OPEN SUB-MANIFOLD OF $\mathbb{R}^n$
(DRAFT)

LARRY SUSANKA

Contents

Date: April 27, 2016.

Manifolds are topological spaces that locally "look like" $\mathbb{R}^n$ for some fixed n, and the simplest example of such a space is an open subset of $\mathbb{R}^n$ itself. Discussion of issues such as differentiability, integrability and solutions of differential equations on manifolds are intended to mimic the more familiar correlates in $\mathbb{R}^n$.

Before embarking on this journey we must make sure they really *are* more familiar, so we will build a "tinker-toy" model of our ultimate goal, confined to open subsets of $\mathbb{R}^n$, and see how to generalize later. Though not in final form, the discussion below would be perfectly adequate *if all we cared about were manifolds which are open subsets of* $\mathbb{R}^n$. *And they illustrate issues that must be handled somehow even with these, the familiar cases from Multivariable Calculus.*

Let $\mathcal{U}$ denote a nonempty open subset of $\mathbb{R}^n$. The set $\mathcal{U}$ is not "reality" but after we all agree on units, coordinate axes, origin and so forth it can be identified with something "real" and we presume this has been agreed-upon and done at the outset of the discussion. $\mathcal{U}$ is "the real thing," distances and angles in $\mathcal{U}$ are "real" and any other calculations we may do in $\mathcal{U}$ are to be taken literally, with their naive interpretation. If we change coordinates in $\mathcal{U}$, as we might do to aid in solving differential equations for instance, we would need to issue consistent instructions about how to translate features involving our new coordinates back to features of $\mathcal{U}$.

1. Differentiability

Functions between manifolds will be important, and we will insist and assume in this section that every function mentioned is infinitely differentiable in the "ordinary calculus" sense on its domain. This is a convenience, allowing us to make statements without continual reference to special cases and caveats.

Recall that if F is a function from any subset S of $\mathbb{R}^n$ to $\mathbb{R}^m$ then $F = (F^1, \ldots, F^m)$ where the F^i are the real-valued coordinate functions of F. This m-tuple is represented as a row with commas as a purely typographical convenience. When involved in matrix operations it is a column, an $m \times 1$ matrix.

If p is in the interior S^o of S, $F'(p)$ is an $m \times n$ matrix, the best linear approximant to F near p in the sense that

$$\lim_{x \to p} \frac{\| F(x) - F(p) - F'(p)\,(x-p) \|}{\| x - p \|} = 0.$$

This implies that the ij-th entry of $F'(p)$ must be $D_j F^i(p)$, the j-th partial derivative of the i-th coordinate function of F.

If the domain S of F is *not* open, we still want to be able to define F' in some cases. There are problems with uniqueness of F' on the topological boundary of S and we deal with this in the simplest possible way.

Suppose $p \in S \cap \delta S$ where δS is the topological boundary of S. It may be that F can be extended to a function whose domain contains some open ball around p, and this extension might be differentiable. We say that F is differentiable at p if there is an extension of this kind that is differentiable at p and if the derivative at

p of *every* extension of this type which is differentiable at p is the same. We denote this derivative $F'(p)$ as before.

The exact conditions under which this can be done may be interesting but will not concern us. For the conditions we will care about, differentiability on the boundary will have an obvious meaning.

And for now we will assume that the domains of our functions are open, thereby sidestepping the whole issue.

The rows and columns of $F'(p)$ have their own important interpretations.

First, the j-th column is the tangent vector at p to the curve obtained by freezing all domain coordinates except the j-th and, at that coordinate, moving through p^j at unit rate in the domain.

The i-th row is row vector, the $1 \times n$ matrix $(F^i)'(p)$, which tells us among other things the direction of "no change" of the i-th coordinate function. For any vector v in $\mathbb{R}^n$ the number $(\, D_1F^i(p) \ \cdots \ D_nF^i(p)\,)\, v = (F^i)'(p)\, v$ is called the directional derivative of F^i in the direction of v at p. The chain rule tells us that if you are passing through p in $\mathbb{R}^n$ on a curve which has velocity vector v at p then, confined to this curve, the values of F^i are changing at this rate. When this number is 0 the curve could be trapped on the constant-F^i surface.

