface.

I am used to such a surface appearing as g(x,y,t) = 0, then I can say things about it. But this is a *parametrized* surface which has

$x_1 = \psi_1(t_1, t_2)$
$x_2 = \psi_2(t_1, t_2)$ or $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ // was $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ in Ch 3
$x_3 = \psi_3(t_1, t_2)$. // was $\mathbf{x} = \mathbf{c}(t)$ in Ch 4, and $\mathbf{x} = \mathbf{c}(\mathbf{t})$ in Ch 5
...
$x_N = \psi_N(t_1, t_2)$ // is $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ here in Ch 6

Well, consider a neighborhood of our point of interest. Then

$dx_1 = (\partial\psi_1/\partial t_1)\, dt_1 + (\partial\psi_1/\partial t_2) dt_2$ etc $\quad d\mathbf{x} = (D\psi)(\mathbf{t})\, d\mathbf{t}$ // matrix times vector
$dx_i = (D\psi)_{ij} dt_j = \partial_j \psi_i\, dt_j = (\partial\psi_i/\partial t_j) dt_j$

If you move just a very small d**t**, since the surface is planar locally, d**x** lies on the surface. So let

$d\mathbf{t} = dt_i\, \mathbf{e}_i$ where $\mathbf{e}_i$ is a unit vector in t-space. // no implied sum

Then we get

$d\mathbf{x}_i = (D\psi)(\mathbf{t})\, [dt_i\, \mathbf{e}_i]$.

These vectors $d\mathbf{x}_i$ for i = 1,2 are therefore tangent to the surface at t or **x**. One could define

$\boldsymbol{\xi}_i \equiv (D\psi)(\mathbf{t})\mathbf{e}_i$ = some un-normalized tangent vectors $\quad i = 1,2...n$.

This then agrees with the arrows shown on the right side of page 68B.

Question: *Why* are the dx_i tangent to the surface? If you move a tiny amount in t-space, you move a tiny amount along the surface of the manifold in x-space, since after all the mapping is from t-space to that surface. If you move a tiny amount along the surface, you are moving tangent to the surface.

What happens now if we write

$(D\psi) = (\mathbf{c}_1, \mathbf{c}_2 ..... \mathbf{c}_n)$ $\quad (D\psi)_{i1} = (\mathbf{c}_1)_i$ = column 1 $\quad (D\psi)_{ij} = (\mathbf{c}_j)_i$

Then,

$$[\xi_i]_k = [\,(D\psi)\,\mathbf{e}_i]_k = (D\psi)_{kj}\,(\mathbf{e}_i)_j = (D\psi)_{kj}\,\delta_{ij} = (D\psi)_{ki} = (c_i)_k \quad \Rightarrow \quad \xi_i = \mathbf{c}_i$$

Fact: the tangent vectors are the same as the columns of $(D\psi)$. Thus, we really have

$$\mathbf{c}_i = (D\psi)(\mathbf{t})\mathbf{e}_i \qquad \text{where} \quad (D\psi) = (\mathbf{c}_1, \mathbf{c}_2 ..... \mathbf{c}_n)$$

The tangent vectors form a basis locally at point **x** on the target space surface.
This is the famous **tangent space** at point **x**.
Since the basis vectors are the columns $\mathbf{c}_i$, this is also a **column space**.
I would say this was an **affine space** unless you translate your basis vectors $\xi_i = \mathbf{c}_i$ to the origin.
Then you get a subspace. [the $\xi_i = \mathbf{c}_i$ are not generally orthogonal, but still span the tangent space]

*** Note added 8.13.15. (see "the meaning of (Df)...doc"). You *want* the tangent base vectors to be linearly independent so they form a full and useful (non-degenerate) space at the point **x**. According to the above, that means you want the columns $\mathbf{c}_i$ to be linearly independent, since these are the tangent vectors. The $(D\psi)$ matrix in the embedding scenario is in general non-square and taller than it is wide. Saying that the columns are linearly independent is the same as saying the $(D\psi)$ matrix has full rank. As shown in the quoted doc, this means that the matrix $(D\psi)$, although not square, nevertheless has a unique left inverse allowing the equation $\mathbf{u} = (D\psi)\mathbf{v}$ to be solved for **v.** That in turn means that the matrix transformation $\mathbf{u} = (D\psi)\mathbf{v}$ which maps $R^n \rightarrow R^N$ is **one-to-one**. The "requirement" that $(D\psi)$ be one-to-one appears as item (ii) in the definition 6.1 of an embedding on page 67. It was this line of logic that eluded me on my first reading.

I guess you could form the normal in our example as

$$\mathbf{n} = \mathbf{c}_1 \times \mathbf{c}_2 .$$

Question: How do I interpret the expression on the RHS of page 67 B? (Need a decoder to read Sja)

$[(D\psi)(\mathbf{t})]$ = a matrix which acts on vectors in t-space which is some U within R^n.

So I certainly could write (recall that U lies in R^n)

ψ:	$U \rightarrow R^N$	$\boldsymbol{\psi}(\mathbf{t}) = \mathbf{x}$	our transformation
$(D\psi)$:	$U \rightarrow R^N$	$[(D\psi)(\mathbf{t})]\,\mathbf{t}' = \mathbf{x}'$	a different mapping

In the first mapping, as **t** runs over all of U, the points **x** run over all of some surface of dimension n inside the space R^N. For example, on page 68 U lies in R^2 and the toroidal surface is a 2D surface and we happen to display in inside $R^N = R^3$ because that is all we know how to draw.

In the second mapping, for the n special points $\mathbf{t'} = \mathbf{e}_i$ in U the mapping (Dψ) produces a set of n tangent vectors at the location **x** (associated with **t** in U) on the nD surface in R^N. Recall that these tangent vectors are the columns $\mathbf{c}_i$ of the matrix (Dψ). You can write a general **t'** in U as $\mathbf{t'} = \Sigma_i t'_i \mathbf{e}_i$ and this would map into a general point **x** lying in the tangent space over in R^N .

$$(D\psi)\,\mathbf{t'} = (D\psi)(\,\Sigma_i t'_i \mathbf{e}_i\,) = \Sigma_i\, t'_i\,(D\psi)\mathbf{e}_i = \Sigma_i\, t'_i\,\mathbf{c}_i = \Sigma_i\, t'_i\,\xi_i \quad = \text{point in tangent space}$$

In this discussion, I have **t** being the reference point, and **t'** being the vector mapped by [(Dψ)(**t**)] and I would write that mapping as [(Dψ)(**t**)](**t'**) .

Alternate Notation: Let's pick some $\mathbf{t}_0$ and corresponding $\mathbf{x}_0$ as reference. There exists a tangent space for $(\mathbf{t}_0, \mathbf{x}_0)$ and that space has tangent vectors $\mathbf{c}_i^{(0)} = (D\psi)(\mathbf{t}_0)\mathbf{e}_i$. Everything is specific to $\mathbf{t}_0$. Now given this sort of fixed reference point, we can consider $\mathbf{t} = \Sigma_i t_i \mathbf{e}_i$ and watch it map into (Dψ) $\mathbf{t} = \Sigma_i\, t_i\, \xi_i$. Thus, any point in U maps into some point in the tangent space associated with $(\mathbf{t}_0, \mathbf{x}_0)$. The point $\mathbf{t} = \mathbf{0}$ then maps into the point $(D\psi)\mathbf{0} = \mathbf{0}$ in the tangent space, so: "the origin of t-space which is U maps into the origin of the tangent space $T_{\mathbf{x}}M$ " . So there are then two different values of the variable **t** to think about here. [correct]. In more detail you might write

$$[(D\psi)(\mathbf{t}_0)]\,\mathbf{t} = \Sigma_i\, t_i\, \xi_i \qquad \text{for any } \mathbf{t} = \Sigma_i t_i \mathbf{e}_i$$

This is true because we showed above that $[(D\psi)(\mathbf{t}_0)]\,\mathbf{e}_i = \xi_i = \mathbf{c}_i$ = basis of $T_{\mathbf{x}0}M$. [correct]

If we now let **t** exhaust all points in U within the t-space R^n, we then would exhaust some portion of the infinite tangent space based on $\mathbf{t}_0$ and $\mathbf{x}_0$. So then

$$[(D\psi)(\mathbf{t}_0)]\ U = \text{a piece of } T_{\mathbf{x}0}M. \qquad \text{// not a piece of M, but a "plane" tangent to M}$$

Then we could say:

$$[(D\psi)(\mathbf{t}_0)] : U \rightarrow T_{\mathbf{x}0}M \qquad \text{where range is part of } T_{\mathbf{x}0}M$$

But since U lies in R^n you *could* regard this as saying

$$[(D\psi)(\mathbf{t}_0)] : R^n \rightarrow T_{\mathbf{x}0}M$$

In fact, if you were to extend U to be all of the t-space R^n, then you would exhaust the tangent space. Of course then you might not have a real manifold due to self-intersection, but ignore that. In this situation, the above line could be written

$$[(D\psi)(\mathbf{t}_0)]\,(R^n) = T_{\mathbf{x}0}M$$

Both sides of this equation are sets which include all points of the tangent space associated with $(\mathbf{t}_0, \mathbf{x}_0)$. Notice that our test variable **t** does not appear in this equation! Suppose we change the test point in U to be called **t'** and we let (**t**, **x**) be the reference point. Then we would rewrite the above equation as

$[(D\psi)(\mathbf{t})]\ (R^{\mathbf{n}}) = T_{\mathbf{x}}M$ // as **t'** exhausts $R^{\mathbf{n}}$, we exhaust the tangent space. (*)

THIS must be what Sjamaar means by page 67 B!! You could also write this as

$[(D\psi)(\mathbf{t})] : R^{\mathbf{n}} \rightarrow T_{\mathbf{x}}M$

where this would be a 1-to-1 mapping which exhaust $R^{\mathbf{n}}$ on the left and exhausts $T_{\mathbf{x}}M$ on the right.

