Elementary Analysis on Manifolds in $\mathbb{R}^n$

Michael D. Taylor
Emeritus Professor
Department of Mathematics
University of Central Florida
Orlando, FL 32816
E-mail: taylorseries@embarqmail.com

May 28, 2017

Abstract

We develop some of the elementary aspects of analysis on $\mathcal{C}^r$ manifolds embedded in $\mathbb{R}^n$, keeping in mind that the differential structure of such manifolds must be compatible with or inherited from that of $\mathbb{R}^n$. The aim of this paper is to provide those working in geometric algebra and geometric calculus with relatively rigorous derivations of basic properties of manifolds in such a setting. It is found convenient to construct a slightly extended version of manifold, one which includes corners, and a concept of the identity function on the manifold which takes account of the manifold's embedding in $\mathbb{R}^n$.

2000 Mathematics Subject Classification. 58A05, 58A99.
Key words and phrases. Differentiable manifolds in $\mathbb{R}^n$, geometric calculus.

Contents

1 Introduction

The inspiration for this paper is the growing field of geometric algebra and geometric calculus. (For example, [1, 3, 9, 20].) Two reasons for interest in this area are the facts that, on the one hand, there is a strong opinion from certain quarters ([6] and [7]) that this machinery ought to be the standard and *unified* way to present a great deal of mathematics (linear algebra, vector analysis, complex analysis, differential geometry, etc.), and, on the other hand, the presentation is, in some sense, more elementary and accessible than in standard approaches.

We are concerned in this paper with the setting for geometric calculus, and it should be understood that manifolds in this setting are (generally) in some $\mathbb{R}^n$ and inherit the inner product of $\mathbb{R}^n$. Thus we are concerned with Riemannian manifolds. It seems there are some problems with the way that setting is handled in geometric calculus.

Broadly speaking, there are at least two paths one can take to devolop the theory of the manifolds. On the first of these paths—the standard one in the mathematical literature—manifolds are taken to be Hausdorff, second countable topological spaces equipped with charts and atlases, and if one is doing differential geometry, they are taken to be $\mathcal{C}^\infty$. (See, for example, [11] and [21].) If, more particularly, one wishes them to be Riemannian manifolds, then they are also equipped with a positive definite inner product.

It is important to notice that this first path is an *intrinsic* one. Manifolds are not assumed to "sit" in any larger space, and any machinery or concepts (differentiation, inner products, etc.) must be defined *within* the manifold.

On the other hand, manifolds in geometric calculus are taken to be sitting in some vector space, and in this paper, we shall take this to be $\mathbb{R}^n$. This is our second path to the theory of Riemannian manifolds: manifolds embedded in $\mathbb{R}^n$. It is a setting in which one has to always be aware of *extrinsic* properties. Fortunately we know from the Nash embedding theorem, [5] and [19], that Riemannian manifolds can always be isometrically embedded in some $\mathbb{R}^n$, so there is no loss of generality to this assumption.

Now here is an important distinction between the two paths:

In mathematical literature, the machinery of calculus on manifolds (tangent vectors, derivatives, differentiability, etc.) is carefully constructed in the standard setting, that is, on manifolds *not* embedded in an ambient $\mathbb{R}^n$. (There may be exceptions where constructions are carried out in $\mathbb{R}^2$ and $\mathbb{R}^3$ to build up the reader's intuition before getting to the "real" definitions.)

I am not aware of a similar construction of such machinery for manifolds contained in $\mathbb{R}^n$. Presumably, one can establish the "same results" in both settings, but there is just enough difference between the two arenas that this is not always clear. For instance, in the standard, topological setting, tangent vectors to a manifold are likely to be differential operators acting on real-valued functions defined in the manifold, a construction which at first sight appears artificial and unduly abstract. However in the $\mathbb{R}^n$ setting, one can simply and naturally take tangent vectors to be particular elements of $\mathbb{R}^n$.

A difficulty that arises in the literature of geometric calculus is a tendency to slide back and forth between the intrinsic, topological setting and the $\mathbb{R}^n$ setting without noting the fact and to assume that what works in one arena must work in the other. This may well be true, but the logical justification may be missing.

For example, in [8], the displayed equation between (19.2) and (19.3) is

$$\partial = \partial_x = I_m^{-1}(I_m \bullet \nabla)$$

where ∂ is the vector derivative on a given manifold, I_m is the tangent pseudoscalar, and ∇ is the derivative in the ambient vector space that contains the manifold. The left side is calculated using directional derivatives of the form

$$\lim_{\tau \to 0} \frac{f(x(\tau)) - f(x)}{\tau}$$

where $x(\tau)$ is a curve in the manifold that approaches x as $\tau \to 0$. This is basically an intrinsic calculation. But the right side is calculated using directional derivatives of the form

$$\lim_{\tau \to 0} \frac{f(x + \tau v) - f(x)}{\tau}$$

where v is a tangent vector to the manifold at x. Notice that in this last expression, $x + \tau v$ is not necessarily on the manifold, so we are assuming that an extrinsic calculation gives the same result as an intrinsic one.

A second place where confusion can arise is with respect to the **identity function** of a manifold. As is remarked in 4.2 of [9], "The identity function on a manifold (which, of course, maps each point to itself) is not the trivial function one might at first suppose." The authors use x as the identity function on a manifold and P as the orthogonal projection map onto the tangent space for a manifold. They then go on to show that for an arbitrary vector a, the directional derivative of x in the direction a must be $P(a)$. Now the reader may be tempted to take a vector a which is not tangent to the manifold, to extend x to the identity on the ambient vector space containing the manifold, and to duplicate the calculation. He or she will not obtain $P(a)$. Clearly something more subtle is going on.

So here is the goal of this paper: It is to develop the basic machinery of calculus for manifolds that are contained in $\mathbb{R}^n$ and to do it in a way that is both moderately rigorous and that takes advantage of the ambient Euclidean space. It is hoped this will be of use to workers in geometric calculus and perhaps in other fields as well.

We make use of only the elementary machinery of multivariable calculus in $\mathbb{R}^n$. This is to insure that our derivations, though rigorous, are accessible to as many people as possible. The results we present are only very basic ones. They are not necessarily the most elegant, perspicuous, or concise, but we hope they are at least correct.

There is little in this paper that ought be considered new or research, though in the course of this work, it has been found convenient to give a definition of manifold a little more general than the usual one. The result is manifolds-with-corners, sometimes an infinite number of them. This is useful in assuring we can differentiate on the boundary of a manifold. We have also developed a particular construction of the identity function connected with a manifold, a construction that we call *the normal identity.* An idea of this sort seems implicit in the literature of geometric calculus. For both of these concepts, we make essential use of the fact that the manifold lies in an ambient $\mathbb{R}^n$.

2 The setting

2.1 Directional derivatives and differentiability

If f is a function with domain in $\mathbb{R}^m$ and range in $\mathbb{R}^n$, x is a point in $\mathrm{dom}(f)$, and a is a vector in $\mathbb{R}^m$, then the *directional derivative of f at the point x in the direction a* is

$$\partial_a f(x) \stackrel{\text{def.}}{=} \lim_{\lambda \to 0} \frac{f(x+\lambda a) - f(x)}{\lambda} \tag{1}$$

where it is understood that λ is a scalar. (The notation $a \bullet \partial f$ instead of $\partial_a f$ is often preferred for the directional derivative in the literature of geometric calculus.)

Remark 1. We will tend to use Greek letters such as $\lambda, \alpha, \phi, \ldots$ for real numbers and real-valued functions and latin letters such as $x, y, f, \ldots$ for points in $\mathbb{R}^n$, vectors, and multivector-valued functions.

Let $e_1, \ldots, e_m$ be the *standard basis* for $\mathbb{R}^m$, that is, $e_i = (0, \ldots, 1, \ldots, 0)$ where the lone 1 occurs in the ith position. Then by the *partial derivative of f with respect to the ith variable*, we mean, of course,

$$\frac{\partial f}{\partial \chi_i} \stackrel{\text{def.}}{=} \partial_{e_i} f.$$

If $f : U \to \mathbb{R}^n$ where U is an open set in $\mathbb{R}^m$ and $r = 1, 2, \ldots$, then we say that f is a $\mathcal{C}^r$ function or, equivalently, that it is *r-times continuously differentiable* provided

$$\frac{\partial^r f}{\partial \chi_{i_1} \cdots \partial \chi_{i_r}}(x)$$

exists and is continuous for for all $x \in U$ and for all $i_1, \ldots, i_r$. If f is $\mathcal{C}^r$ for all positive r, then we say it is $\mathcal{C}^\infty$.

If U and V are open sets of $\mathbb{R}^n$, we say that $f \colon U \to V$ is a $\mathcal{C}^r$ *diffeomorphism of U to V* provided f is one-to-one and onto and both f and f^{-1} are $\mathcal{C}^r$.

Definition 1. If $f \colon A \to \mathbb{R}^n$ where $A \subseteq \mathbb{R}^m$, we say that f is *differentiable* at a point x_0 of A provided there is a (unique) linear transformation $f'(x_0) \colon \mathbb{R}^m \to \mathbb{R}^n$ and a function g defined on an open neighborhood of 0 such that

$$f(x_0 + v) - f(x_0) \;=\; f'(x_0)\, v + g(v)$$

where $f'(x_0)\, v$ means that the linear transformation $f'(x_0)$ is acting on the vector v and where

$$\lim_{v \to 0} \frac{g(v)}{|v|} \;=\; 0.$$

If f is at least $\mathcal{C}^1$ at x_0, it is a standard result that it is differentiable there. A useful relation between differentiability and directional derivatives is the following:

$$\partial_v f(x_0) \;=\; f'(x_0)\, v. \tag{2}$$

2.2 Geometric algebra

It is not necessary for the reader to understand geometric algebra to follow our exposition, but we say a few words here to give an idea of the arena in which we hope these concepts and proofs will be useful. Those who wish to learn more about geometric algebra and geometric calculus can consult, for example, [2, 4, 9, 6, 12, 13, 14, 15, 20].

We call elements of the geometric algebra *multivectors*, and if we refer to f defined on $\mathcal{M}$, we may implicitly have in mind a multivector-valued function over a manifold in $\mathbb{R}^n$. The particular geometric algebra we tend to have in mind is the Clifford algebra $\mathcal{C}\ell_{n,0}(\mathbb{R})$, though the results we discuss should apply equally well to the Clifford algebra $\mathcal{C}\ell_{p,q}(\mathbb{R})$ where $p + q = n$.

We recall that a basis for the geometric algebra over $\mathbb{R}^n$ consists of 1 and all elements of the form $e_{i_1} \cdots e_{i_k}$ where $i_1 < \cdots < i_k$ and where we are taking the geometric product of standard basis elements $e_{i_1}, \ldots, e_{i_k}$. Since $\{e_i\}_{i=1}^n$ is a orthogonal set of vectors, these basis elements can also be written in the form

$$e_{i_1} \cdots e_{i_k} \;=\; e_{i_1} \wedge \cdots \wedge e_{i_k}.$$

It follows that every multivector field has a unique expansion of the form

$$f \;=\; \sum_{k=0}^{n} \sum_{i_1 < \cdots < i_k} \phi_{i_1 \cdots i_k}\, e_{i_1} \cdots e_{i_k}$$

where each $\phi_{i_1 \cdots i_k}$ is a real-valued function. We see from this that f is $\mathcal{C}^k$ if and only if each $\phi_{i_1 \cdots i_k}$ is $\mathcal{C}^k$. (Of course this assertion does not depend on the use of the standard basis $\{e_i\}_{i=1}^n$. We could use the expansion of f in terms of any fixed orthogonal basis $\{u_i\}_{i=1}^n$ for $\mathbb{R}^n$.)

It follows from these considerations that proofs involving f defined on $\mathcal{M}$ can usually be reduced to proofs that refer to real-valued functions ϕ on $\mathcal{M}$.

2.3 Manifolds

A p-dimensional $\mathcal{C}^k$ manifold $\mathcal{M}$ can be defined thus:

1. $\mathcal{M}$ is a second countable, Hausdorff topological space.
2. There exists a family of homeomorphisms $\{x_\alpha \colon U_\alpha \to \mathbb{R}^p,\ \alpha \in A\}$ (an *atlas*) such that each U_α is an open set of $\mathcal{M}$ and $\mathcal{M} = \cup_{\alpha \in A} U_\alpha$.
3. If x_α and x_β are members of the atlas, then $x_\alpha \circ x_\beta^{-1} \colon \mathbb{R}^p \to \mathbb{R}^p$ is a $\mathcal{C}^k$ map.

We will call such a manifold a *standard* manifold.

The maps x_α are called *charts.* Clearly they could be replaced by maps $y_\alpha \colon \mathbb{R}^p \to U_\alpha$; these are usually called *parametrizations* or *local parametrizations.*

For our purposes, it is more convenient to have maps that run from $\mathbb{R}^p$ to $\mathcal{M}$, and we shall feel free to sloppily use either of the terms *chart* or *parametrization* for such maps.

We want to give a somewhat different description of *manifold*, one that takes advantage of working in an ambient $\mathbb{R}^n$. We want our new definition to have the following properties:

1. It should include manifolds-with-corners (such as *cells*, diffeomorphic images of p-dimensional rectangles).

2. We should be able to do *calculus* at boundary points as easily (or almost as easily) as at interior points. In particular, we should be able to take limits and directional derivatives.

3. Our new description should clearly give us pretty much everything that we are already used to calling a manifold.

We will do this by constructing our definition in such a way that at every point x_0 in the boundary of $\mathcal{M}$, we can attach a "small tongue" to $\mathcal{M}$ so that we obtain an extended manifold $\mathcal{M}'$ in which x_0 is now an interior point. To put this another way, we want to be able to extend a chart on the boundary of $\mathcal{M}$ into the ambient $\mathbb{R}^n$.

We start by revisiting the definition of a chart.

Definition 2. We say that x is a $\mathcal{C}^k$ *p-dimensional chart* (*or p-chart*) *in* $\mathbb{R}^n$ (where $k, p \geq 1$) provided the following hold:

1. $x\colon U \to \mathbb{R}^n$ is a one-to-one $\mathcal{C}^k$ map where U is an open set in $\mathbb{R}^p$.

2. $x'(t_0)$ has rank p for all $t_0 \in U$.

Definition 3. We say that $\mathcal{M}$, a subset of $\mathbb{R}^n$, is a $\mathcal{C}^k$*, p-dimensional manifold-with-corners in* $\mathbb{R}^n$ provided the following is true: For every $x_0 \in \mathcal{M}$, there exists a $\mathcal{C}^k$ p-chart $x\colon U \to \mathbb{R}^n$, a point $t_0 \in U$, and a nondegenerate p-simplex P such that

1. $t_0 \in P \subseteq x^{-1}(\mathcal{M}) \cap U$,

2. $x_0 = x(t_0)$,

3. For every open subset U_0 of U, $x(U_0) \cap \mathcal{M}$ is an open subset of $\mathcal{M}$ in the subspace topology that $\mathcal{M}$ inherits from $\mathbb{R}^n$.