2. Tangent Vectors, First Try

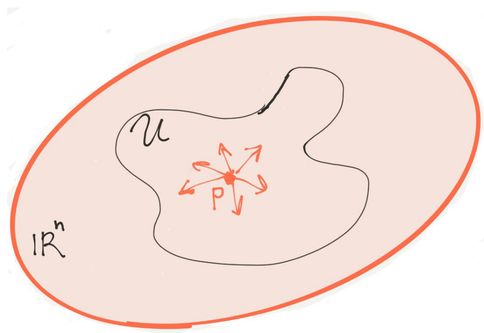
At each point p in $\mathcal{U}$ we need to keep track of all the directions one can move and the speeds in those direction, starting from p, while remaining in $\mathcal{U}$.

There are (at least) two equivalent ways of doing this. The first will make sense to folks who've had a little Linear Algebra, and the second is less natural and somewhat unmotivated. However it is the second way, not the first, that applies to general manifolds so we present both.

Our first try, approach #1(a) (which will not carry enough information for our later purposes), starts off by considering the set

$$\{\, (p, v) \ | \ p \in \mathcal{U} \text{ and } v \in \mathbb{R}^n \,\} \subset \mathbb{R}^n \times \mathbb{R}^n.$$

For each p the subset with first coordinate p may be thought of as the collection of all "arrows" emanating from the point p representing velocity vectors to parameterizations of curves through p. These "arrows starting at p" can be made to form a vector space with the obvious operations, though we make no definition of vector operations involving (p, v) and (q, w) unless $p = q$.



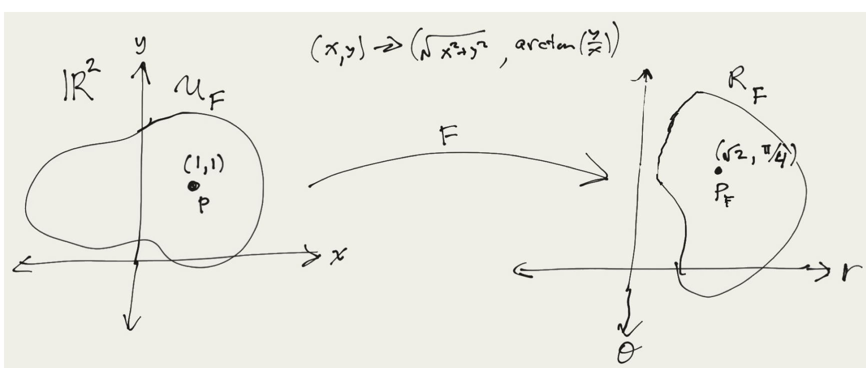
Recall that one of our goals is to solve differential equations or understand physical phenomena such as rotations and, as everyone who has studied such things knows, a judicious choice of coordinates can make an amazing difference in the simplicity of calculations. But such change of coordinates is often only local. For us, that is not a huge problem: derivatives also depend on local behavior. It is essential that we learn how to translate these local results.

The classic example is polar coordinates. Suppose our manifold $\mathcal{U}$ is $\mathbb{R}^2$ and point $p = (1,1)$. Let $\mathcal{R}_F = \{\, (r,\theta) \mid r > 0 \text{ and } -\pi/2 < \theta < \pi/2 \,\}$. So $\mathcal{R}_F$ is also an open subset of $\mathbb{R}^2$ and the right half-plane $\mathcal{U}_F = \{\, (x,y) \mid x > 0 \text{ and } y \in \mathbb{R} \,\}$ contains p and is contained in $\mathcal{U}$. The maps

$$F\colon \mathcal{U}_F \to \mathcal{R}_F \text{ given by } F(x,y) = \left(\sqrt{x^2+y^2},\, \arctan(y/x) \right)$$

$$\text{and} \quad F^{-1}\colon \mathcal{R}_F \to \mathcal{U}_F \text{ given by } F^{-1}(r,\theta) = (\, r\cos(\theta),\, r\sin(\theta)\,)$$

are inverse to each other and $F^{-1}(\sqrt{2},\pi/4) = (1,1) = p$ and $F(p) = (\sqrt{2},\pi/4)$.