Now I have jumped ahead a bit. The surface M is not yet a manifold but we shall still call it M. We do know that we generate this surface by doing M = ψ(U). Thus we could rewrite (*) above as

$[(D\psi)(\mathbf{t})]\ (R^{\mathbf{n}}) = T_{\mathbf{x}}\psi(U)$ // as **t'** exhausts $R^{\mathbf{n}}$, we exhaust the tangent space. (*)'

This then matches page 67 B.

This tangent space and its basis vectors are different as you change $\mathbf{t_0}$ and therefore you change **x**. F

Note added 8.13.15. This notation $(T_{\mathbf{x}}\psi) \equiv (D\psi)$ simply stresses that the matrix $(D\psi)$ when applied to the basis vectors in U (t-space) generates "tangent vectors" in the range space at point **x**. So the letter T is for Tangent (space or vectors), and the letter **x** emphasizes that the tangent vectors generated by $(D\psi)$ are only really tangent vectors at the point **x** on the range space surface. So $(T_{\mathbf{x}}\psi) \equiv (D\psi)$ is just a notation.

$R^{\mathbf{N}}$ is called the **target space**, it holds your nD surface. [page 68 has n = 2 and N = 3]

Finally we come to Sjamaar's definition of an embedding on page 67, in terms of the picture page 68:

Embedding Definition of Sjamaar

6.1. DEFINITION. Let U be an open subset of $\mathbf{R}^n$. An *embedding* of U into $\mathbf{R}^N$ is a C^∞ map $\psi: U \to \mathbf{R}^N$ satisfying the following conditions:

(i) ψ is one-to-one (i.e. if $\psi(\mathbf{t}_1) = \psi(\mathbf{t}_2)$, then $\mathbf{t}_1 = \mathbf{t}_2$);
(ii) $D\psi(\mathbf{t})$ is one-to-one for all $\mathbf{t} \in U$;
(iii) the inverse of ψ, which is a map $\psi^{-1}: \psi(U) \to U$, is continuous.

The set U is **embedded** into $R^{\mathbf{N}}$ (a larger space) to create an (smooth) **embedding** of U into $R^{\mathbf{N}}$. This smooth embedding scenario has three requirements.

(i) The embedded *surface* is defined by $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$.The mapping $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ must be C^∞ so the mapping is **smooth.** If you want this surface to be non-self-intersecting, then you must insist that the mapping $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ [which is $\psi: R^{\mathbf{n}} \rightarrow R^{\mathbf{N}}$] be one-to-one. If you allow self-intersections [which would be the torus case for a larger U] then you are creating an **immersion**, which is a larger concept than an embedding, as I learned from Darling, see "the meaning...doc". Darling also talks about submersion, but I ignore that extra stuff here.

(ii) The mapping (Dψ) must be one-to-one. As noted above ***, this means that for any point **t** in U, at the point **x** in R^N on your embedded surface you will have a full linearly independent set of tangent vectors. The mapping itself would be written $(D\psi):R^n \rightarrow R^N$ and has the same domain U and range space R^N as the mapping ψ.

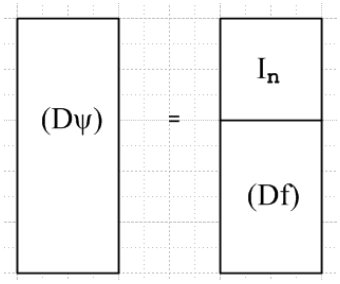
(iii) Since $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ is C^∞, we know that the mapping ψ is continuous in the forward direction, so that a converging sequence of points in U would map into a converging sequence of points on the embedded surface in R^N. In item (iii) we require that the inverse mapping ψ^{-1} must *also* be continuous (though not C^∞). As Sja points out, if you were to consider a sequence of points on the embedded surface which approach some limit point on that surface, the reverse-mapped sequence of $\mathbf{t}_i$ values in U would approach some limiting **t** value in U. This condition just says things are smooth.

So with these three parts of the definition, an embedded surface has "nice" properties: the mapping ψ is continuous in both directions, the tangent vectors at any x are non-degenerate and span a full subspace R^n within R^N, and the surface is not-self intersecting. To make this all work, (Dψ) must have full rank, meaning its column vectors (column space) are linearly independent. This Dψ is my R matrix.

Example 6.2, The domain U is a 2D rectangle, and the embedded surface is a piece of a torus. At a given point on the torus there are two tangent base vectors. Then are correctly shown as $(D\psi)\mathbf{e}_1$ and $(D\psi)\mathbf{e}_2$ which are of course not unit vectors in general.

Example 6.3 Consider a smooth (C^∞) mapping f: U → R^m where U is the same U in R^n discussed above, and now we have a new dimension space R^m in the picture. In terms of this mapping f, we now define the embedding's function **ψ** in this manner: $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t}))$. Therefore we have N = n+m in this example. In general, something like g(x) = (x,f(x)) is called "**the graph of f**".

So for this particular ψ, we can view the matrix (Dψ) in this manner



This picture is the meaning of page 68 E where Sja uses a compressed vector-like notation.

The inverse function (assuming it exists) is defined by

$\mathbf{t} = \boldsymbol{\psi}^{-1}(\mathbf{x}) = \boldsymbol{\psi}^{-1}(\mathbf{t}, \mathbf{f}(\mathbf{t}))$ // since $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t}))$

This appears as page 69 A as the equation $\boldsymbol{\psi}^{-1}(\mathbf{t}, \mathbf{f}(\mathbf{t})) = \mathbf{t}$. Sja is making no claims about how you might write down this inverse function given **f**(**t**), he just says it has this property.

Now is this example really an embedding?

(i) the mapping $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t}))$ is smooth given that $\mathbf{f}(\mathbf{t})$ is smooth. We can see that $\mathbf{t}_1 = \mathbf{t}_2$ implies that $\mathbf{x}_1 = \mathbf{x}_2$. In the reverse direction, suppose $\mathbf{x}_2 = \mathbf{x}_1$. that would require $\mathbf{t}_1 = \mathbf{t}_2$ from the first argument. Therefore the mapping ψ is one-to-one as well as C^∞ so requirement (i) is met

(ii) Regardless of (Df), the picture above shows that the columns of $(D\psi)$ are linearly independent. This is exactly my Mini-Columns Theorem (A.2) in Lagrange doc. Thus, $(D\psi)$ is one to one, so (ii) is met.

(iii) Now we have to show that $\mathbf{t} = \boldsymbol{\psi}^{-1}(\mathbf{x})$ is continuous. Consider a sequence $\mathbf{x}_n \to \mathbf{x}$. Looking at the forward map $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t}))$, we can write $\mathbf{x}_n = (\mathbf{t}_n, \mathbf{f}(\mathbf{t}_n))$. Since the forward map is continuous, we know that $\mathbf{t}_n \to \mathbf{t}$ implies $\mathbf{x}_n \to \mathbf{x}$ for any $\mathbf{t}_n$ sequence. If we have a convergence sequence $\mathbf{x}_n \to \mathbf{x}$, then since $\mathbf{x}_n = (\mathbf{t}_n, \mathbf{f}(\mathbf{t}_n))$, we must also have a convergent sequence $\mathbf{t} \to \mathbf{t}_n$. Thus $\boldsymbol{\psi}^{-1}$ is also continuous. So property iii is met. Therefore this example really is an embedding.

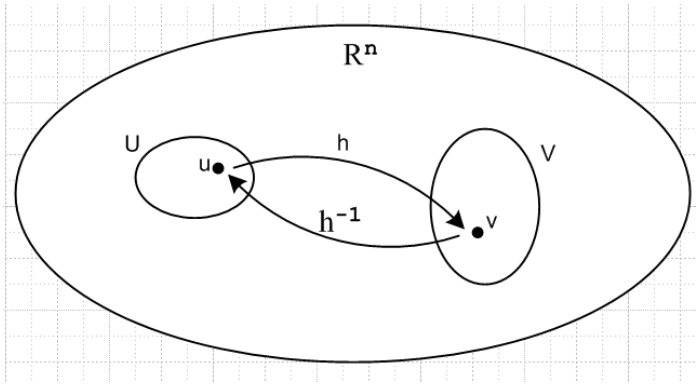
Spivak Digression: His definition of diffeomorphism and two manifold definitions

Before continuing to Sjamaar's definition of a manifold, we digress to another source (just because that is what I did on my first reading attempt). Spivak here describes a certain mapping between two regions within space R^n which mapping I would say was "exceedingly well behaved" and "exceedingly smooth" in every possible sense.