(See Figure 1.) We say that x is a *chart on* $\mathcal{M}$. If a point x_0 of $\mathcal{M}$ lies in the range of a chart x, we use such language as x *covers* x_0 or x_0 *lies in the coordinate patch* x. If there exists a simplex P satisfying condition 1 of Definition 3 such that t_0 is an interior point of P, then we say that x_0 is an *interior point* of $\mathcal{M}$. Otherwise, we say that x_0 is a *boundary point* of $\mathcal{M}$ and write $x_0 \in \partial\mathcal{M}$.

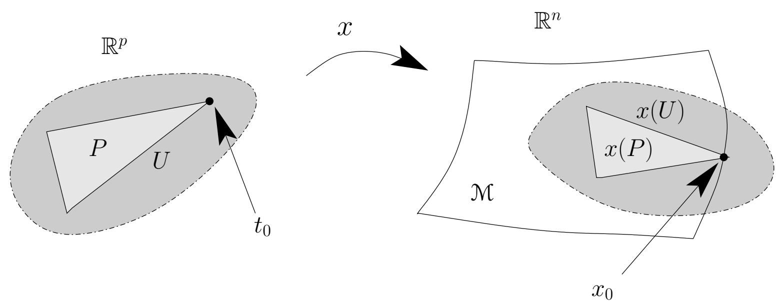

Figure 1: A chart may extend beyond the boundary of a manifold

Remark 2. Notice that if $x\colon U \to \mathbb{R}^n$ is a chart on $\mathcal{M}$ that covers x_0 and V is an open subset of U such that $t_0 \in V$, then the restriction of x to V, $x|_V\colon V \to \mathbb{R}^n$, is still a chart that covers x_0.

Remark 3. We also refer to a chart x as a *coordinate patch* or a (local) *parametrization* of $\mathcal{M}$. If $x_0 = x(\tau_1, \ldots, \tau_p)$, then we say that x assigns x_0 the coordinates $(\tau_1, \ldots, \tau_p)$.

Remark 4. Rather than continually use the clumsy phrase *manifold-with-corners*, we shall refer to objects in the sense of Definition 3 as *manifolds*. The reader should keep in mind that is not quite the thing as the usual definition of manifold which we have labelled a *standard manifold*.

How does our version of a manifold compare with the standard one?

If $\mathcal{M}$ is a standard, $\mathcal{C}^k$, p-manifold $(k \geq 1)$ and $f\colon \mathcal{M} \to \mathbb{R}^n$ is a $\mathcal{C}^k$ embedding, then it is easily checked that $f(\mathcal{M})$, the "version" of $\mathcal{M}$ we have placed in $\mathbb{R}^n$, can be equipped with charts in our sense. Thus manifolds-with-corners will include standard manifolds.

Here is an example to show that our concept of manifold is a bit broader than what one might expect:

Draw a sawtooth curve with an infinite number of teeth which decrease in size and converge to a limit point. Use this as the top of an otherwise rectangular region $\mathcal{M}$. (Figure 2.) $\mathcal{M}$ is readily seen to be a manifold-with-corners in $\mathbb{R}^2$.

The crucial question when dealing with our new version of a manifold is what happens when switching between charts. This is more complicated

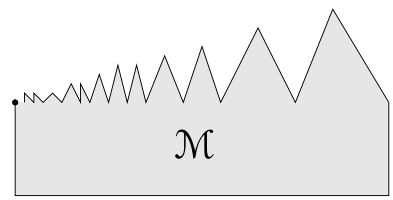

Figure 2: A sawtooth "manifold-with-corners"

than on a standard manifold (because of the boundary points) but sufficiently well-behaved that we get the usual theorems for doing calculus on manifolds. We spell out the details of switching in Proposition 3.

3 Local description of manifolds and charts

In our first result, we show that we may think of p-charts and the manifolds to which they are "pasted" as (locally) lying) in $\mathbb{R}^p$ where $\mathbb{R}^p$ is a factor of $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}$.

In all that follows, we use 0^k to denote the zero vector of $\mathbb{R}^k$.

Proposition 1. *Let $x : W \to \mathbb{R}^n$ be a $\mathcal{C}^r$ $(r \geq 1)$ p-dimensional chart in $\mathbb{R}^n$ where W is an open subset of $\mathbb{R}^p$. Suppose $x_0 = x(t_0)$ where $x_0 \in \mathbb{R}^n$ and $t_0 \in W$. Then there exist open sets U and V in $\mathbb{R}^n$ and a map $X : U \to V$ such that*

1. *X is a $\mathcal{C}^r$ diffeomorphism of U to V.*
2. *$(t_0, 0^{n-p}) \in U \subseteq W \times \mathbb{R}^{n-p}$.*
3. *$X(t, 0^{n-p}) = x(t)$ for all $(t, 0^{n-p}) \in U$. An alternate way to state this condition is*
$$X\big(U \cap (\mathbb{R}^p \times 0^{n-p})\big) \;=\; \{x(t) \,:\, (t, 0^{n-p}) \in U\}.$$
4. *If we further suppose that x is a chart on a $\mathcal{C}^r$ p-manifold $\mathcal{M}$, then U and V may be constructed in such a way that*
$$U \cap X^{-1}(\mathcal{M}) \;=\; \{(t, 0^{n-p}) \in U \,:\, x(t) \in \mathcal{M}\} \;\subseteq\; W \times 0^{n-p}.$$

Proof. If $p = n$, this is simply the inverse function theorem with X being a restriction of x and the last property of the proposition being vacuous. Therefore let us assume $p < n$.

Since x is a $\mathcal{C}^r$ chart, we know that $\{\partial_{e_i} x(t_0)\}_{i=1}^p$ is a set of linearly independent vectors, but it cannot be a basis for $\mathbb{R}^n$ since $p < n$. We can, without loss of generality, suppose that

$$\{\partial_{e_i} x(t)\}_{i=1}^p \cup \{e_{p+1}, \dots, e_n\}$$

is a linearly independent set (hence a basis for $\mathbb{R}^n$) for all t sufficiently close to t_0. (This is equivalent to saying that the wedge product of these vectors is nonzero.)

Let us define $X : W \times \mathbb{R}^{n-p} \to \mathbb{R}^n$ by

$$X(t, t') \;=\; x(t) + \tau_{p+1} e_{p+1} + \cdots + \tau_n e_n \;=\; x(t) + (0^p, t') \tag{3}$$

where $t = (\tau_1, \dots, \tau_p) \in \mathbb{R}^p$ and $t' = (\tau_{p+1}, \dots, \tau_n) \in R^{n-p}$. We see that

$$\partial_{e_i} X(t, t') \;=\; \begin{cases} \partial_{e_i} x(t) & \text{if } i = 1, \dots, p, \\ e_i & \text{if } i = p+1, \dots, n. \end{cases}$$

These are the column vectors of $\det X'(t, t')$, so $\det X'(t, t') \neq 0$ for t sufficiently close to t_0.

We can now invoke the inverse function theorem (see, for example, [10] or [16]) and the form of X in (3) to claim that there exist open sets U_0, V_0 of $\mathbb{R}^n$ such that

1. X is a $\mathcal{C}^r$ diffeomorphism of U_0 to V_0,
2. $(t_0, 0^{n-p}) \in U_0 \subseteq W \times \mathbb{R}^{n-p}$.

We see from (3) that $X(t, 0^{n-p}) = x(t)$ for all $(t, 0^{n-p}) \in W \times \mathbb{R}^{n-p}$. This must also be true for $U_0 \subseteq W \times \mathbb{R}^{n-p}$, so this gives us properties 1–3 in the statement of the proposition.

To obtain property 4 of the proposition, we now "shrink" U_0:

Since U_0 is open in $\mathbb{R}^n$, it follows that

$$W_0 \;=\; \{t \in W \;:\; (t, 0^{n-p}) \in U_0\}$$

is an open subset of W, and we also see that $t_0 \in W_0$. By Definition 3, $x(W_0) \cap \mathcal{M}$ must be an open subset of $\mathcal{M}$. Thus there must be an open subset V_1 of $\mathbb{R}^n$ such that $V_1 \cap \mathcal{M} = x(W_0) \cap \mathcal{M}$. Set $V = V_0 \cap V_1$; this is again an open subset of $\mathbb{R}^n$. Since $x(W_0) \cap \mathcal{M} \subseteq V_0$, we have $V \cap \mathcal{M} = V_1 \cap \mathcal{M} = x(W_0) \cap \mathcal{M}$. Set $U = X^{-1}(V)$; then U is an open subset of U_0 (hence open in $\mathbb{R}^n$), and $x_0 \in U$. Properties 1–3 of the proposition hold trivially for the restricted map $X : U \to V$, and we see that

$$\begin{aligned} U \cap X^{-1}(\mathcal{M}) \ &= \ X^{-1}\big(x(W_0) \cap \mathcal{M}\big) \\ &= X^{-1}\big(X(W_0 \times 0^{n-p}) \cap \mathcal{M}\big) \\ &= (W_0 \times 0^{n-p}) \cap X^{-1}(\mathcal{M}). \end{aligned}$$

It is straightforward to show that

$$(W_0 \times 0^{n-p}) \cap X^{-1}(\mathcal{M}) \ = \ \{(t, 0^{n-p}) \in U \ : \ x(t) \in \mathcal{M}\},$$

and this yields the desired result. $\square$

For our next proposition, we want to show that if we have a p-manifold lying in a q-manifold which in turn lies in $\mathbb{R}^n$, that is, $\mathcal{M} \subseteq \mathcal{N} \subseteq \mathbb{R}^n$, then we may think of them as (locally) lying in the factors $\mathbb{R}^p$ and $\mathbb{R}^q = \mathbb{R}^p \times \mathbb{R}^{q-p}$ of $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{q-p} \times \mathbb{R}^{n-q}$.

Before doing this, we need some auxiliary results which we state as lemmas.

Lemma 1. *Suppose that $\mathcal{M}$ is a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^n$ and U is an open set in $\mathbb{R}^n$ such that $U \cap \mathcal{M} \neq \emptyset$. Then $U \cap \mathcal{M}$ is also a $\mathcal{C}^r$ p-manifold.*

Proof. If $x_0 \in U \cap \mathcal{M}$, then any $\mathcal{C}^r$ p-chart $x : V \to \mathbb{R}^n$ on $\mathcal{M}$ that covers x_0 is automatically such a chart on $U \cap \mathcal{M}$. The only part to check is the existence of the nondegenerate p-simplex as demanded by Definition 3.

Since $\mathcal{M}$ is a manifold, we know there exists a nondegenerate p-simplex P in $\mathbb{R}^p$ such that $t_0 \in P \subseteq V \cap x^{-1}(\mathcal{M})$ where we have assumed $x_0 = x(t_0)$. By the continuity of x, we know that $V \cap x^{-1}(U)$ is an open set in $\mathbb{R}^p$ and $t_0 \in V \cap x^{-1}(U)$. Let $B(t_0, \epsilon)$ be the open ball in $\mathbb{R}^p$ centered at t_0 and having radius $\epsilon > 0$. We can find ϵ so small that

$$t_0 \ \in \ B(t_0, \epsilon) \ \subseteq \ V \cap x^{-1}(U).$$

We see that the set $P \cap B(t_0, \epsilon)$ is sufficiently simple that we can find in it a nondegenerate p-simplex P' such that

$$t_0 \in P' \subseteq V \cap x^{-1}(U \cap \mathcal{M}). \qquad \square$$

Lemma 2. *Suppose U and V are open sets in $\mathbb{R}^n$ and $f\colon U \to V$ is a $\mathcal{C}^r$ diffeomorphism ($r \geq 1$). Then $f'(x_0)$ has rank n at every $x_0 \in U$.*

Proof. Since $f^{-1} \circ f$ is the identity map on U, the chain rule tells us that we have $\left[\left(f^{-1}\right)'(y_0)\right]\left[f'(x_0)\right]v = v$ where $y_0 = f(x_0)$ and v is any vector in $\mathbb{R}^n$. It follows easily that if $\{v_i\}_{i=1}^n$ is an independent set of vectors, then $\{\left[f'(x_0)\right]v_i\}_{i=1}^n$ must also be an independent set. $\square$

Lemma 3. *Suppose U and V are open sets in $\mathbb{R}^n$ and $f\colon U \to V$ is a $\mathcal{C}^r$ diffeomorphism ($r \geq 1$). If $\mathcal{M}$ is a $\mathcal{C}^r$ p-manifold and $\mathcal{M} \subseteq U$, then $f(\mathcal{M})$ is also a $\mathcal{C}^r$ p-manifold.*

Proof. Let $y_0 \in f(\mathcal{M})$. We find $x_0 \in \mathcal{M}$ such that $y_0 = f(x_0)$. If $x\colon U \to \mathbb{R}^n$ is a chart on $\mathcal{M}$ that covers x_0, then it is straightforward to check that $f \circ x$ is a chart on $f(\mathcal{M})$ that covers y_0 and that the requirements of Definition 3 are satisfied. (We use Lemma 2 to check the rank of $(f \circ x)'$ at all points of U.) $\square$

Our last lemma seems an obvious result, and we state it without proof.

Lemma 4. *Suppose that $\mathcal{M} \subseteq \mathbb{R}^m$ and $m \leq n$. Then $\mathcal{M}$ is a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^m$ if and only if $\mathcal{M} \times 0^{n-m}$ is a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$.*

Proposition 2. *Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{C}^r$ ($r \geq 1$) p- and q-manifolds in $\mathbb{R}^n$, $\mathcal{M} \subseteq \mathcal{N}$, and $x_0 \in \mathcal{M}$. Then there exist open sets $U, V \subseteq \mathbb{R}^n$, a $\mathcal{C}^r$ diffeomorphism $Z\colon U \to V$, and $s_0 \in \mathbb{R}^p$ such that*

1. *$Z(s_0, 0^{n-p}) = x_0$,*
2. *$Z^{-1}(\mathcal{N}) \cap U \subseteq \mathbb{R}^q \times 0^{n-q}$,*
3. *$Z^{-1}(\mathcal{M}) \cap U \subseteq \mathbb{R}^p \times 0^{q-p} \times 0^{n-q}$.*

Proof. By Proposition 1 applied to a $\mathcal{C}^r$ q-chart on $\mathcal{N}$ that covers x_0, there is a $\mathcal{C}^r$ diffeomorphism $X\colon U_0 \to V_0$, where U_0, V_0 are open subsets of $\mathbb{R}^n$ and $x_0 \in V_0$, such that

$$X^{-1}(\mathcal{N}) \cap U_0 \subseteq \mathbb{R}^q \times 0^{n-q}. \tag{4}$$

We may suppose that $X(t_0, 0^{n-q}) = x_0$.