F and F^{-1} are both one-to-one, onto and infinitely differentiable. This is about as nice an association of two regions in the plane as you can get other than translation or rotation or scaling by a numerical factor.

But there is no way to extend polar coordinates to a differentiable and invertible function onto $\mathcal{U} = \mathbb{R}^2$.

The best you can do is extend F^{-1} so it maps onto the open subset of $\mathcal{U}$ obtained from $\mathcal{U}$ by deleting a ray through the origin. It is inevitable that some calculations and constructions will require two or more versions of these coordinates.

The moral of this story is that in our applications we must allow for useful coordinate systems ***many of which can be defined only on subsets of our manifold $\mathcal{U}$***: we cannot require that they be defined on all of $\mathcal{U}$. And the families of coordinate systems we want to consider will vary with the point $p \in \mathcal{U}$ in the sense that each should be defined on some neighborhood of p but need not be defined at all away from p.

There needs to be some kind of accounting of the possible coordinate systems that might be useful around p.

3. Tangent Vectors

And after changing coordinates, a velocity vector of a curve corresponds to a *real* velocity vector that we have all agreed should be interpreted literally. We need to keep track of which velocity vectors for the coordinate-changed curves are to be associated with which "real" velocities. This is essential, for instance, in keeping track of physical quantities such as kinetic energy or momentum.

The simple (p, v) from approach $\#1(a)$ doesn't contain any of this information. We'll try to do better with approach $\#1(b)$, which we describe now.

Use the identifier *id* to stand for the easiest possible coordinate system, the identity map on $\mathcal{U}$. Then define $\mathcal{C}_p$ to be the set of all differentiable coordinate systems defined on a neighborhood of p with domain contained in $\mathcal{U}$. A function $x\colon V_x \to R_x$ is in $\mathcal{C}_p$ if and only if $p \in V_x \subset \mathcal{U}$ and V_x is open and x is differentiable, one-to-one and onto $R_x \subset \mathbb{R}^n$.

We now let

$$\mathcal{A}_p = \{ \, (x, p_x, v) \;\mid\; x \in \mathcal{C}_p, \; p_x = x(p), \; v \in \mathbb{R}^n \, \}$$

and make an equivalence relation on $\mathcal{A}_p$ as follows.

We want (x, p_x, w) to be equivalent to (id, p, v) if $x'(p)\, v = w$. Equivalent triples represent the velocity of the same curves viewed from different coordinate systems. *There is no way to know what the vector w represents without either implicit or explicit reference to the coordinate system x.* Explicit is better. This is the extra element we need to smoothly manage coordinate changes.

We now define $[x, p_x, w]$ to be the class of (x, p_x, w). So if (y, p_y, u) and (x, p_x, w) are both in $[id, p, v]$ then

$$(\, x'(p)\,)^{-1}\, w = (\, y'(p)\,)^{-1}\, u = v$$

and it follows that

$$w = x'(p)(\, y'(p)\,)^{-1}\, u = (x \circ y^{-1})'(p_y)\, u.$$

The composite derivative carries the tangent vector to a curve in $\mathcal{U}$ represented in R_y to the tangent vector of the same curve represented in R_x.