Spivak. He opens his Chapter 5 on manifolds with this definition:

> If U and V are open sets in $\mathbf{R}^n$, a differentiable function $h: U \to V$ with a differentiable inverse $h^{-1}: V \to U$ will be called a **diffeomorphism.** ("Differentiable" henceforth means "C^∞".)

Here is a picture to go with the above definition



Here **u** and **v** are vectors in R^n and so I bold them. Now we really have $\mathbf{v} = \mathbf{h}(\mathbf{u})$ so differentiable means in the mult-variable sense, so one is claiming I think that (**Dh**) exists at any point in U and V, and I guess so does (D^2h) and so on, whatever that means (well, $\partial_j \partial_i h$ exists, etc). The fact that an inverse exists

says that this little mapping between U and V must be one-to-one. So at least Spivak defines the word **diffeomorphism** for his poor sucker reader. OK, more words: when you consider $\mathbf{v} = \mathbf{h}(\mathbf{u})$, you can compute $(\mathbf{Dh})(\mathbf{u})$ and all higher vector derivatives (like $\partial_i\partial_j h_k$) and all these derivatives are well defined and smooth. Moreover, the function h^{-1} exists AND you have $\mathbf{u} = \mathbf{h}^{-1}(\mathbf{v})$ and in this direction $(\mathbf{Dh}^{-1})(\mathbf{v})$ exists and is computable and is smooth. The point is that you can differentiate C^∞ in either direction on any variable. So I guess you would say this was a *very smooth mapping*, and since h^{-1} exists, it is 1-to-1.

So in Sja's language we have $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ replacing $\mathbf{v} = \mathbf{h}(\mathbf{u})$ and we have not only ψ being one to one and having an inverse, but $(D\psi)$ exists and $(D\psi^{-1})$ exists. In the embedding definition, we only required that ψ^{-1} be continuous, but in this diffeomorphism in fact ψ^{-1} is C^∞, overkill.

Note that in Spivak's diffeomorphism, both spaces are on an equal footing. So you could refer to either h or h^{-1} as a diffeomorphism.

Comment added 8.12.15. In regular 1D calculus, suppose you have y = f(x). We know (from Buck p 270) that the inverse is obtained by reflecting in the x=y plane and can result in several inverse branch functions as on page 270 or 271. If f(x) is one-to-one, that would be like the right side only of the parabola on page 270 $y = f(x) = x^2$. For just that piece y=f(x) and $x = f^{-1}(y) \equiv g(y)$ would exist, f(x) is invertible, the function is 1-to-1 and all is well. The function $y = F(x) \equiv f'(x) = 2x$ is also 1-to-1. We might refer to this as F(x) = (Df)(x). So for a general smooth situation, one would have both f(x) and (Df)(x) being 1-to-1.

The question is how you extend this smoothness concept to nD calculus, and that is what Spivak is discussing. You then want $\mathbf{y} = \mathbf{f}(\mathbf{x})$ to be 1-to-1 so that $\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y})$ exists. And you also want that $\partial y_i/\partial x_j \equiv \partial_j f_i = (\mathbf{Df})_{ij}(\mathbf{x})$ be one-to-one. In the 1D case, there is only one derivative, so both f(x) and f'(x) are scalars. In the nD case, $\mathbf{f}(\mathbf{x})$ is a vector and $(\mathbf{Df})_{ij}$ is a rank-2 tensor, so things are much more complicated. You can differentiate each vector component in each of many directions. We then want to have "balls" around the points **x** and **y** so we can examine derivatives in all directions.

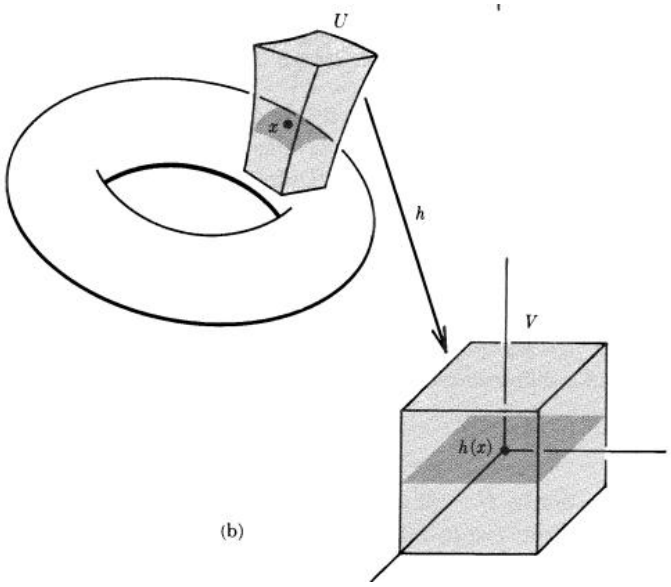
This is what Spivak is doing. The diffeomorphism defined above, the vector function **h** is assumed to have an inverse, so $\mathbf{v} = \mathbf{h}(\mathbf{u})$ is 1-to-1. And it is required that both $\mathbf{v} = \mathbf{h}(\mathbf{u})$ *and* $\mathbf{u} = \mathbf{h}^{-1}(\mathbf{v})$ be differentiable. But it is not required $(\mathbf{Dh})(\mathbf{u})$ be 1-to-1 . Differentiable means all derivatives C^∞.

Comment: A **homeomorphism** is like a diffeomorphism but has a lesser requirement: h and h^{-1} have to exist and be just *continuous* in both directions, not necessarily differentiable. So the Sjamaar embedding conditions require something halfway between a diffeo and a homeo morphism.

Another spoonful of Spivak caster oil which is his **manifold definition**:

> A subset M of $\mathbf{R}^n$ is called a ***k*-dimensional manifold** (in $\mathbf{R}^n$) if for every point $x \in M$ the following condition is satisfied:
>
> (M) There is an open set U containing x, an open set $V \subset \mathbf{R}^n$, and a diffeomorphism $h: U \to V$ such that
>
> $$h(U \cap M) = V \cap (\mathbf{R}^k \times \{0\}) = \{y \in V: y^{k+1} = \cdots = y^n = 0\}.$$

Here is Spivak's corresponding picture where the manifold M is the toroid donut surface.



For his picture, the manifold M is a 2D surface, so k = 2. The dimension n is called N by Sjamaar, it is the dimension of the space in which you are doing the embedding, and in the above picture as well as Sjamaar page 68 this dimension is 3. The dimension k is called n by Sjamaar. In Spivak's picture, both the domain and the range of the mapping are in the larger space of dimension n (Sja N). The dark grey square on the right would be Sja's rectangle on the left of p 68 B. So, the idea is that you can talk about a patch on the manifold as being U∩M where U provides "elbow room in all directions" for any point on M. Then of course h(U∩M) is the dark rectangle on the right. Of course V = h(U) is the light graph box on the right, and this provides "elbow room in all directions" for any point on the right in the dark square.

Now, on the right, the gray rectangle of course has z = 0 (it lies on this plane on the right). For N > 3, all such coordinates would be zero, and the mapping is just that gray rectangle or its generalization to k dimensions. The idea that z = 0 is what Spivak means when he writes R^k x {0). The mapped grey square is the R^k part, and all the z = 0 coordinates are the z = 0 part. Now the grey plane on the right can be large, but we only care about that part of it which corresponds to (is mapped from) the patch of interest on M, and that is then $V \cap [R^k \times \{0\}]$.

So now all of Spivak's notation makes sense, that was the hard part. What he is saying is this: if you have a candidate manifold surface M, to test whether it really is a manifold, you pick a point on M (Sjamaar's x) and you select your little U and V balls (they can be as small as necessary) and you check to see whether the mapping h (which is Sja ψ^{-1}) is a diffeomorphism. So Spivak is requiring that for any tiny patch on M, the mapping ψ must be exceedingly smooth and well behaved, even C^∞ in both directions!

Now that we have seen Spivak's manifold definition, let's look at Sjamaar's on page 60. There are several differences:

(1) Whereas Spivak's ball around the toroid surface point is called U, Sjamaar calls it V, so there is an immediate U↔V change going on. So Spivak had U in R^n around the toroid surface point, while Sjamaar has V in R^N around that same point.

(2) Spivak puts a ball V around the parameter space rectangle where the higher coordinates are all zero. Sjamaar does not bother with those extra dimensions and he calls this thing U. So

Sjamaar		Spivak		
V	dim N	U	dim n	ball around the toroid point in R^N
U	dim n	$V \cap [R^k \times \{0\})]$	dim k	the parameter space
$V \cap M$	dim n	$U \cap M$	dim k	the toroid patch
ψ		h^{-1}		mapping from param space to toroid

(3) Spivak requires that his mapping h^{-1} be a full blown diffeomorphism. Sjamaar only requires that this mapping be an embedding, which is not as supersmooth as the diffeomorphism in the inverse direction. Spivak has the inverse mapping being C^∞, whereas Sjamaar has it being C^1 only.