Notice that $X^{-1}(\mathcal{M}) \cap U_0$ must lie in $\mathbb{R}^q \times 0^{n-q}$. Define

$$\pi \colon \mathbb{R}^n = \mathbb{R}^q \times \mathbb{R}^{n-q} \to \mathbb{R}^q$$

to be the projection $\pi(t, t') = t$ where $t \in \mathbb{R}^q$ and $t' \in \mathbb{R}^{n-q}$. By Lemmas 1, 2, and 4, we know that $\pi \circ X^{-1}(V_0 \cap \mathcal{M})$ is a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^q$.

Let $w \colon W \to \mathbb{R}^n$ be a $\mathcal{C}^r$ p-chart on $\mathcal{M}$ that covers x_0 and assume $x_0 = w(s_0)$. Set $y = \pi \circ X^{-1} \circ w$. The appropriate domain for y is $W \cap w^{-1}(V_0)$, and

$$y(s_0) \;=\; \pi \circ X^{-1}(x_0) \;=\; \pi(t_0, 0^{n-q}) \;=\; t_0.$$

It is straightforward to check that y is a $\mathcal{C}^r$ p-chart, $y \colon W \cap w^{-1}(V_0) \to \mathbb{R}^q$, on the manifold $\pi \circ X^{-1}(V_0 \cap \mathcal{M})$ that covers the point t_0. We now apply Proposition 1 to y and obtain the following: There exist open sets U_1, V_1 in $\mathbb{R}^q$ and a map Y such that

$$\begin{aligned}
&Y \colon U_1 \to V_1 \quad \text{is a } \mathcal{C}^r \text{ diffeomorphism},\\
&Y(s_0, 0^{q-p}) \;=\; t_0,\\
&Y(s, 0^{q-p}) \;=\; y(s) \quad \text{for all } (s, 0^{q-p}) \in U_1),\\
&U_1 \cap Y^{-1}\big(\pi \circ X^{-1}(V_0 \cap \mathcal{M})\big)\\
&\qquad = \; \big\{(s, 0^{q-p}) \in U_1 \;:\; y(s) \in \pi \circ X^{-1}(V_0 \cap \mathcal{M})\big\}\,.
\end{aligned} \tag{5}$$

Let Id be the identity map on $\mathbb{R}^{n-q}$ and notice that $Y \times Id \colon U_1 \times \mathbb{R}^{n-q} \to V_1 \times \mathbb{R}^{n-q}$ is a $\mathcal{C}^r$ diffeomorphism. The point $(t_0, 0^{n-q})$ is common to $V_1 \times \mathbb{R}^{n-q}$ and U_0. Set $U_2 = (V_1 \times \mathbb{R}^{n-q}) \cap U_0$; this is a nonempty open set in $\mathbb{R}^n$. Now let $U = (Y \times Id)^{-1}(U_2)$ and $V = X(U_2)$. Of course U and V are open sets of $\mathbb{R}^n$. We now have the following diagram of maps

$$\begin{array}{ccccc}
U & \xrightarrow{Y \times Id} & U_2 & \xrightarrow{X} & V\\
{\scriptstyle\pi}\big\downarrow & & {\scriptstyle\pi}\big\downarrow & & {\scriptstyle w}\big\uparrow\\
\pi(U) & \xrightarrow{Y} & \pi(U_2) & \xleftarrow{y} & W_0
\end{array} \tag{6}$$

where $W_0 = w^{-1}(V) \subseteq W$.

Set $Z = X \circ (Y \times Id)$. Then $Z \colon U \to V$ is a $\mathcal{C}^r$ diffeomorphism. By (4) and the fact that $U \subseteq U_0$, we see that

$$Z^{-1}(\mathcal{N}) \cap U \;\subseteq\; (Y \times Id)^{-1}(\mathbb{R}^q \times 0^{n-q}) \;\subseteq \mathbb{R}^q \times 0^{n-q}.$$

This is the first of the two desired conclusions of the proposition.

Next choose $(s, s_1, s_2) \in Z^{-1}(\mathcal{M}) \cap U$ where $s \in \mathbb{R}^p$, $s_1 \in \mathbb{R}^{q-p}$, and $s_2 \in \mathbb{R}^{n-q}$. We know that

$$(s, s_1, s_2) \ \in \ Z^{-1}(\mathcal{N}) \cap U \ \subseteq \mathbb{R}^q \times 0^{n-q},$$

so $s_2 = 0^{n-q}$. Next notice that $(s, s_1, 0^{n-q}) \in U$ so that we have

$$(s, s_1) \ \in \ \pi(U) \ \subseteq \ U_1. \tag{7}$$

Now consider that

$$Z(s, s_1, 0^{n-q}) \ = \ X\big(Y(s, s_1), 0^{n-q}\big) \ \in \ \mathcal{M}$$

implies

$$(s, s_1) \ \in \ Y^{-1} \circ \pi \circ X^{-1}(\mathcal{M}).$$

We know from (6) that $X \colon U_2 \to V$, thus

$$X^{-1}(\mathcal{M}) \ = \ U_2 \cap X^{-1}(\mathcal{M}) \ \subseteq X^{-1}(V_0 \cap \mathcal{M}).$$

Therefore

$$(s, s_1) \ \in \ Y^{-1} \circ \pi \circ X^{-1}(V_0 \cap \mathcal{M}). \tag{8}$$

From (7) and (8), we have

$$(s, s_1) \ \in \ U_1 \cap \big(Y^{-1} \circ \pi \circ X^{-1}(V_0 \cap \mathcal{M})\big).$$

It then follows from the last condition of (5) that (s, s_1) must have the form $(s, 0^{q-p})$, that is, $s_1 = 0^{q-p}$.

Therefore we have shown that $Z^{-1}(\mathcal{M}) \cap U \subseteq \mathbb{R}^p \times 0^{n-p}$. $\square$

In standard manifold theory, it is crucial that for two charts x and y on a $\mathcal{C}^k$ manifold the composition $y^{-1} \circ x$ be $\mathcal{C}^k$. Here it is convenient to formulate a somewhat more complex statement.

Proposition 3. *Suppose $\mathcal{M}$ is a $\mathcal{C}^k$ p-manifold in $\mathbb{R}^n$ $(k \geq 1)$ and $x_0 \in \mathcal{M}$. Let $x \colon W_0 \to \mathbb{R}^n$ and $y \colon W_1 \to \mathbb{R}^n$ be two charts which cover x_0 and $x_0 = x(r_0) = y(s_0)$. Then there exist open sets D_0 and D_1 in $\mathbb{R}^p$ and a map $g \colon D_1 \to R^p$ such that the following hold:*

1. *$r_0 \in D_0 \subseteq W_0$ and $s_0 \in D_1 \subseteq W_1$.*

2. *$x|_{D_0}$ and $y|_{D_1}$ are charts on $\mathcal{M}$.*

3. *g is $\mathcal{C}^k$.*

4. *If $N_0 = x^{-1}(\mathcal{M}) \cap D_0$ and $N_1 = y^{-1}(\mathcal{M}) \cap D_1$, then $g\colon N_1 \to N_0$ is one-to-one, onto, and $y = x \circ g$ on N_1.*

Proof. By Proposition 1, there exist open sets U_0, U_1, V_0, V_1 in $\mathbb{R}^n$ and maps $X\colon U_0 \to V_0$ and $Y\colon U_1 \to V_1$ such that the following hold:

1. X and Y are $\mathcal{C}^k$ diffeomorphisms.

2. $(r_0, 0^{n-p}) \in U_0$ and $(s_0, 0^{n-p}) \in U_1$.

3. $X(r, 0^{n-p}) = x(r)$ for all $(r, 0^{n-p}) \in U_0$ and $Y(s, 0^{n-p}) = y(s)$ for all $(s, 0^{n-p}) \in U_1$.

4. $U_0 \cap X^{-1}(\mathcal{M}) \subseteq W_0 \times 0^{n-p}$ and $U_1 \cap Y^{-1}(\mathcal{M}) \subseteq W_1 \times 0^{n-p}$.

Notice that $x_0 = x(r_0) = X(r_0, 0^{n-p}) \in V_0$ and similarly $x_0 \in V_1$. Set

$$\begin{aligned} V &= V_0 \cap V_1, \\ U_2 &= X^{-1}(V) \subseteq U_0, \\ U_3 &= Y^{-1}(V) \subseteq U_1. \end{aligned}$$

Of course, V is an open neighborhood of x_0, and U_2 and U_3 must be open sets in $\mathbb{R}^n$, neighborhoods of $(r_0, 0^{n-p})$ and $(s_0, 0^{n-p})$ respectively. Next set

$$\begin{aligned} D_0 &= \{r \in \mathbb{R}^p \colon (r, 0^{n-p}) \in U_2\}, \\ D_1 &= \{s \in \mathbb{R}^p \colon (s, 0^{n-p}) \in U_3\}. \end{aligned}$$

D_0 and D_1 are open sets in $\mathbb{R}^p$ and are neighborhoods of r_0 and s_0 respectively. It follows that $x|_{D_0}$ and $y|_{D_1}$ are charts in $\mathbb{R}^n$ for $\mathcal{M}$ that cover x_0.

We want to show that

$$X(N_0 \times 0^{n-p}) = V \cap \mathcal{M} = Y(N_1 \times 0^{n-p}). \tag{9}$$

The proof is elementary but sufficiently complicated to warrant showing the details:

Choose $X(z) \in X(N_0 \times 0^{n-p})$. We then make the following string of deductions:

$$\text{There exists } r \in N_0 \text{ such that } z = (r, 0^{n-p}). \tag{10}$$

$r \in x^{-1}(\mathcal{M}) \cap D_0$ by the definition of N_0. (11)

By (11), there exists $x_1 \in \mathcal{M}$ such that $x(r) = x_1$. (12)

$(r, 0^{n-p}) \in U_2 = X^{-1}(V) \subseteq U_0$ by the definition of D_0 and U_2. (13)

$X(z) \in V$ and $z \in U_0$ by (10) and (13). (14)

Thus $X(z) = X(r, 0^{n-p}) = x(r)$ by the definition of U_0. (15)

By (12) and (15), $X(z) = x_1 \in \mathcal{M}$. (16)

So $X(z) \in V \cap \mathcal{M}$ by (14) and (16). (17)

We have thus established $X(N_0 \times 0^{n-p}) \subseteq V \cap \mathcal{M}$.

To establish containment in the other direction, we choose $s \in V \cap \mathcal{M}$ and write down a second string of deductions:

$s \in V_0$ by the definition of V. (18)

By definition of V_0, there exists $z \in U_0$ such that $s = X(z)$. (19)

$z = X^{-1}(s) \in X^{-1}(\mathcal{M})$ by (19) and definition of s. (20)

$z \in U_0 \cap X^{-1}(\mathcal{M})$ by (19) and (20). (21)

By definition of U_0, we have $z \in W_0 \times 0^{n-p}$. (22)

So $z = (r, 0^{n-p})$ for some $r \in W_0$. (23)

Then $(r, 0^{n-p}) = z \in U_0$ by (21). (24)

By (19), (21), (24), and definition of X, $s = X(z) = x(r)$. (25)

$r \in x^{-1}(\mathcal{M})$ by (25) and definition of s. (26)

By definition of s and (25), $z = X^{-1}(s) \in X^{-1}(V) = U_2$. (27)

From (24) and (27), $(r, 0^{n-p}) = z \in U_2$. (28)

So $r \in D_0$ by (28) and definition of D_0. (29)

$r \in x^{-1}(\mathcal{M}) \cap D_0 = N_0$ by (26), (29), and definition of N_0. (30)

So $z = (r, 0^{n-p}) \in N_0 \times 0^{n-p}$ by (30). (31)

It follows from (31) that $s = X(z) \in X(N_0 \times 0^{n-p})$. (32)

We thus have $V \cap \mathcal{M} \subseteq X(N_0 \times 0^{n-p})$.

The other half of (9) is shown similarly.

We now introduce the maps

$$\pi \colon \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p} \to \mathbb{R}^p \text{ where } \pi(r, r') = r,$$
$$J \colon \mathbb{R}^p \to \mathbb{R}^p \times 0^{n-p} \subseteq \mathbb{R}^n \text{ where } J(r) = (r, 0^{n-p})$$

and note that

$$x = X \circ J \text{ on } D_0, \text{ and } y = Y \circ J \text{ on } D_1. \tag{33}$$

We now have the following maps:

$$D_1 \xrightarrow{J} D_1 \times 0^{n-p} \xrightarrow{Y} Y(D_1 \times 0^{n-p}) \subseteq Y(U_3) = V,$$
$$V \xrightarrow{X^{-1}} U_2 \xrightarrow{\pi} \mathbb{R}^p.$$

We see that we can define $g\colon D_1 \to \mathbb{R}^p$ by

$$g \;=\; \pi \circ X^{-1} \circ Y \circ J$$

and that this will be a $\mathcal{C}^k$ map. Consider the restriction of g to N_1,

$$N_1 \xrightarrow{J} N_1 \times 0^{n-p} \xrightarrow{Y} V \cap \mathcal{M} \xrightarrow{X^{-1}} N_0 \times 0^{n-p} \xrightarrow{\pi} N_0$$

and notice that this is a composition of one-to-one, onto maps. (We have used (9).) Hence $g\colon N_1 \to N_0$ is a one-to-one, onto map.

Finally let $s \in N_1$. Then

$$x \circ g(s) \;=\; \big(X \circ J \circ \pi \circ X^{-1} \circ Y \circ J\big)(s) \;=\; y(s). \qquad \square$$

4 Tangent spaces

We now define tangent spaces via an appeal to charts, so our first concern is to show this definition does not depend on our choice of a chart.

Let $\mathcal{M}$ be a $\mathcal{C}^k$ $(k \geq 1)$ p-manifold in $\mathbb{R}^n$ and suppose that $x_0 \in \mathcal{M}$. If $x\colon U \to \mathbb{R}^n$ is a chart on $\mathcal{M}$ that covers x_0 and $x_0 = x(r_0)$, then we define

$$T\mathcal{M}(x_0, x) \;=\; \{x'(r_0)\, v \;:\; v \in \mathbb{R}^p\}$$

where $x'(r_0)\, v$ indicates the linear transformation $x'(r_0)$ operating on the vector v.

Proposition 4. *If $x\colon U \to \mathbb{R}^n$ and $y\colon V \to \mathbb{R}^n$ are charts on $\mathcal{M}$ that cover x_0, then $T\mathcal{M}(x_0, x) = T\mathcal{M}(x_0, y)$.*

Proof. We may assume $x_0 = x(r_0) = y(s_0)$. By Proposition 3, there exist open sets D_0 and D_1 in $\mathbb{R}^p$ and a $\mathcal{C}^k$ map $g\colon D_1 \to \mathbb{R}^p$ such that

1. $g\colon N_1 \to N_0$ in one-to-one and onto where $N_0 = x^{-1}(\mathcal{M}) \cap D_0$ and $N_1 = y^{-1}(\mathcal{M}) \cap D_1$,
2. $r_0 \in D_0 \subseteq U$ and $s_0 \in D_1 \subseteq V$,
3. $y = x \circ g$ on N_1.