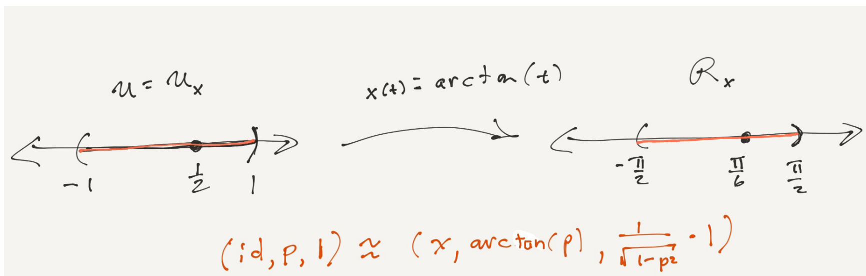
This structure allows us to create vector operations on *on this set of equivalence classes* by

$$[id, p, v] + [id, p, w] = [id, p, v + w] \quad \text{and} \quad c\,[id, p, w] = [id, p, c\,w].$$

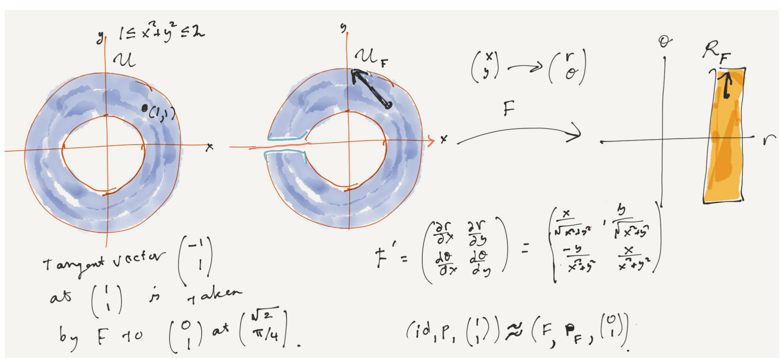
Focusing on the representative (id, p, v) of any $[id, p, v]$ this is just the "arrow" (p, v) from approach $\#1(a)$ with the same vector operations we were discussing before, but now incorporating also the effect on velocities created by all possible coordinate changes. Vector operations can be carried out using any fixed coordinate system in $\mathcal{C}_p$.

4. A Few Examples

Here is a one dimensional example.



And here is a two dimensional example.



5. The Tangent Bundle and Vector Fields

We define $\boldsymbol{\mathcal{S}(\mathcal{U})_p}$ to be the set of equivalence classes with these vector operations, called the **tangent space** to $\mathcal{U}$ at p, and the collection of all of these tangent spaces,

$$\boldsymbol{\mathcal{S}(\mathcal{U})} = \bigcup_{p \in \mathcal{U}} \mathcal{S}(\mathcal{U})_p,$$

is called the **tangent bundle.**

Suppose $g \colon \mathcal{U} \to \mathbb{R}^n$ is a vector valued function. We will call

$$\overline{g} \colon \mathcal{U} \to \mathcal{S}(\mathcal{U}) \text{ given by } \overline{g}(p) = [\, id, p, g(p)\,]$$

a **vector field on $\boldsymbol{\mathcal{U}}$** or a **section of the tangent bundle $\boldsymbol{\mathcal{S}(\mathcal{U})}$**.

In applications, we will only be interested in vector fields for which g is differentiable.

6. Cotangent Vectors

We have worked with curves in $\mathcal{U}$, functions whose domain is an interval in $\mathbb{R}$ with values in $\mathcal{U}$, and now we will consider the dual concept, namely real *valued* functions defined and differentiable on some open set contained in $\mathcal{U}$.