Spivak goes on to give an *alternate* manifold definition which looks closer to the Sja one:

There is an alternative characterization of manifolds which is very important.

5-2 Theorem. *A subset M of $\mathbf{R}^n$ is a k-dimensional manifold if and only if for each point $x \in M$ the following "coordinate condition" is satisfied:*

(C) There is an open set U containing x, an open set $W \subset \mathbf{R}^k$, and a 1-1 differentiable function $f: W \to \mathbf{R}^n$ such that

(1) $f(W) = M \cap U$,
(2) $f'(y)$ has rank k for each $y \in W$,
(3) $f^{-1}: f(W) \to W$ is continuous.

[Such a function f is called a **coordinate system** around x (see Figure 5-2).]

Here is a translation table for this alternative definition

Sjamaar		Spivak		
V	dim N	U	dim n	ball around the toroid point in R^N
U	dim n	W	dim k	the parameter space
$V \cap M$	dim n	$U \cap M$	dim k	the toroid patch
ψ		f		mapping from param space to toroid

Here is the Sjamaar manifold definition,

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 $\psi(U) = V \cap M$.

Sjamaar's last bullet says that ψ(U) = V∩M and that ψ is an embedding. In Spivak, since the word embedding is not used, Spivak has to write out the embedding requirements. Spivak (1) is part of the last bullet of Sjamaar. Spivak (2) uses notation f'(y) to represent Sjamaar's (Dψ) matrix. Spivak requires f'(y) to have full rank k, which is like Sjamaar having (Dψ) be full rank n, so it is one-to-one. Spivak (3) is the embedding requirement that the inverse function as well as the forward function be continuous.

The notion of a chart and atlas (Sjamaar page 67, so I have gone backwards)

The inverse map ψ^{-1} for a patch of some size is called a **chart** or **coordinate map**. This is the map from the fancy torus patch to its parameter space rectangle. This is like the mapping ψ^{-1} from a piece of the surface of the Earth to a chart (map) of that piece of the earth.

PL Example 1 of a manifold, charts and an atlas.

The manifold here is a circle, and in this example wiki considers four different patches which in fact overlap (patch = piece of circle), and there are then four charts which make up one atlas of the circle. The dimensions here are: manifold is a 1D surface (n = 1), parameter space is 1D (n = 1) which they call x for two of the charts and y for the other two charts. The circle is embedded in R^2 so N = 2.

Motivational examples [edit]

Circle [edit]

Main article: Circle

After a line, the circle is the simplest example of a topological manifold. Topology ignores bending, so a small piece of a circle is treated exactly the same as a small piece of a line. Consider, for instance, the top part of the unit circle, $x^2 + y^2 = 1$, where the y-coordinate is positive (indicated by the yellow circular arc in *Figure 1*). Any point of this arc can be uniquely described by its x-coordinate. So, projection onto the first coordinate is a continuous, and invertible, mapping from the upper arc to the open interval $(-1,1)$:

$$\chi_{\text{top}}(x, y) = x.$$

Such functions along with the open regions they map are called *charts*. Similarly, there are charts for the bottom (red), left (blue), and right (green) parts of the circle:

$$\chi_{\text{bottom}}(x, y) = x$$
$$\chi_{\text{left}}(x, y) = y$$
$$\chi_{\text{right}}(x, y) = y.$$

Together, these parts cover the whole circle and the four charts form an atlas for the circle.

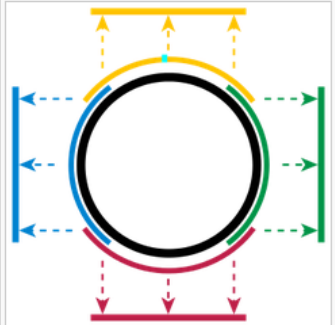
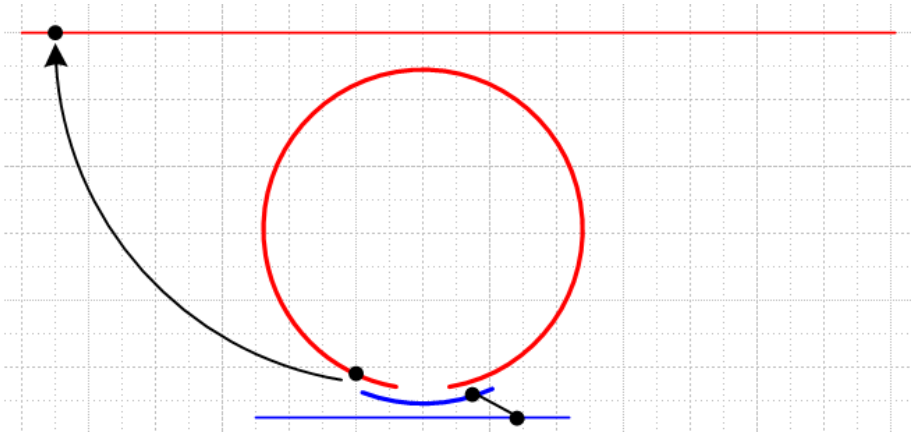


Figure 1: The four charts each map part of the circle to an open interval, and together cover the whole circle.

In this example, one has 4 charts to cover the circle, so the atlas has four charts. The wiki says that almost always you need more than one chart to cover a manifold. It also says that you could do the circle with two charts like this:



where the line segment at the top is the E^1 space to which we are mapping the red part of the circle. The mapping has to be one to one! If you try to cover the circle with just one chart by extending the red piece and shrinking away the blue piece, you end up with a mapping from circle to line segment which is NOT one to one. The two ends of the line segment map to the same (bottom) point on the circle. So this same idea would apply to any closed curve including the one Sja shows page 69D. So it the problem has noting to do with tangent spaces or places where the curve has the same slope.

To summarize: a simple closed curve is a manifold, but you cannot cover it with one chart. You need two or more charts.

This of course all relates to mapping the globe that is the earth. You try to do it this way



But now the left and right edge map to the same great semicircle on the globe, so the mapping is not one to one! The right thing to do is have a whole book of charts, where each chart covers some local section of the earth, and that book of local maps is indeed called an atlas.

Review of things prior to page 70

1. Cover the entire M with patches small enough so the above works for each patch. It is OK for the patches to overlap, and probably desirable that they do so. Then

$M = P_1 \cup P_2 \cup ... \cup P_n$ P = Patch (= chart)

$\psi_1: U_1 \to R^N$ is an embedding making use of V_1
$\psi_2: U_2 \to R^N$ is an embedding making use of V_2
.....

If you can do this, then M is a **manifold**.

2. Fact: if the surface M self-intersects even just at one point **x**, then for any patch containing that point x, you cannot create an embedding because $\psi(t)$ won't be one-to-one since two points on the surface map back to same t. Thus, this one point excludes surface M from being a manifold.

3. The manifold has dimension n within R^N, so the **codimension** of M is N-n. For the torus example, that codimension is 1.

4. For a given one of the patch embeddings $\psi^{(i)}$ that includes **x** on M, Sja now wants to say that the notation for the tangent space associated with that **x** is written this way:

$T_x M = [(D\psi^{(i)})(\mathbf{t})]\, U^{(i)} \approx$ the tangent space at **x** $\approx R^n$

but Sja still uses that strange notation I don't understand, see p 69 C. [but now I do understand it]

5. He says now that the collection of all the embeddings $\{\psi^{(i)}\}$ which cover M is what you should call an **atlas**. But I like the idea of thinking of the $U^{(i)}$ as being the **charts** for each embedding, and the atlas is the set of all those charts. A chart is the inverse map from the M patch to the U domain, and its image is that U domain. So OK, semantics here.

He then repeats the Buck idea that the words curve, surface and hypersurface imply that these things are manifolds. To emphasize this fact, you can add the word "**smooth**" such as "smooth surface".

Sja is really hammering hard on this tangent space idea at every opportunity, but we don't yet know why he is doing that. He is setting us up for something!

The rest of this Section 6.1 is treated one page at a time:

Page 70

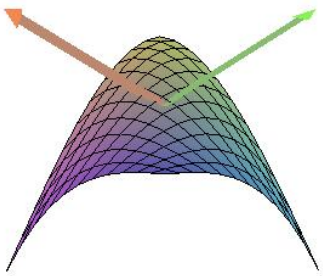
Example 6.5. What happens if $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = \mathbf{t}$? Here ψ is the identity map. Then $\boldsymbol{\psi}: U \to U$ and surely this is a valid embedding, and only one embedding is needed to cover all of U (all of U is "the patch") and then this U within R^n is a manifold. $(D\psi) = I_n$ so the tangent space is spanned by the same $\mathbf{e}_i$ that span U, so tangent space = R^n.

Conclusion: Any open region U in R^n is in fact a manifold of dimension n and codimension 0.

Example 6.6. What happens if, for $\psi:R^n \to R^N$ we have $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t}) = (t_1,t_2, \ldots t_n, 0,0,\ldots 0)$? Is this an embedding? This seems a slight generalization of 6.5 above. The higher dimensions of R^N do nothing, and we basically have $\boldsymbol{\psi}$: U → U happening within the context of this larger R^N. So yes, this seems to be an embedding. This situation is "an n-manifold within R^N". This reminds me of the original Spivak manifold definition stated above.