By Definition 3, there exists a nondegenerate p-simplex P such that $s_0 \in P \subseteq y^{-1}(\mathcal{M}) \cap V$. Since $s_0 \in D_1$ and D_1 is open, we may, if we wish, choose P so small that $P \subseteq N_1$.

We can choose points $s_1, \ldots, s_p$ in P such that if v_i is the vector from s_0 to s_i, then $\{v_1, \ldots, v_p\}$ is a linearly independent set. Since $y'(s_0)$ is one-to-one, it follows that $\{y'(s_0)v_i\}_{i=1}^p$ is a basis for $T\mathcal{M}(x_0, y)$. From $y = x \circ g$ on N_1, we deduce that

$$y'(s_0)\, v_i \;=\; x'(r_0)\, g'(s_0)\, v_i \;=\; x'(r_0)\, u_i$$

where $u_i = g'(s_0)\, v_i$. This means that each $y'(s_0)\, v_i$ lies in $T\mathcal{M}(x_0, x)$, so $T\mathcal{M}(x_0, y) \subseteq T\mathcal{M}(x_0, x)$. Containment in the other direction holds by symmetry, so the result is established. □

Remark 5. Notice that the proof Proposition 4 does not require us to pay attention to whether x_0 is an interior point or boundary point of $\mathcal{M}$.

We now abandon the $T\mathcal{M}(x_0, x)$ notation for one that depends only on $\mathcal{M}$ and x_0:

Definition 4. By the *tangent space to the* $\mathcal{C}^k$ *p-manifold* $\mathcal{M}$ $(k \geq 1)$ *at the point* $x_0 \in \mathcal{M}$, we mean

$$T_{x_0}\mathcal{M} \;=\; \{x'(r_0)\, v \;:\; v \in \mathbb{R}^p\}$$

where $x\colon U \to \mathbb{R}^n$ is any chart that covers x_0 and $x_0 = x(r_0)$.

Proposition 5. *If* $\mathcal{M}$ *is a* $\mathcal{C}^1$ *p-manifold and* $x_0 \in \mathcal{M}$, *then* $T_{x_0}\mathcal{M}$ *is a vector space of dimension p.*

Proof. This is clearly a vector space, so we check only the dimension. Let $x\colon U \to \mathbb{R}^n$ be a p-chart that covers x_0. U must be an open subset of $\mathbb{R}^p$ and there must exist $t_0 \in U$ such that $x_0 = x(t_0)$. Set $u_i = \big[x'(t_0)\big]e_i$ for $i = 1, \ldots, p$. If u is a tangent vector to $\mathcal{M}$ at x_0, that is, $u \in T_{x_0}\mathcal{M}$, then by Definition 4, we have $u = \big[x'(t_0)\big]v$ for some $v \in \mathbb{R}^p$. We can write v as a linear combination of the standard basis vectors of $\mathbb{R}^p$, $v = \sum_{i=1}^p \lambda_i e_i$, so we must have $u = \sum_{i=1}^p \lambda_i u_i$. That is, every element of $T_{x_0}\mathcal{M}$ is a linear combination of the u_i vectors. We also know from Definition 2 that $x'(t_0)$ has rank p, so $\{u_i\}_{i=1}^p$ must be a linearly independent set of vectors, a basis for $T_{x_0}\mathcal{M}$. $\square$

Another way to try to define tangent vectors to a manifold at a point is by means of curves in the manifold running through the point.

Definition 5. By a curve in $\mathbb{R}^n$, we mean a continuous map $c\colon J \to \mathbb{R}^n$ where J is a nondegenerate interval in $\mathbb{R}$. We say the curve is $\mathcal{C}^k$ provided $\big(d^k c/d\tau^k\big)(\tau_0)$ exists at every point $\tau_0 \in J$. In the case where J is a closed or half-closed interval, then we calculate derivatives at endpoints as one-sided limits. Thus, for example, if $J = (\alpha, \beta]$, then

$$c'(\beta) \;=\; \frac{dc}{d\tau}(\beta) \;=\; \lim_{\tau\to\beta^-} \frac{c(\beta) - c(\tau)}{\tau}.$$

We say c passes through x_0 provided there is some τ_0 such that $x_0 = c(\tau_0)$. A curve c lies in a manifold $\mathcal{M}$ provided $c(\tau) \in \mathcal{M}$ for all τ.

Definition 6. If $\mathcal{M}$ is a $\mathcal{C}^k$ manifold $(1 \le k)$ and x_0 is a point in $\mathcal{M}$, then we say that v is a *tangent vector to* $\mathcal{M}$ *at* x_0 *generated by a curve in the manifold* provided there is a $\mathcal{C}^1$ curve $c\colon J \to \mathcal{M}$ satisfying $x_0 = c(\tau_0)$ and $v = c'(\tau_0)$ for some $\tau_0 \in J$.

This way of defining tangent vectors is, of course, equivalent to ours except possibly at boundary points as we explain below.

Proposition 6. *Let $\mathcal{M}$ be a $\mathcal{C}^1$ p-manifold in $\mathbb{R}^n$, let x_0 be a point in $\mathcal{M}$, and let $x\colon U \to \mathbb{R}^n$ be a chart covering x_0. Then v is a tangent vector to $\mathcal{M}$ at x_0 if and only if there is a $\mathcal{C}^1$ curve $c\colon (-\epsilon, \epsilon) \to x(U)$ such that $c(0) = x_0$ and $c'(0) = v$.*

The proof follows easily from finding $r_0 \in U$ such that $x(r_0) = x_0$ and $u \in \mathbb{R}^p$ such that $x'(r_0)\,u = v$ and then setting $c(\lambda) = x(r_0 + \lambda u)$.

Remark 6. If x_0 is an interior point of $\mathcal{M}$, then by choosing ϵ sufficiently small, we may assume the curve lies completely in $\mathcal{M}$, not just in $x(U)$. However this may not be true if x_0 is a boundary point. For example, if x_0 is a vertex of a 2-simplex $\mathcal{M}$ in $\mathbb{R}^2$ and v is a vector that points from x_0 out of the 2-simplex as in Figure 3, then there may not be any curve at all in $\mathcal{M}$ that generates v. In the next proposition, we show a sense in which every

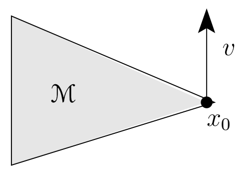

Figure 3: A tangent vector that does not correspond to a curve in $\mathcal{M}$

tangent vector to a manifold is generated by curves in the manifold, even in the case of boundary points.

Proposition 7. *Let $\mathcal{M}$ be a $\mathcal{C}^1$ p-manifold in $\mathbb{R}^n$, let x_0 be a point in $\mathcal{M}$, and let $x\colon U \to \mathbb{R}^n$ be a chart covering x_0. If u is a tangent vector to $\mathcal{M}$ at x_0, then there exist tangent vectors $u_1, \ldots, u_p$ to $\mathcal{M}$ at x_0 and scalars $\lambda_1, \ldots, \lambda_p$ such that each u_i is generated by a $\mathcal{C}^1$ curve in $\mathcal{M}$ and $u = \lambda_1 u_1 + \cdots + \lambda_p u_p$.*

Proof. We need consider only the case where x_0 is a boundary point of $\mathcal{M}$. There must exist $t_0 \in U$ such that $x_0 = x(t_0)$ and a vector $v \in \mathbb{R}^p$ such that $u = [x'(t_0)]v$. By Definition 3, there exists a nondegenerate p-simplex P in $\mathbb{R}^p$ such that $t_0 \in P \subseteq x^{-1}(\mathcal{M}) \cap U$. There exist linearly independent vectors $v_1, \ldots, v_p$ in $\mathbb{R}^p$ with basepoint t_0 such that for all i, we have $t_0 + \lambda v_i \in P$ when λ is sufficiently small. There must exist scalars $\lambda_1, \ldots \lambda_p$ such that $v = \sum_{i=1}^{p} \lambda_i v_i$. (See Figure 4.) Set $u_i = [x'(t_0)]v_i$ for $i = 1, \ldots, p$. Then $u = \sum_{i=1}^{p} \lambda_i u_i$. Next let us set $c_i(\lambda) = x(t_0 + \lambda v_i)$. This defines a $\mathcal{C}^1$ curve in $\mathcal{M}$ since $t_0 + \lambda v_i \in P$ for λ sufficiently small. Notice that $c_i(0) = x_0$ and

$$c_i{}'(0) \;=\; [x'(t_0)]v_i \;=\; u_i.$$

Thus each u_i is generated by a curve in $\mathcal{M}$. □

We note that maps between manifolds induce a mapping of tangent vectors to tangent vectors.

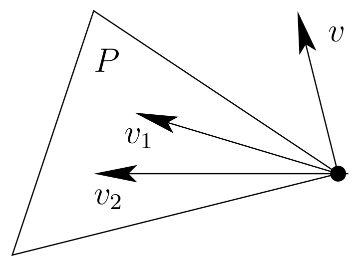

Figure 4: v is a linear combination of vectors in the simplex

Proposition 8. *Suppose that $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{C}^1$ manifolds in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively and $x_0 \in \mathcal{M}$. If $f\colon \mathcal{M} \to \mathcal{N}$ is $\mathcal{C}^1$ and u is a tangent vector to $\mathcal{M}$ at x_0, then $\partial_u f(x_0) = \big[f'(x_0)\big]u$ is a tangent vector to $\mathcal{N}$ at $y_0 = f(x_0)$.*

Proof. Supposing $\mathcal{M}$ to be a p-manifold, we know from Proposition 7 that we can find tangent vectors $u_1, \ldots, u_p$ to $\mathcal{M}$ at x_0 and scalars $\lambda_1, \ldots, \lambda_p$ such that each u_i is generated by a $\mathcal{C}^1$ curve c_i lying in $\mathcal{M}$ and $u = \sum_{i=1}^p \lambda_i u_i$. Without loss of generality, we may assume each c_i defined on an interval in $\mathbb{R}$ that contains 0 and that $x_0 = c_i(0)$ and $u_i = c_i'(0)$. We note that for each i, $f \circ c_i$ is a $\mathcal{C}^1$ curve lying in $\mathcal{N}$ and that $y_0 = f(x_0) = \big(f \circ c_i\big)(0)$. It follows from Proposition 6 that the vector

$$v_i \stackrel{\text{def.}}{=} (f \circ c_i)'(0) \;=\; \big[f'(x_0)\big]\, c_i'(0) \;=\; \big[f'(x_0)\big] u_i$$

is a tangent vector to $\mathcal{N}$ at y_0. We then conclude that

$$\big[f'(x_0)\big]u \;=\; \big[f'(x_0)\big]\Big(\sum_{i=1}^{p} \lambda_i u_i\Big) \;=\; \sum_{i=1}^{p} \lambda_i v_i$$

must be a tangent vector to $\mathcal{N}$ at y_0. □

Remark 7. When working with a chart x on a manifold where the associated coordinates are $(\tau_1, \ldots, \tau_p)$, we may prefer to use the suggestive symbolism

$$\frac{\partial x}{\partial \tau_i}(x_0) \stackrel{\text{def.}}{=} \partial_{e_i} x(t_0). \tag{34}$$

Each $\big(\partial x/\partial \tau_i\big)(x_0)$ is of course a tangent vector to $\mathcal{M}$ at the point x_0 corresponding to the curve in $\mathcal{M}$ along which τ_i increases while all the other coordinates remain constant. The set $\{\big(\partial x/\partial \tau_i\big)(x_0)\}_{i=1}^p$ is a basis for the tangent space to $\mathcal{M}$ at x_0.

5 Two definitions of directional derivatives

Before we start on this section, we remind readers, as was pointed out in Section 2.2, it is sufficient to suppose each function f is real-valued. Yet the results also apply to vector-valued and multivector-valued functions.

We know what directional derivatives $\partial_u f$ are in $\mathbb{R}^n$ from Equation (1). However if f is defined on a manifold $\mathcal{M}$ in $\mathbb{R}^m$, then in works such as [9], [17], or [18], directional derivatives are defined by constructing curves in $\mathcal{M}$ and making use only of points in the manifold. We make this notion explicit:

Let $\mathcal{M}$ be a manifold in some $\mathbb{R}^m$ that is at least $\mathcal{C}^1$, let $x_0 \in \mathcal{M}$, and let u be a tangent vector to $\mathcal{M}$ at x_0. If $A \subseteq \mathbb{R}$ and $y\colon A \to \mathcal{M}$ is a $\mathcal{C}^1$ curve in $\mathcal{M}$ such that $y(\tau_0) = x_0$ and

$$\lim_{\tau \to \tau_0} \frac{y(\tau) - y(\tau_0)}{\tau - \tau_0} \;=\; u,$$

then we set

$$\partial_{\mathcal{M},u} f(x_0) \;\overset{\text{def.}}{=}\; \lim_{\tau \to \tau_0} \frac{\big(f \circ y\big)(\tau) - \big(f \circ y\big)(\tau_0)}{\tau - \tau_0}. \tag{35}$$

Of course there may be more than one way to choose the curve y, but it follows from our next consideration that this is unimportant.