Localizing, we let $\mathcal{F}_p$ denote the set of all such functions whose domain contains p. The derivatives of such functions evaluated at p are linear transformations which are represented as row matrices, elements of $\mathbb{R}^{n*}$.

$$f'(p) = (\, D_1 f(p)\, D_2 f(p) \,\cdots\, D_n f(p)\,).$$

Note that if c is any parametrized curve in the domain of f passing through p at α then $f \circ c$ is a real valued function defined on an interval and

$$(f \circ c)'(t) = f'(c(t))c'(t) \quad \Longrightarrow \quad (f \circ c)'(\alpha) = f'(p)c'(\alpha).$$

If σ is the row vector $f'(p) \in \mathbb{R}^{n*}$ and $v \in \mathbb{R}^n$ is the column $c'(\alpha)$ then

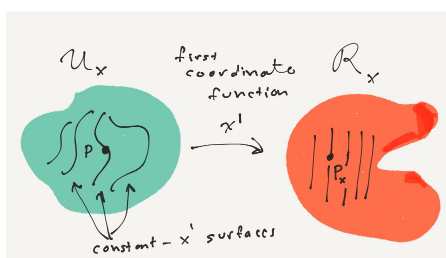
$$(f \circ c)'(\alpha) = \sigma\, v.$$

This corresponds to a directional derivative: if you are moving through p in the direction and at the pace determined by c the values of f are changing at rate $\sigma\, v$.

Define $\mathcal{A}_p^* = \{\, (x, p_x, \sigma) \;|\; x \in \mathcal{C}_p,\ \sigma \in \mathbb{R}^{n*}\,\}$.

Declare (x, p_x, σ) equivalent to (id, p, τ) if $\tau = \sigma\, x'(p)$.

Equivalent triples describe the rate of change of the same collection of real valued functions viewed from different coordinate systems. Coordinate system x is incorporated explicitly as part of (x, p_x, σ) *because there is no way to know what σ means without it.*



Suppose $f'(p) = \tau$ and $(f \circ x^{-1})'(p_x) = \sigma$. Then

$$\sigma = (f \circ x^{-1})'(p_x) = f'(p)\,(x^{-1})'(p_x) = \tau\,(x'(p))^{-1}.$$

We now let $[x, p_x, \sigma]$ denote the class of (x, p_x, σ).

If (y, p_y, μ) and (x, p_x, σ) are both in $[id, p, \tau]$ then

$$\tau = \sigma\, x'(p) = \mu\, y'(p) \quad \Longrightarrow \quad \sigma = \mu\, y'(p)\,(x'(p))^{-1} = \mu\,(y \circ x^{-1})'(p_x).$$

Define vector operations on these classes by

$$[id, p, \tau] + [id, p, \sigma] = [id, p, \tau + \sigma] \quad \text{and} \quad c\,[id, p, \tau] = [id, p, c\,\tau].$$

7. The Cotangent Bundle and One-Forms

Define $\boldsymbol{\mathcal{S}^*(\mathcal{U})_p}$ is the set of equivalence classes with these vector operations, called the **cotangent space** to $\mathcal{U}$ at p, and the collection of all of these cotangent spaces,

$$\boldsymbol{\mathcal{S}^*(\mathcal{U})} = \bigcup_{p \in \mathcal{U}} \mathcal{S}^*(\mathcal{U})_p,$$

is called the **cotangent bundle.**

If $g \colon \mathcal{U} \to \mathbb{R}^{n*}$ is a covector valued function we will call

$$\underline{g} \colon \mathcal{U} \to \mathcal{S}(\mathcal{U}) \text{ given by } \underline{g}(p) = [\, id, p, g(p)\,]$$

a **covector field** or **one-form on $\boldsymbol{\mathcal{U}}$** or a **section of the cotangent bundle $\boldsymbol{\mathcal{S}^*(\mathcal{U})}$**.

As before we will only be interested in one-forms for which g is differentiable.

8. Push Forward, Pull Back

We now suppose $F \colon \mathcal{U}^1 \to \mathcal{U}^2$ is a differentiable map from n-manifold $\mathcal{U}^1$ to m-manifold $\mathcal{U}^2$.

We don't suppose F is one-to-one or onto, though it often will be.