Example 6.7. Here we are back to the (**t**,f(**t**)) **graph** thing, now written (**x**,**f**(**x**)). The second example would then have ψ = (x,y,f(x,y)) as a column vector. He computes the (Dψ) matrix and I think he has the wrong results displayed. [agreed on 8/13/15] The graphs on the next page just show two views of the same 3D graph, and he is displaying tangent vectors at a few points. I wonder if these are wrong too?

How do you display a simple arrow in Maple? I learned how and here is my version of his graph



where I just picked one point on the surface to show tangent vectors. I used my formulas, and my arrows are in fact tangent.

Page 71

Example 6.8. Here $\psi(t) = e^t(\cos t,\sin t)$ so we have a spiral thing for $\psi:R \to R^2$. As $t \to -\infty$, it spirals around the origin an infinite number of times which seems reasonable to me. It never gets there. He shows that this $\psi(t)$ does define an embedding (the surface is the spiral curve in E^2). He claims this curve is a manifold as long as you stay away from the origin by using large but finite negative t endpoint.

Page 72

[72] Example 6.9. I already saw that a circle needs two charts or more. Here the unit sphere S^{n-1} is considered, where I think n-1 is the dimensionality of the sphere's surface, so a circle is S^1 and a sphere is S^2. For the general case S^{n-1} we have **t** in U in R^{n-1} . He writes a mapping into R^n (where sphere lives),

$$\psi(\mathbf{t}) = \frac{1}{\|\mathbf{t}\|^2 + 1}\left(2\mathbf{t} + (\|\mathbf{t}\|^2 - 1)\mathbf{e}_n\right)$$

This function somehow has an image which is the sphere surface, details in Exercise B.7 which I have not done.

His mapping is the one used for a stereographic projection of the earth. He does a north pole one and then a south pole one, and so you need two charts to cover the sphere, in any number of dimensions. Because you can cover the sphere in these two embeddings, it is a manifold by our piecing-together definition.

And so finally I have finished this very painful section 6.1

6.2 The regular value theorem

Warning:

In the previous section 6.1 we dealt with $\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$ and matrix $(D\psi)$ had columns $\mathbf{c}_i$.
In Section 3.2 we had $\mathbf{y} = \boldsymbol{\varphi}(\mathbf{x})$ in our initial pullback discussion with pullback operator φ^*
and in the pullback structure we encountered the matrix object $(D\varphi)$
In Section 4.1 we had c* being this pullback operator for 1-forms only.
So we have had the symbols φ and c floating around in various ways.

Now below we have a new $\mathbf{c}$ and a new $\boldsymbol{\varphi}(\mathbf{x})$ which we take just *a-priori*, and there is of course a $(D\varphi)$ matrix going with this φ.

[72] Suppose $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ is m equations in N unknowns. The points $\mathbf{x}$ here lie in U within $R^{\mathbf{N}}$.

Warning: In Section 6.1 we had $x = \psi(t)$ and $\psi : R^{\mathbf{n}} \to R^{\mathbf{N}}$ with U in $R^{\mathbf{n}}$. Here we have $c = \varphi(x)$ with φ: $R^{\mathbf{N}} \to R^{\mathbf{m}}$ with U in $R^{\mathbf{N}}$, so beware that the letters are all shuffled around.

We could compute the following differentials:

$$dc_i = (D\varphi)_{ij}(\mathbf{x})dx_j = \partial_j\varphi_i\, dx_j = (\partial\varphi_i/\partial x_j)dx_j \qquad (D\varphi)_{ij} \text{ has i = m rows and j = N columns.}$$

Sjamaar does not say anything yet about the relative sizes of the integers N and n. To be clear, the vectors $\boldsymbol{\varphi}$ and $\mathbf{c}$ have m components, while the vector $\mathbf{x}$ has N components. ok to here

Suppose that, for a given fixed vector **c,** the equation $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ has a solution locus for $\mathbf{x}$ such that for all $\mathbf{x}$ in this locus, $(D\varphi)$ has full row rank m. I know that this means $(D\varphi)$ also has column rank m and as a whole this non-square matrix has rank m. This matter all cleared up in Lagrange Appendix A. This would then suggest that $N \geq m$ in Section 6.2, On page 73 top we see mention of N-m as a space dimension, so again Sja is intending that $N \geq m$. This is the scenario of a wide R matrix, as in Lagrange doc.

Note added 2.14.16. You can of course compute $[(D\varphi)(\mathbf{x})]$ ignoring $\mathbf{c}$ and you get some matrix. You can evaluate this matrix at any $\mathbf{x}$ in $R^{\mathbf{N}}$ that you want. However, $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ defines some region within $R^{\mathbf{N}}$ which is our candidate manifold M. The idea is that this region will be a manifold if $[(D\varphi)(\mathbf{x})]$ is full rank for all points $\mathbf{x}$ *on the surface M*.

Now Sja's notation for "the solution locus of x of $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ " is the slightly strange $\varphi^{-1}(\mathbf{x})$. So if for all **x** in $\varphi^{-1}(\mathbf{x})$ the matrix $(D\varphi)(\mathbf{x})$ has full rank (which is rank m), then **c** is called a "**regular value** of φ". And if there is somewhere in that locus where that rank falls below m, then **c** is a "**singular value**".

Recall that Buck refer to a point where f(**x**) has all partials 0 is called a "**critical point**", so that is a slightly different idea. However, Bucks also use the term "critical point" (p293) for a place where a transformation matrix R drops below full rank. so the Buck's "critical point" seems to be the same as Sjamaar's "singular value" concept.

Aside: Go back to the solution set $\boldsymbol{\varphi}^{-1}(\mathbf{x})$ for a given **c**. For every **c** in its space $R^{\mathbf{m}}$ we are generating some whole solution surface for **x** in its space $R^{\mathbf{N}}$ for each value of **c**. You can think of these spaces as forming a set of **level surfaces** within $R^{\mathbf{N}}$ (such as a set of concentric spheres). Sja and others like to call these **level sets**. The idea that you have a whole surface for each value of **c** is related to the idea of a **fiber bundle**. Think of a hairbrush with handle a cylinder and radial tongs. For each point on the cylinder (for each value of c in one space) you have a whole other space (here a linear "fiber" which is the brush tong). So in this example you have a "bundle of fibers" making up the brush.

Comment: This seems similar to the idea of a Tangent Space $T_{\mathbf{x}}\psi$ in Section 6.1 where for each value of **x**, you have a "whole space" which is the tangent space at that point. So maybe the tangent space concept is also a fiber bundle situation? But in this case, we don't have a surface for each x, we have a space with axes and so on, so maybe this is not a fiber situation.

So here are the claims of the theorem 6.10.

[73] Claims of Theorem 6.10. Consider $\boldsymbol{\varphi}$: U → $R^{\mathbf{m}}$ as appropriate for the $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$. This is then $\boldsymbol{\varphi}$: $R^{\mathbf{N}}$ → $R^{\mathbf{m}}$ with U an open set in $R^{\mathbf{N}}$. In Section 6.1, we had U having the smaller dimension, but now it has the larger dimension (so the R matrix is wide as in Lagrange doc, not tall). Think of the candidate manifold as $M = \boldsymbol{\varphi}^{-1}(\mathbf{x})$ = the solution set $\{\mathbf{x}\}$ for $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$, as defined above. Clearly this M is something embedded in $R^{\mathbf{N}}$ since the **x** are in $R^{\mathbf{N}}$. If **c** is a regular point, *then* the claims of this theorem are:

(1) M is a manifold in $R^{\mathbf{N}}$ of dimension N-m and thus codimension m.
(2) Furthermore, the tangent space at point **x** is the nullspace of $(D\varphi)$.

In English, this says: for surfaces implicitly defined by a function, there is a simple test to see if that surface is really a manifold or not. Write the function in the form $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ and check to see if **c** is a regular point by seeing if the mxN R matrix $(D\varphi)$ has full rank m, which is to say, there are m linearly independent rows. If yes, then the surface is a manifold and the tangent space at some **x** is the nullspace of the matrix $(D\varphi)$. The manifold has dimensionality n and thus the tangent space at any point on the manifold has dimension n. The codimension is N-n = m.

Comment: Since we have m equations in N unknowns $\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$ defining our manifold, we certainly need to have $m \leq N$ in order to be defining a solution space in the first place, so again, $N \geq m$.

We want to show that M = $\boldsymbol{\varphi^{-1}(x)}$ is a manifold. From page 69, our first task is to put a ball V around a point **x** on the manifold M. Then we put another ball U in our parameter space. We need to show that the mapping ψ: U → $R^{\mathbf{M}}$ is an embedding and that $\psi(U) = V \cap M$.

The complicated proof of this Theorem has something like 24 little steps (each got a red check).
I will outline the first few steps:

Proof of Claim 1

1. Think of **x** = (**u**,**v**) where dim(x) = N, dim(u) = n, dim(v) = m, and so N = n+m. We are in effect saying to think of $R^{\mathbf{N}}$ as being $R^{\mathbf{n}}$x $R^{\mathbf{m}}$.