The near-equivalence of the $\partial_u f(x_0)$ and $\partial_{\mathcal{M},u} f(x_0)$ is not surprising and hinges on the fact that if y defines a curve on $\mathcal{M}$ passing through $x_0 = y(\tau_0)$ and $u = y'(t_0)$ is a tangent vector to $\mathcal{M}$, then

$$\begin{aligned} \partial_{\mathcal{M},u} f(x_0) \;&=\; \big(f \circ y\big)'(\tau_0) \;=\; f'(x_0)\, y'(\tau_0) \\ &= f'(x_0)\, u \;=\; \partial_u f(x_0) \end{aligned}$$

where we appealed to Equation (2). We then have the following:

Proposition 9. *Let $\mathcal{M}$ be a $\mathcal{C}^1$ manifold in $\mathbb{R}^m$. Suppose that x_0 is a point in $\mathcal{M}$, u is a vector in $\mathbb{R}^m$, and f is a $\mathcal{C}^1$ function on $\mathcal{M}$. If u is tangent to $\mathcal{M}$ at x_0 and the directional derivative in the sense of (35) exists, then it also exists in the sense of (1) and*

$$\partial_u f(x_0) \;=\; \partial_{\mathcal{M},u} f(x_0).$$

At interior points of $\mathcal{M}$, the definitions of Equations (1) and (35) are equivalent. However there is a problem with $\partial_u f(x_0)$ in case $x_0 \in \partial\mathcal{M}$, that is, in case x_0 is a boundary point of $\mathcal{M}$. Recall that if a function is $\mathcal{C}^k$ on $\mathcal{M}$ ($k \geq 1$), then it must possess local $\mathcal{C}^k$ extensions to points not in $\mathcal{M}$. The extension of f beyond $\mathcal{M}$ is not in general unique. So when evaluating

$$\partial_u f(x_0) \ = \ \lim_{\lambda \to 0} \frac{1}{\lambda} \big(f(x_0 + \lambda u) - f(x_0) \big),$$

we may have different values of $f(x_0 + \lambda u)$ depending on which extension of f we appeal to. It turns out that if u is a *tangent* vector to $\mathcal{M}$ there is no problem:

Proposition 10. *Let $\mathcal{M}$ be a $\mathcal{C}^1$ p-manifold in $\mathbb{R}^m$. Suppose that $x_0 \in \mathcal{M}$ and u is a tangent vector to $\mathcal{M}$ at x_0. Let f_1 and f_2 be two $\mathcal{C}^1$ functions on $\mathcal{M}$ such that their restrictions to $\mathcal{M}$ agree, that is, $f_1|_{\mathcal{M}} = f_2|_{\mathcal{M}}$. Then*

$$\partial_u f_1(x_0) \ = \ \partial_u f_2(x_0).$$

Proof. By Proposition 7, there must exist tangent vectors $u_1, \ldots, u_p$ to $\mathcal{M}$ at x_0 and scalars $\lambda_1, \ldots, \lambda_p$ such that $u = \lambda_1 u_1 + \cdots + \lambda_p u_p$ and such that each u_i is generated by a curve c_i lying in $\mathcal{M}$. Since $\partial_u f_j(x_0)$ is linear in u for $j = 1, 2$, it follows that

$$\partial_u f_j(x_0) \ = \ \sum_{i=1}^{p} \lambda_i \left(\partial_{u_i} f_j(x_0) \right).$$

Then by Proposition 9, we have $\big(\partial_{u_i} f_j\big)(x_0) = \big(\partial_{\mathcal{M},u_i} f_j\big)(x_0)$. Since $f_1 = f_2$ at all points of $\mathcal{M}$ and the curves c_i lie in $\mathcal{M}$, it follows that $\big(\partial_{\mathcal{M},u_i} f_1\big)(x_0) = \big(\partial_{\mathcal{M},u_i} f_2\big)(x_0)$. Therefore,

$$\begin{aligned} \partial_u f_1(x_0) \ &= \ \sum_{i=1}^{p} \lambda_i \partial_{u_i} f_1(x_0) \ = \ \sum_{i=1}^{p} \lambda_i \partial_{\mathcal{M},u_i} f_1(x_0) \\ &= \ \sum_{i=1}^{p} \lambda_i \partial_{\mathcal{M},u_i} f_2(x_0) \ = \ \sum_{i=1}^{p} \lambda_i \partial_{u_i} f_2(x_0) \ = \ \partial_u f_2(x_0). \qquad \square \end{aligned}$$

Because of these results, we shall have no further use for the notation $\partial_{\mathcal{M},u} f$ and refer only to $\partial_u f$. However there is the annoying fact that if u is not tangent to $\mathcal{M}$ at x_0, then we can get different answers for $\partial_u f(x_0)$ depending on which extension of f from $\mathcal{M}$ we appeal to. We shall see a way to deal with this in the section on the *normal identity*.

6 When is a map between manifolds $\mathcal{C}^q$?

We know what it means to say $f\colon U \to \mathbb{R}^n$ is $\mathcal{C}^q$ when U is an open subset of $\mathbb{R}^m$. We want to extend that definition to the case $f\colon \mathcal{M} \to \mathcal{N}$ where $\mathcal{M}$ and $\mathcal{N}$ are manifolds.

We also need to take into account the fact that there is already a way of defining this notion for standard manifolds, a definition which does not require the manifolds to be embedded in some $\mathbb{R}^n$. Therefore we give two definitions for $f\colon \mathcal{M} \to \mathcal{N}$ being $\mathcal{C}^q$, one corresponding to what is usually done for standard manifolds, the other making essential use of the fact that the manifolds are embedded in an ambient $\mathbb{R}^n$, and we show the equivalence of these definitions.

We first give a version of the usual definition of a function between standard manifolds being $\mathcal{C}^q$. Since our manifolds-with-corners are a bit more general than standard manifolds and our definition of charts requires an ambient $\mathbb{R}^n$, this is a slight modification of the standard definition which requires no ambient $\mathbb{R}^n$. It is, however, in essence, the same.

Definition 7. Suppose that $f\colon \mathcal{M} \to \mathcal{N}$ where $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{C}^k$ manifolds lying in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. We say that *f is $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^q$ at $x_0 \in \mathcal{M}$* (where $0 \le q \le k$) provided the following is true: For all charts $x\colon U \to \mathbb{R}^m$ and $y\colon V \to \mathbb{R}^n$ on $\mathcal{M}$ and $\mathcal{N}$ respectively such that x covers x_0 and y covers $f(x_0)$, there exists U_0 and $g\colon U_0 \to \mathbb{R}^m$ such that

1. U_0 is an open subset of U,
2. g is $\mathcal{C}^q$,
3. $g(t) = \big(y^{-1} \circ f \circ x\big)(t)$ for all $t \in U_0$.

We say *f is $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^q$ on $\mathcal{M}$* provided it is $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^q$ at all $x_0 \in \mathcal{M}$.

We contrast Definition 7 with a definition of a map $f\colon A \to B$ between arbitrary sets being $\mathcal{C}^q$ which is an extension of what this would be if f were mapping between open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$:

Definition 8. Suppose that $f\colon A \to B$ where A and B are sets lying in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. We say that *f is $\mathcal{C}^q$ at $x_0 \in A$* (where $0 \le q$) provided the following is true: There exists an open neighborhood U of x_0 in $\mathbb{R}^m$ and a function $F\colon U \to \mathbb{R}^n$ such that F is $\mathcal{C}^q$ on U and $F(t) = f(t)$ for all $t \in U \cap A$. We say that *f is $\mathcal{C}^q$ on A* provided f is $\mathcal{C}^q$ at all points $x_0 \in A$.

We are mainly concerned to apply this definition in the case where A and B are manifolds.

Example 1. Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{C}^k$ manifolds lying in $\mathbb{R}^n$ and $\mathcal{M} \subseteq \mathcal{N}$. For example, $\mathcal{M}$ might be an arc in $\mathcal{N}$. Let $I\colon \mathcal{M} \to \mathcal{N}$ be the inclusion map $I(x) = x$. We extend I to the identity map $I\colon \mathbb{R}^n \to \mathbb{R}^n$ and see that by Definition 8 the inclusion of a $\mathcal{C}^k$ submanifold into a $\mathcal{C}^k$ manifold must be a $\mathcal{C}^k$ map.

We want to show Definitions 7 and 8 are equivalent on manifolds. Here is a convenient preliminary step:

Proposition 11. *Let $\mathcal{M}$ be a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^n$, and suppose that ϕ is a real-valued function that is $(\mathcal{M}, \mathbb{R})$-$\mathcal{C}^q$ on $\mathcal{M}$ (where $1 \le q \le r$). Then for every $x_0 \in \mathcal{M}$, there exist $W \subseteq \mathbb{R}^n$ and $\Phi : W \to \mathbb{R}$ such that*

1. *W is an open neighborhood of x_0 in $\mathbb{R}^n$,*
2. *Φ is $\mathcal{C}^q$ on W,*
3. *$\Phi(z) = \phi(z)$ for all $z \in \mathcal{M} \cap W$.*

Proof. We know if we choose $\mathcal{C}^q$ charts $x\colon U_0 \to \mathbb{R}^n$ and $y\colon V_0 \to \mathbb{R}$ on $\mathcal{M}$ and $\mathbb{R}$ respectively such that x covers x_0, then $y^{-1} \circ \phi \circ x$ must be $\mathcal{C}^q$ on r_0 where $x_0 = x(r_0)$. Since y is a chart that maps between subsets of $\mathbb{R}$, we know that $y' \neq 0$ and hence y^{-1} is $\mathcal{C}^q$. It follows that $\phi \circ x$ must be $\mathcal{C}^q$ at r_0. We may, without loss of generality, suppose that $\phi \circ x$ is $\mathcal{C}^q$ on U_0.

We now appeal to Proposition 1: There exist open subsets U and V of $\mathbb{R}^n$ and a map $X : U \to V$ such that

1. X is a $\mathcal{C}^r$ diffeomorphism of U to V,
2. there exists $t_0 \in \operatorname{dom} x$ such that $x_0 = X(t_0, 0^{n-p})$,
3. $X(t, 0^{n-p}) = x(t)$ for all $(t, 0^{n-p}) \in U$,
4. for all $(t, t') \in U$, where $t \in \mathbb{R}^p$ and $t' \in \mathbb{R}^{n-p}$, if $X(t, t') \in \mathcal{M}$, then $t' = 0^{n-p}$.

Let $\pi : \mathbb{R}^n \to \mathbb{R}^p$ be the projection $\pi(t, t') = t$. We notice that $\big(U_0 \times \mathbb{R}^{n-p}\big) \cap U$ is an open neighborhood of $(t_0, 0^{n-p})$ in $\mathbb{R}^n$. We set

$$W \;=\; X\big[\big(U_0 \times \mathbb{R}^{n-p}\big) \cap U\big],$$

and we know this is an open neighborhood of x_0 in $\mathbb{R}^n$. Now define $\Phi : W \to \mathbb{R}$ by

$$\Phi \;=\; \phi \circ x \circ \pi \circ X^{-1}.$$

This is $\mathcal{C}^q$ since π and X^{-1} are $\mathcal{C}^\infty$ and $\mathcal{C}^r$ respectively.

We need only show that Φ is a (local) $\mathcal{C}^q$ extension of ϕ. Choose $z \in \mathcal{M} \cap W$. There exists a unique $(t, t') \in U$ such that $t \in U_0$, $t' \in \mathbb{R}^{n-p}$, and $z = X(t, t')$. Since $z \in \mathcal{M}$, we must have $t' = 0^{n-p}$; thus $z = X(t, 0^{n-p}) = x(t)$. Then

$$\begin{aligned}\Phi(z) \;&=\; \phi \circ x \circ \pi \circ X^{-1}(z)\\ &= \phi \circ x \circ \pi(t, 0^{n-p})\\ &= \phi \circ x(t)\\ &= \phi(z).\end{aligned}$$

□

Now here is the equivalence of Definitions 7 and 8:

Proposition 12. *Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{C}^k$ p- and q-manifolds in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. If we have a map $f \colon \mathcal{M} \to \mathcal{N}$, then f is $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^r$ at $x_0 \in \mathcal{M}$ (where $1 \le r \le k$) if and only if f is $\mathcal{C}^r$ at x_0.*

Proof. Suppose that f is $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^r$ at x_0.

Let $x \colon U \to \mathbb{R}^m$ and $y \colon V \to \mathbb{R}^n$ be charts on $\mathcal{M}$ and $\mathcal{N}$ respectively such that $x_0 = x(t_0)$ for $t_0 \in U$. We know that y and $y^{-1} \circ f \circ x$ are $\mathcal{C}^r$, so we may assume $f \circ x = y \circ y^{-1} \circ f \circ x$ is $\mathcal{C}^r$ on an open neighborhood U' of t_0 where $U' \subseteq U$.

We can write $f = \sum_{i=1}^n \phi_i e_i$ where $\{e_i\}_{i=1}^n$ is the standard basis on $\mathbb{R}^n$ and each ϕ_i is a real-valued function on $\mathcal{M}$. Since $f \circ x = \sum_i (\phi_i \circ x)\, e_i$, we can assume each $\phi_i \circ x$ is a real-valued $\mathcal{C}^r$ function on U'. Now let $z \colon J \to J'$ be a $\mathcal{C}^r$ chart on $\mathbb{R}$ where J and J' are open subsets of $\mathbb{R}$ and z covers $\phi_i(x_0)$. We know from the definition of chart that $z' \neq 0$, so z^{-1} is also $\mathcal{C}^r$. Hence $z^{-1} \circ \phi_i \circ x$ is $\mathcal{C}^r$, and it follows that ϕ_i is $(\mathcal{M}, \mathbb{R})$-$\mathcal{C}^r$.

By Proposition 11, there exists in $\mathbb{R}^m$ an open neighborhood W of x_0 such that for all i there is a real-valued $\mathcal{C}^r$ function Φ_i defined on W having the property that $\Phi_i(t) = \phi_i(t)$ for all $t \in \mathcal{M} \cap W$. If we set $F = \sum_{i=1}^n \Phi_i e_i$, it follows that F is a $\mathcal{C}^r$ function on W such that $F(t) = f(t)$ for all $t \in \mathcal{M} \cap W$. Therefore f is $\mathcal{C}^r$ at x_0.

Now assume that f is $\mathcal{C}^r$ at x_0.

Let $x\colon U \to \mathbb{R}^m$ and $y\colon V \to \mathbb{R}^n$ be $\mathcal{C}^k$ charts on $\mathcal{M}$ and $\mathcal{N}$ respectively such that x covers x_0. We must show that $y^{-1} \circ f \circ x$ is $\mathcal{C}^r$ on some open neighborhood of t_0 where $x_0 = x(t_0)$.

We may assume that $y_0 = f(x_0)$ and that $y_0 = y(s_0)$ for $s_0 \in V$. Since f is $\mathcal{C}^r$ at x_0, we know there exist an open neighborhood W of x_0 in $\mathbb{R}^m$ and a $\mathcal{C}^r$ function $F\colon W \to \mathbb{R}^n$ such that $F(z) = f(z)$ for all $z \in \mathcal{M} \cap W$.

We now appeal to Proposition 1: There exist open sets U_0, U_1 in $\mathbb{R}^m$, open sets V_0, V_1 in $\mathbb{R}^n$, and maps $X : U_0 \to U_1$ and $Y : V_0 \to V_1$ such that

1. X and Y are $\mathcal{C}^k$ diffeomorphisms,
2. $(t_0, 0^{m-p}) \in U_0$ and $(s_0, 0^{n-q}) \in V_0$,
3. $X(t, 0^{m-p}) = x(t)$ for all $(t, 0^{m-p}) \in U_0$ and $Y(s, 0^{n-q}) = y(s)$ for all $(s, 0^{n-q}) \in V_0$,
4. $U_0 \cap X^{-1}(\mathcal{M}) \subseteq U \times 0^{m-p}$ and $V_0 \cap Y^{-1}(\mathcal{N}) \subseteq V \times 0^{n-q}$.

We may assume W and U_0 chosen so "small" that $F(W) \subseteq V_1$ and $X(U_0) = U_1 \subseteq W$.

Now set
$$U' \;=\; \{t \in \mathbb{R}^p \;:\; (t, 0^{m-p}) \in U_0\}.$$
This is an open neighborhood of t_0 in $\mathbb{R}^p$. Next introduce the $\mathcal{C}^\infty$ maps

$$\pi\colon \mathbb{R}^n = \mathbb{R}^q \times \mathbb{R}^{n-q} \to \mathbb{R}^q \quad \text{and} \quad J\colon \mathbb{R}^p \to \mathbb{R}^p \times \mathbb{R}^{m-p} = \mathbb{R}^m$$

where π is the projection onto the factor $\mathbb{R}^q$ and $J(t) = (t, 0^{m-p})$. We then define
$$g \;=\; \pi \circ Y^{-1} \circ F \circ X \circ J.$$
Clearly g is $\mathcal{C}^r$, and by checking domains and ranges, we see that $g\colon U' \to \mathbb{R}^q$. We claim that g is the desired extension of $y^{-1} \circ f \circ x$.