Every tangent vector $[id, p, v] \in \mathcal{S}(\mathcal{U}^1)_p$ can be "pushed forward" by F to a tangent vector in $\mathcal{S}(\mathcal{U}^2)_{F(p)}$ by

$$[\, id, p, v\,] \longrightarrow [\, id, F(p), F'(p)\, v\,].$$

The map which we denote $F_{*p} \colon \mathcal{S}(\mathcal{U}^1)_p \to \mathcal{S}(\mathcal{U}^2)_{F(p)}$ is a vector space homomorphism, and the dimension of the range is the rank of $F'(p)$. If F has differentiable inverse we must have $m = n$ and this provides an isomorphism between these two vector spaces.

The function $F_* \colon \mathcal{S}(\mathcal{U}^1) \to \mathcal{S}(\mathcal{U}^2)$ is defined at each tangent space by F_{*p}. If $F'(p)$ is invertible for every $p \in \mathcal{U}^1$ then $F(\mathcal{U}^1)$ must be an open subset of $\mathcal{U}^2$ and if F is one-to-one F_* provides a way of identifying $\mathcal{S}(\mathcal{U}^1)$ with $\mathcal{S}(\, F(\mathcal{U}^1)\,)$ which can, itself, be regarded as a subset of $\mathcal{S}(\mathcal{U}^2)$.

And in that case (F is one-to-one and F' invertible) if $\overline{g}$ is a differentiable vector field on $\mathcal{U}^1$ then $F_* \circ \overline{g}$ is a choice of a tangent vector for each point in $F(\mathcal{U}^2)$ given at each $F(p)$ by $[\, id, F(p), F'(p)g(p)\,]$.

A particularly interesting case is when F is a one-to-one parameterization of a curve in $\mathcal{U}^2$. Specifically, we want $F \colon (a, b) \to \mathcal{U}^2$ to be differentiable. Let $g(c)$ denote the unit vector in the positive direction in $\mathbb{R}$ for each $c \in (a, b)$. This is, of course, just the number 1 conceived of as a rightward arrow one unit long for each c.

Then $\overline{g}(c) = [id, c, 1]$ for each c. This represents the tangent vectors for the identity map, a rightward arrow one unit long "attached" to each $c \in (a, b)$.

So $F_* \circ \overline{g}(c) = [\, id, F(p), F'(p)\,]$ which is just the collection of tangent vectors to F along the curve.

To repeat: the tangent vectors along the curve are the pushforward of the positive unit vector field on the domain.

Similarly, F can be used to "pull back" covectors and covector fields from $\mathcal{U}^2$ to $\mathcal{U}^1$.

For each $[\,id, F(p), \tau\,]$ with $\tau \in \mathbb{R}^{m*}$ define $F_p^*(\,[\,id, F(p), \tau\,]\,) = [\,id, p, \tau\, F'(p)\,]$.

Again, the dimension of the range of each F_p^* in $\mathcal{S}^*(\mathcal{U}^1)_p$ is the rank of $F'(p)$.

And if F is one-to-one we can define $F^* \colon \mathcal{S}^*(\mathcal{U}^2) \to \mathcal{S}^*(\mathcal{U}^1)$ at each p and apply it to covector fields as well.

$$F^* \circ \underline{g}(p) = [\,id, p, g(p)\, F'(p)\,].$$

An interesting case is when F is a differentiable real valued function defined on $\mathcal{U}^1$.

In the 1×1 case, covectors and vectors coincide: they are both real numbers.

More generally, choosing $g(p) = 1$ for every p we have

$$F^* \circ \underline{g}(p) = [\,id, p, g(p)\, F'(p)\,] = [\,id, p, F'(p)\,]$$

and this covector is normally denoted $dF(p)$.

dF is the pullback of the positive unit vector field on the range of F.

9. Tangent Vectors as Equivalence Classses of Curves

10. Cotangent Vectors as Equivalence Classes of Real Valued Functions

11. Tangent Vectors as Point Derivations

12. Cotangent Vectors as a Quotient Ring of Germs