2. Our reference point of interest on the candidate manifold is $\mathbf{x_0} = (\mathbf{u_0},\mathbf{v_0})$

3. The matrix (Dφ) which is "wide" certainly has a non-zero minor (mxm) , since the m rows are linearly independent. I showed this in Lagrange doc App B. This minor is an invertible m x m matrix. This is a square piece of the R matrix (the Dφ matrix).

I proceed down the proof about 1 1/2 inches and we see that Sja is constructing exactly the scenario which is studied in Appendix B.4 and for which I have detailed notes. The idea is that you think of having **x** = (**u**,**v**) where **u** has n components, and you then have **φ**(**u**,**v**) = **c** (here **c** and **φ** and **v** have m components, while **v** has n components). You wonder if there is an "implicit function" **v** = **f**(**u**) which you could write out which solves this equation. If there were an implicit function **v** = **f**(**u**) then you might want to know its R matrix which is (**Df**)(**u**). The Implicit Function Theorem says that this matrix is given by:

$$(\mathbf{Df})(\mathbf{u}) = -(\mathbf{D_v\varphi})^{-1}(\mathbf{u},\mathbf{f(u)})\ (\mathbf{D_u\varphi})(\mathbf{u},\mathbf{f(u)})$$ // which is p 132 B

where ($D_{\mathbf{v}}\varphi$) and ($D_{\mathbf{u}}\varphi$) are partial Jacobians of dimension mxm and mxn (both have m rows). Here the object (**u**,**f(u)**) is simply an argument, exactly the same as (**u**,**v**) = **x**. Since the ($\mathbf{D_v\varphi}$) Jacobian is square, you normally could invert it as shown above. The idea is that if this thing is invertible at some ($\mathbf{u_0},\mathbf{v_0}$), then it is going to be invertible for (**u**,**v**) nearby. In our proof the ($\mathbf{D_v\varphi}$) matrix is called A, and it is invertible because it has full rank and my Lagrange Theorem 8 (Cramer) says that fact makes A invertible. The Implicit Function Theorem B.4 then says that we indeed then have an implicit function **v** = **f**(**u**) and **φ**(**u**,**f**(**u**)) = **c** . Note that **u** is in U and **v** is in V, so we then have f: U → V within f: $R^{\mathbf{n}} \rightarrow R^{\mathbf{m}}$. [this is a different U] One can regard φ: **x** in $R^{\mathbf{N}}$ → **c** in $R^{\mathbf{m}}$ as φ: $R^{\mathbf{n}}$x$R^{\mathbf{m}} \rightarrow R^{\mathbf{m}}$ or specifically φ: UxV→ $R^{\mathbf{m}}$ (the space of **c**). In our application, we in fact have **x** = (**u**,**v**) and $\mathbf{x_0} = (\mathbf{u_0},\mathbf{v_0})$ being points in $R^{\mathbf{N}}$. As **u** runs over all of U, **v** runs over all of V according to our implicit function **v** = **f**(**u**).

Confusion: In the theorem premise, Sja says U in $R^{\mathbf{N}}$ so **u** should be in $R^{\mathbf{N}}$, but *here* we are saying **u** in $R^{\mathbf{m}}$ which is a subspace of $R^{\mathbf{N}}$. Are these different U's ? I think really it is two different U's ! The second **u** has n coordinates only.

So why do we care about the fact that an implicit function **v** = **f**(**u**) exists in the above context? Hopefully that will be answered in the following barrage of notes.

Now note and just hold this idea that this column vector we associate with "the graph of f":

$(\mathbf{u},\mathbf{v}) = (\mathbf{u},\mathbf{f}(\mathbf{u}))$ is, according to page 68, the "**graph of f**".
n m

Next: what is the meaning of $M \cap (U \times V)$?

The manifold M is all $(\mathbf{u},\mathbf{v})$ pairs of such points $\mathbf{x} = (\mathbf{u},\mathbf{v})$ *where* $\mathbf{v} = \mathbf{f}(\mathbf{u})$ [which means $\boldsymbol{\varphi}(\mathbf{u},\mathbf{v}) = \mathbf{c}$]. The space U x V inside $R^{\mathbf{N}}$ is bigger than M because UxV contains $(\mathbf{u},\mathbf{v})$ pairs where $\mathbf{v} \neq \mathbf{f}(\mathbf{u})$. We know then that $M \subset UxV$ and therefore I think that $M \cap (U \times V) = M$ except maybe for the boundary of M.

Why does Sja state that " $M \cap (U \times V)$ is the graph of f " ?

Well as shown above graph of f is the set of points in $R^{\mathbf{N}}$ of the form $(\mathbf{u},\mathbf{f}(\mathbf{u}))$ and this is basically M, again apart from boundary issues.

Waypoint: At this point in the proof, we have shown that M is the graph of f for this implicit function f which we showed must exist. But earlier we showed that the graph of f is always an embedding, so that is the main thrust of this proof.

Now, Exercise 6.3 showed that the graph of f [which is $\boldsymbol{\psi}(\mathbf{u}) = (\mathbf{u},\mathbf{f}(\mathbf{u}))$] defines a valid embedding $\boldsymbol{\psi}$ of U into $R^{\mathbf{N}}$ and I guess this means that M is a manifold with a single chart since it is the image of this embedding $\mathbf{x} = \boldsymbol{\psi}(\mathbf{u}) = (\mathbf{u},\mathbf{f}(\mathbf{u}))$.

What is the dimensionality of this manifold M? It exists in $R^{\mathbf{N}}$, yes we know that. But when we write the $\mathbf{x}$ values as $\mathbf{x} = (\mathbf{u},\mathbf{f}(\mathbf{u}))$, when we run $\mathbf{u}$ over U, we are running over $R^{\mathbf{n}}$ so I would say that the surface M has dimensionality n, and thus codimension = m as claimed.

Proof of Claim 2

What is the tangent space here?

I agree that one can write $\boldsymbol{\varphi}(\boldsymbol{\psi}(\mathbf{u})) = \mathbf{c}$ since $\boldsymbol{\psi}(\mathbf{u}) = (\mathbf{u},\mathbf{f}(\mathbf{u})) = \mathbf{x}$.

I agree that $\boldsymbol{\psi}(\mathbf{u})$ is our embedding of U in $R^{\mathbf{N}}$.

The tangent space is spanned by the column vectors of $(D\psi)$, bottom p 67.

Comment on relation between φ and ψ: We have these two facts:

$\mathbf{x} = \boldsymbol{\psi}(\mathbf{t})$	in earlier notes	$\mathbf{x} = \boldsymbol{\psi}(\mathbf{u})$	in current context
$\mathbf{c} = \boldsymbol{\varphi}(\mathbf{x})$	in current context		

Combining these we can say,

$$\mathbf{c} = \boldsymbol{\varphi}(\boldsymbol{\psi}(\mathbf{u}))$$

This gives the relation between functions $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$. The relation is surely complicated, but doing a derivative with the chain rule is how we learn more about it.

Fact: $\mathbf{c} = \boldsymbol{\varphi}(\boldsymbol{\psi}(\mathbf{u})) \equiv \boldsymbol{\varphi}'(\mathbf{u})$ [some new functional form] $\Rightarrow$ $\mathbf{Dc} = \mathbf{0} = (\mathbf{D}\varphi')(\mathbf{u})$.

Fact: The chain rule says [note that φ' does NOT mean a derivative, it is just a new function name]

$(\mathbf{D}\boldsymbol{\varphi}')(\mathbf{u}) = (\mathbf{D}\boldsymbol{\varphi})(\boldsymbol{\psi})\ (\mathbf{D}\boldsymbol{\psi})(\mathbf{u})$.

Therefore combining these last two facts:

$(\mathbf{D}\boldsymbol{\varphi})(\boldsymbol{\psi}(\mathbf{u}))\ (\mathbf{D}\boldsymbol{\psi})(\mathbf{u}) = \mathbf{0}$. // the product of two matrices. this is page 73 B

Replace $\boldsymbol{\psi}(\mathbf{u}) = \mathbf{x}$ to get

$(\mathbf{D}\boldsymbol{\varphi})(\mathbf{x})\ (\mathbf{D}\boldsymbol{\psi})(\mathbf{u}) = \mathbf{0}$. (**)

If we evaluate this at $\mathbf{u} = \mathbf{u}_0$, I think you should get

$(\mathbf{D}\boldsymbol{\varphi})(\mathbf{x}_0)\ (\mathbf{D}\boldsymbol{\psi})(\mathbf{u}_0) = \mathbf{0}$ close to p 73 C

As I showed earlier, the tangent vectors are the columns of $(\mathbf{D}\boldsymbol{\psi})(\mathbf{u}_0)$ at point $\mathbf{x}_0$ on M.
I had in fact that $\boldsymbol{\xi}_i \equiv (D\psi)(\mathbf{u})\mathbf{e}_i$ = column i.

So I agree one can write

tangent vector $\mathbf{v}$ to M at $\mathbf{x}_0$ has the form $\mathbf{v} = (D\psi)\mathbf{a}$ where $\mathbf{a}$ is in U so is in R^n

Now write

$(\mathbf{D}\boldsymbol{\varphi})\ \mathbf{v} = (\mathbf{D}\boldsymbol{\varphi})\ (\mathbf{D}\boldsymbol{\psi})\mathbf{a}$.