To see this, choose $t \in U' \cap x^{-1}(\mathcal{M})$ and note that

$$F \circ X \circ J(t) \;=\; f \circ x(t) \;\in\; V_1 \cap \mathcal{N}.$$

Then we have

$$Y^{-1} \circ f \circ x(t) \;\in\; V_0 \cap Y^{-1}(\mathcal{N}) \;\subseteq\; V \times 0^{n-q}$$

so that

$$Y^{-1} \circ f \circ x(t) \;=\; (s, 0^{n-q}) \text{ for some } (s, 0^{n-q}) \;\in\; V_0 \cap Y^{-1}(\mathcal{N}).$$

It follows that

$$f \circ x(t) \;=\; Y(s, 0^{n-q}) \;=\; y(s).$$

We see from these results that

$$\pi \circ Y^{-1} \circ f \circ x(t) \;=\; s \;=\; y^{-1} \circ f \circ x(t).$$

Thus

$$g(t) \;=\; y^{-1} \circ f \circ x(t) \quad \text{for all } t \in U' \cap x^{-1}(\mathcal{M}),$$

and hence f is $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^r$ at x_0. □

Because of this result, we now abandon the notation $(\mathcal{M}, \mathcal{N})$-$\mathcal{C}^q$ and simply talk of $f\colon \mathcal{M} \to \mathcal{N}$ being $\mathcal{C}^q$.

7 Diffeomorphisms and charts

Definition 9. Suppose that A and B are sets in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. We say that $f\colon A \to B$ is a $\mathcal{C}^k$ *diffeomorphism* if and only if f is one-to-one, onto, and both f and f^{-1} are $\mathcal{C}^k$.

Again we are primarily interested in the case where A and B are manifolds.

Example 2. Suppose ϕ is a real-valued, $\mathcal{C}^k$ function on $\mathcal{M} = \mathbb{R}^m$ and $f\colon \mathbb{R}^m \to \mathbb{R}^{m+1}$ is defined by $f(x) = \big(x, \phi(x)\big)$, that is,

$$f(\chi_1, \ldots, \chi_m) \;=\; \big(\chi_1, \ldots, \chi_m, \phi(\chi_1, \ldots, \chi_m)\big).$$

Set

$$\mathcal{N} \;=\; f(\mathbb{R}^m) \;=\; \big\{\big(x, \phi(x)\big) \,:\, x \in \mathbb{R}^m\big\},$$

that is, $\mathcal{N}$ is the graph of ϕ. Clearly $f\colon \mathcal{M} \to \mathcal{N}$ is one-to-one, onto, and $\mathcal{C}^k$. Since $f^{-1}\big(x, \phi(x)\big) = x$, we see that f^{-1} extends to the projection map $\pi\colon \mathbb{R}^{m+1} \to \mathbb{R}^m$ defined by

$$\pi(\chi_1, \ldots, \chi_m, \chi_{m+1}) \;=\; (\chi_1, \ldots, \chi_m)$$

which is $\mathcal{C}^\infty$ on $\mathbb{R}^{m+1}$; thus f^{-1} is $\mathcal{C}^k$. Therefore f is a $\mathcal{C}^k$ diffeomorphism of $\mathcal{M}$ to $\mathcal{N}$.

Remark 8. When we talk about f^{-1} being $\mathcal{C}^k$, we mean some (local) extension G of f^{-1} is $\mathcal{C}^k$, and it often happens, as in Example 2, that G cannot possibly be one-to-one; therefore it is not a true inverse of any function. To save ourselves introducing extra functions, we shall usually not mention G. We shall, instead, abuse notation and keep talking about f^{-1} and $(f^{-1})'$ even when, in the back of our mind, we mean the extended function G.

Example 3. If $f\colon \mathcal{M} \to f(\mathcal{M})$ is one-to-one and $\mathcal{C}^k$, this does not insure that it is a $\mathcal{C}^k$ diffeomorphism. To see this, let $\mathcal{M}$ be the half-open interval $[0,1)$ in $\mathbb{R}$ and let $f\colon \mathcal{M} \to \mathbb{R}^2$ be

$$f(\theta) \;=\; \cos(2\pi\theta)e_1 + \sin(2\pi\theta)e_2.$$

Then f maps $\mathcal{M}$ to a unit circle in $\mathbb{R}^2$ in a one-to-one $\mathcal{C}^\infty$ manner, however f^{-1} is discontinuous at the point $(1,0)$.

Proposition 13. *Every $\mathcal{C}^r$ chart is locally a $\mathcal{C}^r$ diffeomorphism. That is, if $x : W \to \mathbb{R}^n$ is a p-dimensional $\mathcal{C}^r$ $(r \geq 1)$ chart in $\mathbb{R}^n$ and $x_0 \in W$ such that $x_0 = x(t_0)$, then there exists an open neighborhood U of t_0 in $\mathbb{R}^p$ such that $U \subseteq W$ and $x\colon U \to x(U)$ is a $\mathcal{C}^r$ diffeomorphism.*

Proof. We need only show that x^{-1} has a $\mathcal{C}^r$ extension to an open neighborhood of x_0 in $\mathbb{R}^n$.

We know by Proposition 1 that there exist open sets U_0 and V_0 in $\mathbb{R}^n$ and a map $X : U_0 \to V_0$ such that

1. X is a $\mathcal{C}^r$ diffeomorphism of U_0 to V_0.
2. $(t_0, 0^{n-p}) \in U_0$.
3. $X(t, 0^{n-p}) = x(t)$ for all $(t, 0^{n-p}) \in U_0$.

Let $\pi\colon \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p} \to \mathbb{R}^p$ be the projection map $\pi(t,t') = t$ where $t \in \mathbb{R}^p$ and $t' \in \mathbb{R}^{n-p}$. Notice that the map $\pi \circ X^{-1}\colon V_0 \to \mathbb{R}^p$ is $\mathcal{C}^r$.

We have $(t_0, 0^{n-p}) \in U_0$, so $x_0 \;=\; x(t_0) \;=\; X(t_0, 0^{n-p}) \in V_0$. Hence V_0 is an open neighborhood of x_0 in $\mathbb{R}^n$.

Set $U = \{t \,:\, (t, 0^{n-p}) \in U_0\}$. We see that U is an open neighborhood of t_0 in $\mathbb{R}^p$. If $t \in U$, then $(t, 0^{n-p}) \in U_0$ and $x(t) = X(t, 0^{n-p}) \in V_0$. Therefore $x(U) \subseteq V_0$.

Let $y \in x(U)$, and set $t = x^{-1}(y)$. Since $x(U) \subseteq V_0$, we can apply $\pi \circ X^{-1}$ to y. We see that

$$\pi \circ X^{-1}(y) \;=\; \pi \circ X^{-1} \circ X(t, 0^{n-p}) \;=\; t \;=\; x^{-1}(y).$$

Therefore $\pi \circ X^{-1}$ is a $\mathcal{C}^r$ extension x^{-1} to V_0. □

Proposition 14. *Suppose $f\colon \mathcal{M} \to \mathcal{N}$ is a $\mathcal{C}^1$ onto diffeomorphism where $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{C}^1$ manifolds. Let x_0 be a point in $\mathcal{M}$ and set $y_0 = f(x_0)$. Let u and v be tangent vectors to $\mathcal{M}$ at x_0 and to $\mathcal{N}$ at y_0 respectively. Then*

$$v \;=\; \big[f'(x_0)\big]u \quad \textit{if and only if} \quad u \;=\; \big[\big(f^{-1}\big)'(y_0)\big]v. \tag{36}$$

Thus $f'(x_0)\colon T_{x_0}\mathcal{M} \to T_{y_0}\mathcal{N}$ is an isomorphism of the first tangent space onto the second one and $\big[f'(x_0)\big]^{-1} = \big(f^{-1}\big)'(y_0)$.

Proof. If we establish the validity of (36), then the conclusion follows from Proposition 8.

Assume $v = \big[f'(x_0)\big]u$.

We first consider the case where u is a tangent vector generated by a $\mathcal{C}^1$ curve c in $\mathcal{M}$. We may suppose that $x_0 = c(0)$ and $u = c'(0)$. Notice that $f \circ c$ must be a $\mathcal{C}^1$ curve in $\mathcal{N}$. Further, $y_0 = f(x_0) = \big(f \circ c\big)(0)$ and

$$\big(f \circ c\big)'(0) \;=\; \big[f'(x_0)\big]c'(0) \;=\; \big[f'(x_0)\big]u \;=\; v.$$

That is, v is generated in $\mathcal{N}$ by the curve $f \circ c$. Then by Proposition 9, and Definition 6,

$$\begin{aligned}
\big[\big(f^{-1}\big)'(y_0)\big]v \;&=\; \partial_v\big(f^{-1}\big)'(y_0)\\
&= \lim_{\lambda \to 0} \frac{1}{\lambda}\Big(f^{-1}\big(f \circ c(\lambda)\big) - f^{-1}\big(f \circ c(0)\big)\Big)\\
&= \lim_{\lambda \to 0} \frac{1}{\lambda}\big(c(\lambda) - c(0)\big) \;=\; u.
\end{aligned}$$

This gives us one of the implications of (36) but only in the case where u is generated by a curve in $\mathcal{M}$.

In the general case, we know by Proposition 7 that we can write $u = \sum_{i=1}^{p} \lambda_i u_i$ where each u_i is generated by a curve in $\mathcal{M}$. Set $v_i = \big[f'(x_0)\big]u_i$. By our first case, $u_i = \big[\big(f^{-1}\big)'(y_0)\big]v_i$. Then

$$u \;=\; \sum_{i=1}^{p} \lambda_i\big[\big(f^{-1}\big)'(y_0)\big]v_i \;=\; \big[\big(f^{-1}\big)'(y_0)\big]v.$$

This completes the proof of the first implication of (36).

The second implication is obtained by a similar argument. □

Proposition 15. *A $\mathcal{C}^1$ diffeomorphism between $\mathcal{C}^1$ manifolds preserves dimension.*

Proof. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{C}^1$ p- and q-manifolds respectively and suppose $f\colon \mathcal{M} \to \mathcal{N}$ is a $\mathcal{C}^1$ diffeomorphism of $\mathcal{M}$ onto $\mathcal{N}$. Set $y_0 = f(x_0)$. Proposition 5 tells us that $T_{x_0}\mathcal{M}$ and $T_{y_0}\mathcal{N}$ have dimensions p and q, and by Proposition 14, $f'(x_0)$ is an isomorphism of $T_{x_0}\mathcal{M}$ onto $T_{y_0}\mathcal{N}$. Therefore $p = q$. □

8 Directional derivatives and coordinates

Suppose that $\mathcal{M}$ is a $\mathcal{C}^1$ p-manifold in $\mathbb{R}^n$ and f is a function defined on $\mathcal{M}$. It is sometimes useful to express f in terms of coordinates and then compute derivatives with respect to those coordinates. There are two obvious sets of coordinates in this situation:

First, there are the coordinates $(\chi_1, \ldots, \chi_n)$ attached to every point x of $\mathbb{R}^n$. In this case,

$$\frac{\partial f}{\partial \chi_i}(\chi_1, \ldots, \chi_n) \;=\; \partial_{e_i} f(\chi_1, \ldots, \chi_n) \tag{37}$$

where $\{e_i\}_{i=1}^n$ is the standard basis of $\mathbb{R}^n$ and we are back at Equation (1) and calculating the familiar partial derivatives of introductory calculus.

Second, if we have a $\mathcal{C}^1$ chart $x\colon U \to \mathbb{R}^n$ on $\mathcal{M}$ and x induces coordinates $(\tau_1, \ldots, \tau_p)$ on $\mathcal{M}$, then we can ask for the derivative of f with respect to τ_i. Let $x_0 = x(t_0) = x(\tau_1, \ldots, \tau_p) \in \mathcal{M}$. We set

$$\frac{\partial f}{\partial \tau_i}(x_0) \;\stackrel{\text{def.}}{=}\; \partial_{e_i}\big(f \circ x\big)(t_0) \;=\; \big[\big(f \circ x\big)'(t_0)\big]\, e_i. \tag{38}$$

Notice that (37) is a special case of (38) if we take $\mathcal{M}$ to be $\mathbb{R}^n$ and the chart x to be the identity map on $\mathbb{R}^n$.

If we have a function defined on a manifold in terms of one set of coordinates, then we may replace it by one expressed in the other set of coordinates. That is, $f(\chi_1, \ldots, \chi_n)$ may be replaced by $g(\tau_1, \ldots, \tau_p) = \big(f \circ x\big)(\tau_1, \ldots, \tau_p)$, and $g(\tau_1, \ldots, \tau_p)$ may be replaced by $\big(g \circ x^{-1}\big)(\chi_1, \ldots, \chi_n)$ where x is our chart. In connection with this, remember (Proposition 13) that every $\mathcal{C}^r$ chart is locally a $\mathcal{C}^r$ diffeomorphism.

Recall that in Remark 7 we defined the vector field $\partial x/\partial\tau_i$ on $\mathcal{M}$ by

$$\frac{\partial x}{\partial\tau_i}(x_0) \;=\; \bigl[x'(t_0)\bigr]\, e_i \quad \text{where } x_0 = x(t_0) \text{ and } e_i \in \mathbb{R}^p \tag{39}$$

and $\bigl\{\bigl(\partial x/\partial\tau_i\bigr)(x_0)\bigr\}_{i=1}^{p}$ is a basis for the tangent space $T_{x_0}\mathcal{M}$.

The reader may notice that (38) and (39) do not seem consistent notations because the domains and ranges of f and x play different roles: The domain of f is $\mathcal{M}$ while $x\colon U \to \mathbb{R}^n$ where U is open in $\mathbb{R}^p$ and $\operatorname{ran}(x) \cap \mathcal{M} \neq \emptyset$. We can remedy this inconsistency by assigning *two different interpretations* to the symbol x. Sometimes we will continue to think of it as the chart $x\colon U \to \mathbb{R}^n$ on $\mathcal{M}$ and sometimes we will think of it as the *identity map* $x\colon \mathcal{M} \to \mathcal{M}$ on $\mathcal{M}$.

If we wish, we can introduce a special symbol for the identity map on $\mathcal{M}$, namely $I_{\mathcal{M}}\colon \mathcal{M} \to \mathcal{M}$, so that part of the time we have $x = I_{\mathcal{M}}$. Notice that the definition of (38) is applicable to $I_{\mathcal{M}}$. This raises the question, do we have

$$\frac{\partial x}{\partial\tau_i} \;=\; \frac{\partial I_{\mathcal{M}}}{\partial\tau_i} \tag{40}$$

when x is a chart and we are using the definition of Equation (39) on the left while $I_{\mathcal{M}}$ is the identity map on $\mathcal{M}$ and we are using the definition of Equation (38) on the right? We answer this in the affirmative thus:

Proposition 16. *Equation (40) is valid.*

Proof. Using x as a chart and remembering that $x_0 = x(t_0)$, we calculate thus:

$$\begin{aligned}
\frac{\partial I_{\mathcal{M}}}{\partial\tau_i}(x_0) \;&=\; \partial_{e_i}\bigl(I_{\mathcal{M}} \circ x\bigr)(t_0) \\
&= \lim_{\lambda\to 0} \frac{1}{\lambda}\, \bigl[x(t_0 + \lambda e_i) - x(t_0)\bigr] \\
&= \bigl[x'(t_0)\bigr]\, e_i \\
&= \frac{\partial x}{\partial\tau_i}(x_0).
\end{aligned}$$

□

This seems a reasonable point to establish a second, closely related fact which is useful when differentiating with respect to a coordinate:

Proposition 17. *If f is a $\mathcal{C}^1$ function on $\mathcal{M}$ and x_0 lies in the domain of f, then*

$$\frac{\partial f}{\partial \tau_i}(x_0) \;=\; \partial_{u_i} f(x_0) \;=\; \big[f'(x_0)\big]\, u_i$$

where x is a $\mathcal{C}^1$ chart on $\mathcal{M}$ covering x_0, $x_0 = x(t_0)$, and

$$u_i \;=\; \frac{\partial x}{\partial \tau_i}(x_0) \;=\; \big[x'(t_0)\big]\, e_i.$$

Proof. We call on Equations (38) and (2):

$$\frac{\partial f}{\partial \tau_i}(x_0) \;=\; \partial_{e_i}\big(f\circ x\big)(t_0) \;=\; \big[\big(f\circ x\big)'(t_0)\big]\, e_i.$$

Appealing to the chain rule, we see that

$$\begin{aligned}\big[\big(f\circ x\big)'(t_0)\big]\, e_i \;&=\; \big[f'(x_0)\big]\,\big[x'(t_0)\big]\, e_i\\ &= \big[f'(x_0)\big]\, u_i\\ &= \big(\partial_{u_i} f\big)(x_0)\end{aligned}$$

which is the desired result. □

This trick of assigning two interpretations to x does not seem to cause trouble in practice. See also the remarks at the end of Section 9.

We end with remarks about derivatives of functions which are expressed in terms of the coordinates induced on a manifold.

We suppose we have a $\mathcal{C}^r$ p-manifold $\mathcal{M}$ ($r \geq 1$) in $\mathbb{R}^n$. Assume the coordinates $(\tau_1, \dots, \tau_p)$ are induced on $\mathcal{M}$ by a $\mathcal{C}^r$ chart $x\colon U \to \mathbb{R}^n$.

We know that $(\tau_1, \dots, \tau_p)$ is a p-tuple of real numbers, the coordinates of a point of $\mathcal{M}$ induced by the chart x. However we can also consider each τ_i to be a map which assigns to a point of $\mathcal{M}$ the ith component of its induced coordinates. To be more precise, suppose that $x_0 \in \mathcal{M}$, that $x_0 = x(t_0)$, and that $t_0 = (\tau_{01}, \dots, \tau_{0p})$. Notice that we have a well-defined function

$$x_0 = x(\tau_{01}, \dots, \tau_{0p}) \;\mapsto\; (\tau_{01}, \dots, \tau_{0p}) \;\mapsto\; \tau_{0i}.$$

We denote this map as τ_i and have $\tau_i(x_0) = \tau_{0i}$. Obviously $\tau_i = \pi_i \circ x^{-1}$ where π_i is the projection of an ordered p-tuple to its ith component. Since, by Proposition 13, x^{-1} is locally $\mathcal{C}^r$, the same must be true of the map τ_i. We immediately deduce the following:

Proposition 18. *Assuming τ_i is $\mathcal{C}^1$,*

$$\frac{\partial \tau_i}{\partial \tau_j} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Proof. Suppose that $x_0 = x(t_0)$. By Equation (38), we have

$$\frac{\partial \tau_i}{\partial \tau_j}(x_0) \;=\; \partial_{e_j}\big(\tau_i \circ x\big)(t_0).$$

Since $\tau_i = \pi_i \circ x^{-1}$, we see that $\tau_i \circ x = \pi_i$ and compute that

$$\partial_{e_j}\big(\tau_i \circ x\big)(t_0) \;=\; \big(\partial_{e_j}\pi_i\big)(t_0) \;=\; \delta_{ij}$$

where δ_{ij} is Kronecker's delta. □

Suppose we want to talk about ϕ being a real-valued function on $\mathcal{M}$ defined in terms of the induced coordinates $(\tau_1, \ldots, \tau_p)$ on $\mathcal{M}$. We are led to consider the expression $\phi(\tau_1, \ldots, \tau_p)$. Our chart maps thus: $x\colon U \to \mathbb{R}^n$ where U is an open subset of $\mathbb{R}^p$. If we are thinking of $\tau_1, \ldots, \tau_p$ as real numbers, then $\phi(\tau_1, \ldots, \tau_p)$ is defined on U, not on $\mathcal{M}$. However we want to think of $\phi(\tau_1, \ldots, \tau_p)$ as being defined on $\mathcal{M}$; therefore we shall not think of $\tau_1, \ldots, \tau_p$ as real numbers but instead as the corresponding coordinate functions which are defined on $\mathcal{M}$. Notice that if we do this, it makes perfect sense to write expressions such as $\phi(\tau_1, \tau_2) = \tau_1^3 \sin(4\tau_2)$.

If we adopt this approach, we find that computing $\partial\big(\phi(\tau_1, \ldots, \tau_p)\big)/\partial\tau_i$ reduces to the sort of computations we learned in introductory calculus. Suppose, for instance, we want to compute $\partial\big(\phi(\tau_1, \tau_2)\big)/\partial\tau_1$. Choose $x_0 \in \mathcal{M}$ and suppose that $x_0 = x(\tau_{01}, \tau_{02})$. Then appealing carefully to our definitions and the fact that $\tau_i = \pi_i \circ x^{-1}$, we have

$$\begin{aligned} \frac{\partial\big(\phi(\tau_1, \tau_2)\big)}{\partial \tau_1}(x_0) \;&=\; \partial_{e_1}\big(\phi(\tau_1, \tau_2) \circ x\big)(t_0) \\ &= \partial_{e_1}\big(\phi(\pi_1, \pi_2)\big)(t_0) \\ &= \lim_{\lambda \to 0} \frac{1}{\lambda}\,\big(\phi(\tau_{01} + \lambda,\, \tau_{02}) - \phi(\tau_{01}, \tau_{02})\big). \end{aligned}$$

This last expression is exactly the introductory calculus formula for computing $\partial\phi(\tau_1, \tau_2)/\partial\tau_1$. We see from this that if we have, for example, $\phi(\tau_1, \tau_2) = \tau_1^3 \sin(4\tau_2)$, where τ_1, τ_2 are the coordinate functions on a manifold, then

$$\frac{\partial\big(\phi(\tau_1, \tau_2)\big)}{\partial \tau_1} \;=\; 3\tau_1^2 \sin(4\tau_2),$$

$$\frac{\partial\big(\phi(\tau_1,\tau_2)\big)}{\partial\tau_2} \;=\; 4\tau_1^3\cos(4\tau_2).$$

9 The normal identity

If $\mathcal{M}$ is a manifold in $\mathbb{R}^n$, one sometimes wishes to use calculus-type operations on the identity map $I\colon \mathcal{M} \to \mathcal{M}$. In particular, one wants to calculate directional derivatives. If one extends I to the identity on $\mathbb{R}^n$, this can lead to different results depending on how $\mathcal{M}$ is embedded in $\mathbb{R}^n$. Another way to extend I—a way which we call the *normal identity on* $\mathcal{M}$—leads to results that have a more intrinsic character and are in more accord with the behavior of the identity map in geometric calculus.

9.1 The idea of the normal identity

Suppose that $\mathcal{M}$ is a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^n$ (where $r \geq 2$). Recall that $T_x\mathcal{M}$ is the *tangent space to* $\mathcal{M}$ *at* x. By $N_x\mathcal{M}$, *the normal space to* $\mathcal{M}$ *at* x, we shall mean the orthogonal complement to $T_x\mathcal{M}$. That is, $N_x\mathcal{M}$ is the vector subspace of $\mathbb{R}^n$ consisting of those vectors v such that $v \bullet u = 0$ for all $u \in T_x\mathcal{M}$.

We then introduce the $(n-p)$-dimensional hyperplane $H_x\mathcal{M}$ by

$$H_x\mathcal{M} \;\stackrel{\text{def.}}{=}\; \{x + v \,:\, v \in N_x\mathcal{M}\} \;=\; x + N_x\mathcal{M}.$$

We think of this as the hyperplane that passes through x and is orthogonal to $\mathcal{M}$ at x. (See Figure 5 for the cases in $\mathbb{R}^3$ where $p = 1$ and $p = 2$.)

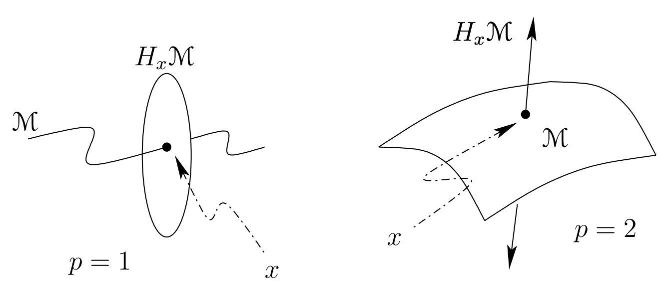

Figure 5: $H_x\mathcal{M}$ for 1- and 2-manifolds

It is now easy to explain what we want the *normal identity* to be: For $y \in \mathbb{R}^n$, provided y is sufficiently close to $\mathcal{M}$, we set

$$I_{\mathcal{M}}(y) \;=\; x \quad \text{when } y \in H_x\mathcal{M}$$

where $I_{\mathcal{M}}$ is our symbol for the normal identity map associated with $\mathcal{M}$. It is clear that for $x \in \mathcal{M}$, we have $I_{\mathcal{M}}(x) = x$.

9.2 Construction of the normal identity

In what follows, if x_0 is a point in $\mathbb{R}^n$ and ξ is a positive number, then by $B(x_0, \xi)$ we mean the set of $x \in \mathbb{R}^n$ such that $|x - x_0| < \xi$. That is, $B(x_0, \xi)$ is the *open ball* in $\mathbb{R}^n$ centered at x_0 with radius ξ. If $t \in \mathbb{R}^p$ and $t' \in \mathbb{R}^q$, then $B\big((t, t'), \delta\big)$ is the open ball in $\mathbb{R}^{p+q}$ that is centered at $(t, t') \in \mathbb{R}^{p+q}$.

Proposition 19. *Suppose that $\mathcal{M}$ is a $\mathcal{C}^r$ p-manifold (where $r \geq 2$ and $p \geq 1$) in $\mathbb{R}^n$ and $x_0 \in \mathcal{M}$. Then for every $\xi > 0$ there is an open neighborhood V in $\mathbb{R}^n$ of x_0 and a map $I_{\mathcal{M}} : V \to V$ such that the following hold:*

1. *$V \subseteq B(x_0, \xi)$.*
2. *For all $x \in V \cap \mathcal{M}$ and $y \in V$, $I_{\mathcal{M}}(y) = x$ if and only if $y \in V \cap H_x\mathcal{M}$.*
3. *$I_{\mathcal{M}}$ is $\mathcal{C}^{r-1}$.*

Remark 9. If $x \in V \cap \mathcal{M}$, then $x \in V \cap H_x\mathcal{M}$ so that $I_{\mathcal{M}}(x) = x$. Thus $I_{\mathcal{M}}$ is an extension of the identity map on $\mathcal{M}$.

We also note that the map $I_{\mathcal{M}}$ is *almost* uniquely defined. If both the conditions $y \in V \cap H_x\mathcal{M}$ and $x \in \mathcal{M}$ are satisfied, then $I_{\mathcal{M}}(y) = x$; that is, $I_{\mathcal{M}}$ is unique under those conditions. However if $y \in V$ but $y \notin H_x\mathcal{M}$ for some $x \in \mathcal{M}$—a situation which can occur for y close to the boundary of $\mathcal{M}$—then there may be more than one choice for $I_{\mathcal{M}}$. We shall say more about this in the proof of the proposition.

Proof of Proposition 19. If $p = n$, then $N_x\mathcal{M} = \{0\}$ whenever $x \in \mathcal{M}$, and thus $H_x\mathcal{M} = \{x\}$. If we take $I_{\mathcal{M}}$ to be the identity map on $\mathbb{R}^n$, then the result is trivially true.

Therefore, from this point on, we assume $p < n$.

Choose a $\mathcal{C}^r$ chart $x\colon W \to \mathbb{R}^n$ on $\mathcal{M}$ with $x_0 = x(t_0)$. We know that $\{\partial_{e_i} x(t_0)\}_{i=1}^p$ is a basis for $T_{x_0}\mathcal{M}$. Recalling that $\{e_i\}_{i=1}^n$ is the standard basis for $\mathbb{R}^n$, we may suppose, without loss of generality, that

$$\{\partial_{e_i} x(t)\}_{i=1}^p \cup \{e_{p+1}, \ldots, e_n\} \tag{41}$$

is a linearly independent set (hence a basis for $\mathbb{R}^n$) at $t = t_0$. More than that, by continuity, (41) must be a basis for $\mathbb{R}^n$ for all t sufficiently close to t_0.

Next construct the reciprocal frame $\{m_i(x)\}_{i=1}^n$ for (41) where $x = x(t)$. By the definition of a reciprocal frame, this is a basis for $\mathbb{R}^n$ such that

$$\delta_{ij} \;=\; \begin{cases} \big(\partial_{e_i} x\big) \bullet m_j & \text{for } i \leq p, \\ e_i \bullet m_j & \text{for } p < i, \end{cases}$$

where δ_{ij} is Kronecker's delta. Of course, each m_j is a function of x and thus of t, $m_j = m_j(x) = m_j\big(x(t)\big)$. Since

$$\big(\partial_{e_i} x\big) \bullet m_j \;=\; 0 \quad \text{for } i \leq p \text{ and } j > p, \tag{42}$$

we see that

$$\{\partial_{e_i} x(t)\}_{i=1}^p \cup \{m_{p+1}(x), \ldots, m_n(x)\} \tag{43}$$

must be a basis for $\mathbb{R}^n$ for t sufficiently close to t_0.