But in (**) I just showed that $(\mathbf{D}\boldsymbol{\varphi})(\mathbf{x})\ (\mathbf{D}\boldsymbol{\psi})(\mathbf{u}) = \mathbf{0}$ for *any* $\mathbf{u}$ in U, so it must be true for $\mathbf{u} = \mathbf{a}$. Thus

$(\mathbf{D}\boldsymbol{\varphi})\ \mathbf{v} = (\mathbf{D}\boldsymbol{\varphi})\ (\mathbf{D}\boldsymbol{\psi})\mathbf{a} = 0$.

Since then $(\mathbf{D}\boldsymbol{\varphi})\ \mathbf{v} = \mathbf{0}$, and since $\mathbf{v}$ is any tangent space vector, we know that these vectors lie within the nullspace of the matrix $(\mathbf{D}\boldsymbol{\varphi})$. Basically then

$T_x M$ = tangent space of M at point x $\subseteq$ nullspace of $(\mathbf{D}\boldsymbol{\varphi})$ which is a matrix of type N x m

But we want to show that $T_x M$ = nullspace, not just $T_x M \subseteq$ nullspace. If we can show that $T_x M$ and nullspace have the same dimension, then we have shown that $T_x M$ = nullspace .

What is the dimensionality of this nullspace of (Dφ) ? I show here that it is n.

Sylvester for non square matrices says: for k x $\boldsymbol{\ell}$ matrix A, nullity(A) + rank(A) = $\boldsymbol{\ell}$

for m x N matrix (Dφ), nullity(Dφ) + rank(Dφ) = N

The tangent space dimensionality is the number of columns of (Dφ) which has n columns and N rows. Thus dim(tangent space) = n. The rank of (Dφ) is m because **c** is a regular point. Thus,

nullity(Dφ) + m = N

Therefore

nullity(Dφ) = n and so

tangent space of M at point x = nullspace of (**Dφ**) = space of dimension n

This concludes the proof of Claim 2.

[74] The rest of Chapter 6 provides **examples** of the regular value theorem idea.

The first several examples relate to surfaces defined by a **single equation** φ(**x**) = c (rather than m equations). As noted earlier, an equation like this defines level surfaces in R^n , the space of **x**. We are used to such level curves when c = φ(x,y) and you imagine φ is height or temperature (topo lines or isotherms).

How does the above regular value theorem apply to a single function like this? Well, Dφ is just a row vector whose transpose is grad φ. Only if all elements of the row vector vanish does Dφ drop below its maximum rank of 1, and if all entries vanish (all gradient components vanish) then you have a singular point for c in φ(**x**) = c. This is then the Buck idea that a "critical point" is a place where all first partials vanish, so we have made now a connection between (1) singular points c in the general theorem **c** = **φ**(**x**) and (2) critical points for c = φ(x) which is to say f(x) = φ(x)-c.

What about dimensions of things for a single equation? We have N = N, m=1 (one equation), n = N-1. The surfaces are N-1 dimensional (candidate manifolds). Codimension is 1.

What about the tangent space which is the nullspace of (Dφ) = (grad φ)^T. Well, the vectors of that nullspace are those for which $(\nabla\varphi)^T$ v = 0, which is to say **∇**φ • **v** = 0. Here is the idea that **∇**φ is normal to the surface. So the null space is $(\text{grad}\varphi)^{\perp}$, it all makes fine sense.

Note added 2.14.16. Suppose $\mathbf{e}_i$ form a basis for the tangent space $T_x M$. Then we know $(D\varphi)^T \mathbf{e}_i = 0$ There are n of these $\mathbf{e}_i$ since $\dim T_x M = n$. Dφ has n columns and N rows. Write this out:

$$\begin{array}{lll} \partial_1\varphi_1 & \partial_2\varphi_1 \ldots\ldots\ldots\partial_n\varphi_1 & (e_i)^1 = 0 \\ \partial_1\varphi_2 & \partial_2\varphi_2 \ldots\ldots\ldots\partial_n\varphi_2 & (e_i)^2 \\ \partial_1\varphi_3 & \partial_2\varphi_3 \ldots\ldots\ldots\partial_n\varphi_3 & (e_i)^3 \\ \ldots\ldots & & \text{---} \end{array}$$

$\Sigma_j \partial_j\varphi_k (e_i)^j = 0$ but a far cry from what I am looking for: $\mathbf{e}^k \bullet \mathbf{e}_i = \delta^k{}_i$
which would say $\Sigma_j (e^k)_j (e_i)^j = \delta^k{}_i$

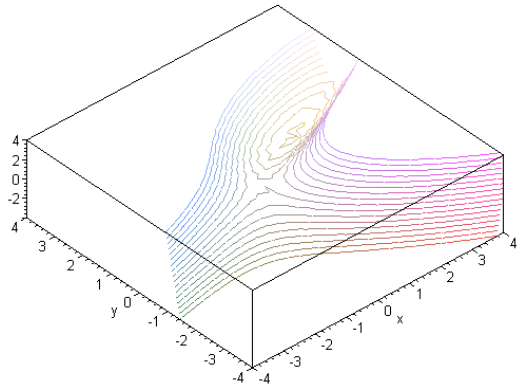
but wait: that works if you select $(e^k)_j = \partial_j \varphi_k = \partial\varphi_k/\partial x^j$. But I don't think this is my silver bullet.

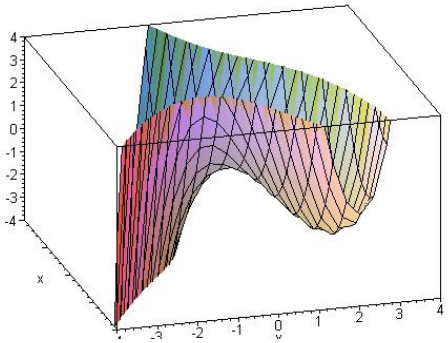
Now we start into the examples.

Example 6.12, $\varphi(x,y) = xy$ so $D\varphi = \nabla\varphi = \begin{pmatrix} y \\ x \end{pmatrix}$. Here $R^N = R^2$ and n = 1 so surfaces are just curves. This $\nabla\varphi$ vanishes only at the origin, so all other points are regular so the surface $\varphi(x,y) = c$ (that is, $xy = c$) is a manifold for any c other than 0. Each level curve is two hyperbola sections in the x-y plane. Each of these two piece curves is a manifold, but Sja does not mention the two pieces issue -- the fact that it seems these manifolds are disconnected. At the singular point = 0, the solution $\varphi^{-1}(0)$ is the two axes. This space self intersects and that confirms that it is not a manifold. Plotted as c = z = xy, we get a normal saddle point at x = 0. At this point, both partials of $\varphi = xy$ vanish, so it is a Buck critical point. But also that makes rank $D\varphi$ be 0 so that is a Sjamaar singular point. His plots of the level curves and gradients are accurate. I did it myself just for fun but threw it out.

Example 6.13, $\varphi(x,y) = x^3+y^3-3xy$. This similar to the example of Buck page 356. $\nabla\varphi = D\varphi = \begin{pmatrix} 3(x^2-y) \\ 3(y^2-x) \end{pmatrix}$.

This example has m = 1 equations and N = 2 variables. So $\varphi(x,y) = c$ specifies a curve in 2D space, which curve is a 1D manifold,. I have plotted these curves on the left below for different values of c. Each of these curves is a horizontal slice of the 3D surface shown on the right -- each slice is a different c (z = c).
.
Looking at $D\varphi$, singular point then is when $x^2 = y$ *and* $y^2 = x$, so $y^4 = y$ so $y^3 = 1$ so y = 1 and then x = 1. So one singular point is at (x,y) = (1,1) and this then is the place Buck shows a minimum, that is, a critical point. Another singular point is at (x,y) = (0,0) and that is a saddle point. Here are my Buck pictures.





What is the tangent space here? We must solve $(D\varphi)\mathbf{v} = 0$ so TS is the nullspace of $D\varphi$. But we know that $\nabla\varphi$ is the normal to the surface, so we are just solving these equations

$$(x^2-y)v_1 + (y^2-x)v_2 = 0$$

and this tells you the ratio of v_2/v_1.

$v_2 = -[(x^2-y)/(y^2-x)]v_1$ so $\mathbf{v} = v_1(1,[(x^2-y)/(y^2-x)])$ = a tangent vector

Remember that you draw these in the level curve space on the left above, not on the right. He has them all plotted.

Example 6.14. Here for $\mathbf{x}$ in R^N we take $\varphi = \mathbf{x}\bullet\mathbf{x}$ and $\nabla\varphi = 2\mathbf{x}$ so only singular point is at $\mathbf{x} = 0$. There is m = 1 equation, N = N, so solution manifold will be dimension N-1. The manifold $\varphi^{-1}(c=r^2)$ is a hypersphere of radius r, and for c< 0 the solution set is empty. Easier to think $\nabla\varphi = 2\mathbf{r}$ so the gradient is normal to the surface, and the tangent vectors are such that $2r\bullet v = 0$. That is to say, to find the tangent space we solve $(D\varphi)\mathbf{v} = 0$, but this IS the equation $2r\bullet v = 0$. Of course $\mathbf{r}\bullet\mathbf{v} = 0$ describes the obvious tangent space for $\mathbf{v}$. So hypersphere is a manifold certainly for any c>0.

Comment. In the above example, the hypersphere surface is N-1 dimensional and codimension = 1 as usual. If you select c = 0, sphere is a point. This is not a manifold of dimension N-1 but it happens to be one of dimension 0. So technically you can have a critical point correspond to a manifold, but that manifold will be "the wrong dimension", here 0 instead of N-1.

[76] **Example 6.15.** Now we have **two φ functions** (m=2) on R^4 as shown. Then Dφ has 2 rows and 4 columns as shown. Critical point would be where all 2x2 dets vanish. We see this requires $x_1 = x_2 = 0$ and then that in turn requires $x_3 = x_4 = 0$ and we end up with $\mathbf{x} = 0$ being the only value where Dφ drops below rank 2, and this corresponds to c = 0 so c = 0 is the only singular point. Staying away from that, we have $M = \varphi^{-1}(c)$ being an n=2 manifold (a 2-manifold) within R^4 so codimension = 2.

The tangent space requires $(D\varphi)\mathbf{v} = 0$ and he shows that $\mathbf{v} = \mathbf{e}_2, \mathbf{e}_4$ are tangent space spanning vectors at all points $\mathbf{x}$ on M (unusual). Hard to plot this in R^4 to see what that looks like.

Since this is a 2-manifold in R^4, could you plot it in R^3? For example, let coords by x,y,z,w. Then for c = (1,0) we would get

$x^2+y^2= 1$
$xz-yw = 0$ $\quad w = xz/y$.

In x,y,z space you have a cylinder of radius 1 parallel to the z axis, z can take any value and then w as shown. For any point on the cylinder, you have normal vector in the x-y plane, and tangent vectors in that plane. But these have nothing to do with the normal and tangent vectors in R^4.

Example 6.16. This does not really fit into the picture because now we have φ(A) = C where A and C are square *matrices* which are n x n. So we now have n^2 equations, I guess that is OK. The particular function φ of interest in this example is $\varphi(A) = A^TA$ so our equation set is $A^TA = C$. Since A^TA is symmetric, we must have C = a symmetric matrix. So our "variable $\mathbf{x}$" here is matrix A which you could I suppose write as a column vector if you wanted.

Now what does (Dφ) look like? It is (Dφ(A)) and so is some kind of tensor since φ(A) is itself a matrix. Maybe write

$$C_{ij} = [\varphi(A)]_{ij} = (A^TA)_{ij} = (A^T)_{ik}A_{kj} = A_{ki}A_{kj}$$

$(A)_{ij} = A_{ij}$

$(A + dA)_{ij} = A_{ij} + (dA)_{ij}$ small variation in A

Now define

$dA = h\,B$ where h is very small.

Then

$\varphi(A+dA) = \varphi(A+hB) = (A+hB)^T(A+hB) = (A^T + hB^T)(A+hB) = A^TA + h(B^TA + A^TB) + h^2B^TB$

$\varphi(A+dA) - \varphi(A) = h(B^TA + A^TB) + h^2B^TB$

$D\varphi(A) = B^TA + A^TB$ if $D\varphi(A) \equiv [\varphi(A+dA) - \varphi(A)]/h$

This is the change in $\varphi(A)$ if we go in the B direction, so to speak.

Try components: recall $(D\varphi)_{ij} = \partial_j\varphi_i$ so first index on $(D\varphi)_{ij}$ is the function label.

$[D\varphi(A)]_{k\ell,ij} = \partial[\varphi(A)]_{k\ell}/\partial A_{ij} = \partial[A^TA]_{k\ell}/\partial A_{ij}$

$= \Sigma_s\, \partial[\, A_{sk}A_{s\ell}\,]/\partial A_{ij}$

$= \Sigma_s\, [\, A_{sk}\,(\partial A_{s\ell}/\partial A_{ij}) + (\partial A_{sk}/\partial A_{ij})A_{s\ell}\,]$

$= \Sigma_s\, [\, A_{sk}\,\delta_{s,i}\delta_{\ell,j} + \delta_{s,i}\delta_{k,j}A_{s\ell}\,]$

$= [\, A_{ik}\delta_{\ell,j} + \delta_{k,j}A_{i\ell}\,]$

So far then I have

$[D\varphi(A)]_{k\ell,ij} = \partial[\varphi(A)]_{k\ell}/\partial A_{ij} = \delta_{\ell,j}\,A_{ik} + \delta_{k,j}A_{i\ell}$ = a rank-4 tensor

Try doing this

$\Sigma_{ij}\,[D\varphi(A)]_{k\ell,ij}\,B_{ij} = \Sigma_{ij}\,[\,\delta_{\ell,j}\,A_{ik} + \delta_{k,j}A_{i\ell}]\,B_{ij} = [\Sigma_{ij}\;\delta_{\ell,j}\,A_{ik}B_{ij} + \Sigma_{ij}\,\delta_{k,j}A_{i\ell}B_{ij}]$

$= [\Sigma_i\; A_{ik}B_{i\ell} + \Sigma_iA_{i\ell}B_{ik}] = (A^TB)_{k\ell} + (B^TA)_{k\ell} = [\, A^TB + B^TA]_{k\ell}$

Then as a matrix equation we can write

$[D\varphi(A)]\,B = A^TB + B^TA$ // this is the meaning of p 77 A

4 2

where on the left I am contracting a rank-4 tensor with a rank-2 tensor to get a rank-2 tensor. Using the other method above I got

$D\varphi(A) = B^TA + A^TB \qquad dA = hB$

So at least both methods give the same thing,

We end up with a matrix which is n x n and I guess we want to show that it is always full rank n so we have no singular points.

Question: does a symmetric matrix always have full rank? No, obviously. Somehow you would need to lay out the matrix in a linear n^2 vector and talk about this situation.

Where does "rank" determine something?

OK, now 2 PM Monday, I am out of time. This example is not very important, I was just fiddling here. Today is 7/20/15 and I won't be able to resume until Fri 7/31/15 assuming I get back here.

Resume 8/13/15 (off by 14 days!) So we have shown that $C = A^TA$ with C being a symmetric constant matrix results in $[(D\varphi)(A)] = B^TA + A^TB$ = a matrix which let us call D. The matrix B is that matrix in whose "direction" A is allowed to change while maintaining $C = A^TA$. Sja names this matrix C which is confusing since we already have a matrix named C! Well, he avoids writing $C = A^TA$, but it is confusing.

So consider the matrix $D \equiv B^TA + A^TB$. We want to know if this has full rank or not. Seems to me full rank would mean rank = n since these are nxn matrices. But I don't think this is the right rank concept for this example. The equation $\varphi(A) = A^TA = C$ is a mapping from the space $R^n x R^n$ into the space of all symmetric matrices which he calls W. And he refers to $R^n x R^n$ as V, so then $\varphi: V \rightarrow W$ and of course W is a subset of V. The dimension of V seems to be n^2. For space W we have $C_{ij} = C_{ji}$ so how many equations is that? A triangular part of C has $(n^2-n)/2$ dimension, so there are then $(n^2-n)/2$ equations, so I would guess that the dimension of W is $\dim(W) = n^2 - (n^2-n)/2 = n^2 - (n^2/2) + n/2 = (n^2 + n)/2$. Call this quantity s so $s = (n^2 + n)/2$. Then $\varphi: V \rightarrow W$ is then $\varphi: R^{n*n} \rightarrow R^s$. If I linearize everything into vectors, I suppose I could obtain $(D\varphi)$ as a huge matrix and then ask about its rank in the same sense as all our previous example. Maybe nullity = 0 here so rank+nullity = range-dim, so need rank = s and this would be the full rank. I have seen in "the meaning of" how having full rank means you can solve the equation v = Mu and thus declare things to be 1-to-1. Somehow Sja uses this concept to say that D will have full rank if you can solve $D \equiv B^TA + A^TB$ for B, but I cannot follow that argument. He goes on to show than that the solution is B = (1/2)DA. Is that really true?

$B^TA + A^TB = [(1/2)DA]^TA + A^T(1/2)DA = (1/2)A^TD^TA + 1/2A^TDA = A^TDA$

So in fact B = (1/2)DA does not solve this equation even if $A^TA = 1$. I see now that this example is only interested in C = 1, I did not realize that at first (orthogonal matrices). Let's try B = (1/2)AD instead:

$$B^T A + A^T B = [\,(1/2)AD]^T A + A^T[\,(1/2)AD] = (1/2)\, D^T A^T A + (1/2) A^T AD = D$$

so that then is the correct solution. So somehow the fact that $D \equiv B^T A + A^T B$ with symmetric D has a solution for B tells us that our effective DA thing has "full rank" and as a result our set of orthogonal matrices forms a manifold. I will have some notes to give him on this I think.

This finally concludes Chapter 6, a long voyage. Now we need some meta notes I suppose. And starting meta notes at 7:40 PM is not a good plan, so wait till tomorrow.