On an open neighborhood of the point $(t_0, 0^{n-p})$, we now construct a map Y between subsets of $\mathbb{R}^n$. Let $t = (\tau_1, \ldots, \tau_p) \in \mathbb{R}^p$ and $t' = (\tau_{p+1}, \ldots, \tau_n) \in \mathbb{R}^{n-p}$. Set

$$Y(t, t') \;=\; x(t) + \tau_{p+1} m_{p+1}(x) + \cdots \tau_n m_n(x) \tag{44}$$

where we understand that $x = x(t)$. Notice that $x_0 = Y(t_0, 0^{n-p})$. We know that the map x is $\mathcal{C}^r$ and each $\partial_{e_i} x$ must be at least $\mathcal{C}^{r-1}$. Since the reciprocal vectors m_i are constructed from the vectors $\partial_{e_i} x$ by an algebraic process, we see that Y must be at least $\mathcal{C}^{r-1}$.

It is straightforward to calculate that

$$\frac{\partial Y}{\partial \tau_i}(t_0, 0) \;=\; \begin{cases} \partial_{e_i} x(t_0) & \text{if } i = 1, \ldots, p \\ m_i(x_0) & \text{if } i = p+1, \ldots, n. \end{cases}$$

Since this is a linearly independent set and $\partial_{e_i} x$ and m_j are $\mathcal{C}^{r-1}$ (with $r \geq 2$), we see that $\det(Y')$ must be nonzero for (t, t') sufficiently close to $(t_0, 0)$.

Therefore by the inverse function theorem, there exist open sets U_1 and V_1 of $\mathbb{R}^n$ such that $Y\colon U_1 \to V_1$ is a $\mathcal{C}^{r-1}$ diffeomorphism and $(t_0, 0^{n-p}) \in U_1$.

Case 1. Let us now suppose that x_0 is an interior point of $\mathcal{M}$ and that $\xi > 0$ is a given value.

Because x and Y are continuous maps, there exists $\epsilon > 0$ such that

$$\begin{aligned} B(t_0, \epsilon) &\subseteq W = \text{domain of } x, \\ B\big((t_0, 0^{n-p}), \epsilon\big) &\subseteq U_1, \\ Y\left(B\big((t_0, 0^{n-p}), \epsilon\big)\right) &\subseteq V_1 \cap B(x_0, \xi). \end{aligned} \tag{45}$$

It follows that $x_0 \in x\big(B(t_0, \epsilon)\big) \subseteq \mathcal{M}$. (See Figure 6 for how these open balls relate to one another and how they map under Y.)

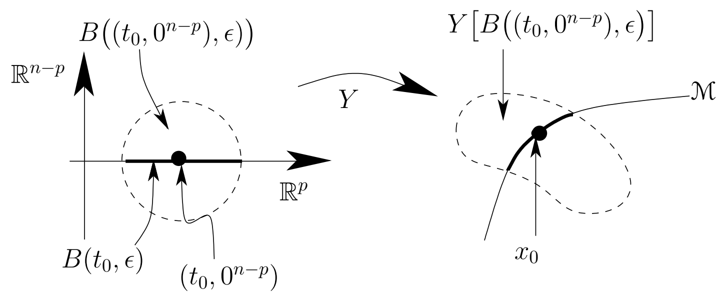

Figure 6: Y mapping open balls

Now let

$$\begin{aligned} U &= B\big((t_0, 0^{n-p}), \epsilon\big), \\ V &= Y(U). \end{aligned}$$

Since $U \subseteq U_1$ and $V \subseteq V_1$, we see that $Y\colon U \to V$ is still a $\mathcal{C}^{r-1}$ diffeomorphism. Because x_0 is an interior point of $\mathcal{M}$, $x(W) \cap \mathcal{M}$ is an open subset of $\mathcal{M}$, and Y is continuous, we can impose a further restriction on ϵ and require that

$$\begin{aligned} V \cap x(W) &= Y\big(B\big((t_0, 0^{n-p}), \epsilon\big)\big) \cap x(W) \\ &= V \cap \mathcal{M}. \end{aligned} \tag{46}$$

We now develop the ways in which certain sets are related to one another.

Suppose that $x = x(t)$ where $t \in B(t_0, \epsilon)$. We know that $\dim T_x\mathcal{M} = p$, so $\dim N_x\mathcal{M} = n - p$. Since $m_{p+1}(x), \ldots, m_n(x)$ are linearly independent and orthogonal to $\mathcal{M}$ at x, it follows that $\{m_i(x)\}_{i=p+1}^n$ must be a basis for $N_x\mathcal{M}$. We know that $H_x\mathcal{M} = x + N_x\mathcal{M}$, and in (44), the definition of Y, the scalars $\tau_{p+1}, \ldots, \tau_n$ range over all values of $\mathbb{R}$; it follows that for t held constant, we have

$$Y(t \times \mathbb{R}^{n-p}) \;=\; H_x\mathcal{M}.$$

However since we have the constraint $Y\colon U \to V$, the relation we really want is

$$Y\big(U \cap (t \times \mathbb{R}^{n-p})\big) \;=\; V \cap H_x\mathcal{M} \quad \text{where } t \in B(t_0, \epsilon). \tag{47}$$

Next notice that

$$U \cap (\mathbb{R}^p \times 0^{n-p}) \;=\; B\big((t_0, 0^{n-p}), \epsilon\big) \cap (\mathbb{R}^p \times 0^{n-p}) \;=\; B(t_0, \epsilon) \times 0^{n-p} \tag{48}$$

and

$$Y\big(B(t_0, \epsilon) \times 0^{n-p}\big) \;=\; x\big(B(t_0, \epsilon)\big),$$

hence

$$Y(U \cap (\mathbb{R}^p \times 0^{n-p}) \;=\; x\big(B(t_0, \epsilon)\big). \tag{49}$$

From (48), we have $U \cap (\mathbb{R}^p \times 0^{n-p}) \;\subseteq\; W \times 0^{n-p}$, so we must have

$$U \cap (\mathbb{R}^p \times 0^{n-p}) \;=\; U \cap (W \times 0^{n-p}).$$

By this fact and (46), we see that

$$Y\big(U \cap (\mathbb{R}^p \times 0^{n-p})\big) \;=\; V \cap \mathcal{M}. \tag{50}$$

Combining (49) and (50) gives us

$$x\big(B(t_0, \epsilon)\big) \;=\; V \cap \mathcal{M}. \tag{51}$$

We are now ready to define the normal identity $I_{\mathcal{M}}\colon V \to V \cap \mathcal{M}$. Let $\pi\colon \mathbb{R}^n \to \mathbb{R}^p$ be the map $\pi(t, t') = t$. Notice that in the following diagram, all the maps are onto:

$$V \xleftarrow{Y} U \xrightarrow{\pi} B(t_0, \epsilon) \xrightarrow{x} V \cap \mathcal{M}.$$

We set

$$I_{\mathcal{M}} \;\overset{\text{def.}}{=}\; x \circ \pi \circ Y^{-1}. \tag{52}$$

By construction, $I_{\mathcal{M}}$ is $\mathcal{C}^{r-1}$, and we know from (45) that $V \subseteq B(x_0, \xi)$, so the only thing we need to check is the relation of $I_{\mathcal{M}}$ to H_x.

Choose $y \in V$ and $x \in V \cap \mathcal{M}$. We know that

$$U \cap (t \times \mathbb{R}^{n-p}) \;=\; B\big((t_0, 0^{n-p}), \epsilon\big) \cap (t \times \mathbb{R}^{n-p}),$$

and this will be nonempty if and only if $(t, 0^{n-p})$ lies in $B\big((t_0, 0^{n-p}), \epsilon\big)$, that is, if and only if $t \in B(t_0, \epsilon)$. Thus we can write

$$U \;=\; \bigcup_{t \in B(t_0,\epsilon)} \big(U \cap (t \times \mathbb{R}^{n-p})\big),$$

and by (47) this becomes

$$V \;=\; \bigcup_{t \in B(t_0,\epsilon)} \big(V \cap H_{x(t)}\mathcal{M}\big).$$

Since the sets $t \times \mathbb{R}^{n-p}$ are pairwise disjoint, the same must be true for the sets $V \cap H_{x(t)}\mathcal{M}$. Thus for our $y \in V$, there exists a unique $t \in B(t_0, \epsilon)$ such that $y \in H_{x(t)}\mathcal{M}$. We can write this y in the form

$$y \;=\; x(t) + \tau_{p+1} m_{p+1}(x(t)) + \cdots + \tau_n m_n(x(t)) \;=\; Y(t, t')$$

where $t' = (\tau_{p+1}, \ldots, \tau_n) \in \mathbb{R}^{n-p}$ is uniquely determined for y. By the definition of $I_{\mathcal{M}}$, we have $I_{\mathcal{M}}(y) = x(t)$. We see that $I_{\mathcal{M}}(y) = x$ if and only if $V \cap H_x = V \cap H_{x(t)}$. Thus $I_{\mathcal{M}}(y) = x$ if and only if $y \in H_x$.

Case 2. Suppose that x_0 is not an interior point of $\mathcal{M}$, that it is a boundary point. We know that we can construct a slightly larger $\mathcal{C}^r$ manifold $\mathcal{M}^+$ containing $\mathcal{M}$ such that x_0 is an interior point of $\mathcal{M}^+$. We now find an open neighborhood V of x_0 in $\mathbb{R}^n$ and construct the normal identity $I_{\mathcal{M}^+} : V \to V$ for $\mathcal{M}^+$. This will be a $\mathcal{C}^{r-1}$ map and will automatically have the properties desired for $I_{\mathcal{M}}$, so we may take $I_{\mathcal{M}^+}$ to be our $I_{\mathcal{M}}$.

We note that this map is not uniquely defined since there must exist an infinite number of ways to expand $\mathcal{M}$ to $\mathcal{M}^+$. $\square$

9.3 Properties of the normal identity

From this point on, we assume $\mathcal{M}$ is a $\mathcal{C}^r$ p-manifold in $\mathbb{R}^n$ with $r \geq 2$.

We must not confuse $I_{\mathcal{M}}$ with an orthogonal projection onto a vector subspace. The connection with subspace orthogonal projections is this:

Corollary 1. *Let* $P_x \colon \mathbb{R}^n \to T_x\mathcal{M}$ *be the orthogonal projection map onto the tangent space of* $\mathcal{M}$ *at the point* $x \in \mathcal{M}$. *Then*

$$I_{\mathcal{M}}(y) = x_0 \quad \textit{implies} \quad P_{x_0}(y) = P_{x_0}(x_0).$$

Proof. By Proposition 19, if $I_{\mathcal{M}}(y) = x_0$, it follows that $y \in H_{x_0}\mathcal{M}$. From this we see that we can write $y = x_0 + v$ where v is orthogonal to $T_{x_0}\mathcal{M}$. Therefore

$$P_{x_0}(y) \;=\; P_{x_0}(x_0) + P_{x_0}(v) \;=\; P_{x_0}(x_0).$$

□

Unless we say otherwise below, we shall use P_{x_0} for the orthogonal projection map $\mathbb{R}^n \to T_{x_0}\mathcal{M}$ where $\mathcal{M} \subseteq \mathbb{R}^n$.

Corollary 2. *If* $x_0 \in \mathcal{M}$ *and* $a \in \mathbb{R}^n$, *then*

$$\partial_a I_{\mathcal{M}}(x_0) \;=\; P_{x_0} a.$$

Proof. We treat x_0 as an interior point of $\mathcal{M}$.

Define a function h from an open neighborhood of 0 in $\mathbb{R}$ into $\mathcal{M}$ by $h(\lambda) = I_{\mathcal{M}}(x_0 + \lambda\, a)$. Since $I_{\mathcal{M}}$ is $\mathcal{C}^{r-1}$, we see that h defines a $\mathcal{C}^{r-1}$ path in $\mathcal{M}$ passing through x_0 at $\lambda = 0$. Then

$$\partial_a I_{\mathcal{M}}(x_0) \;=\; \lim_{\lambda \to 0} \frac{1}{\lambda} \Big(I_{\mathcal{M}}(x_0 + \lambda\, a) - I_{\mathcal{M}}(x_0) \Big) \;=\; h'(0). \tag{53}$$

Since $h(\lambda) = I_{\mathcal{M}}(x_0 + \lambda\, a)$, by Corollary 1, $P_{h(\lambda)}\big(h(\lambda)\big) = P_{h(\lambda)}(x_0 + \lambda\, a)$. This amounts to

$$P_{h(\lambda)} \left(\frac{h(\lambda) - h(0)}{\lambda} - a \right) \;=\; 0. \tag{54}$$

It is straightforward to construct an orthonormal basis $\{u_i(x)\}_{i=1}^{p}$ for $T_x\mathcal{M}$ for all x in some neighborhood in $\mathcal{M}$ of x_0 such that $x \mapsto u_i(x)$ is $\mathcal{C}^{r-1}$. For such x, the orthogonal projection map is given by

$$P_x(v) \;=\; \sum_{i=1}^{p} \big(v \bullet u_i(x)\big) u_i(x).$$

Then Equation (54) becomes

$$\sum_{i=1}^{p} \left[u_i\big(h(\lambda)\big) \bullet \left(\frac{h(\lambda) - h(0)}{\lambda} - a \right) \right] u_i\big(h(\lambda)\big) \;=\; 0.$$

Using the facts that all the functions involved are $\mathcal{C}^{r-1}$ and $h'(0)$ is a tangent vector to $\mathcal{M}$, we let $\lambda \to 0$ and obtain $h'(0) = P_{x_0}(a)$. This gives us the desired result. $\square$

Given an f defined on $\mathcal{M}$ which is $\mathcal{C}^r$ in terms of charts on $\mathcal{M}$, we know from Proposition 12 that we can extend f (locally) to a $\mathcal{C}^r$ function on $\mathbb{R}^n$. The normal identity gives us another way to extend an f from $\mathcal{M}$ to open subsets of $\mathbb{R}^n$.

Definition 10. Given a manifold $\mathcal{M}$ in $\mathbb{R}^n$ and a function f whose domain lies in $\mathbb{R}^n$, we say f is *normally extended from* $\mathcal{M}$ provided $f = f \circ I_{\mathcal{M}}$ in neighborhoods of points of $\mathcal{M}$.

Notice that if f and $\mathcal{M}$ are $\mathcal{C}^r$, then the most we can hope for from $f \circ I_{\mathcal{M}}$ is that it is $\mathcal{C}^{r-1}$ on an open set $\mathbb{R}^n$.

Corollary 3. *If f is a normally extended $\mathcal{C}^r$ function on $\mathcal{M}$, x_0 is a point on $\mathcal{M}$, a is a vector in $\mathbb{R}^n$, and $b = P_{x_0}a$, then*

$$\partial_a f(x_0) \ = \ \partial_b f(x_0).$$

Proof. This follows from Corollary 2 and

$$\big[f'(x_0)\big]a \ = \ \big[f'(x_0)\, I'_{\mathcal{M}}(x_0)\big]a \ = \ \big[f'(x_0)\big]b. \qquad \square$$

It is trivially true that if we have $\mathcal{C}^r$ manifolds where one contains the other, $\mathcal{M} \subseteq \mathcal{N}$, then at points sufficiently close to $\mathcal{M}$, we have $I_{\mathcal{N}} \circ I_{\mathcal{M}} = I_{\mathcal{M}}$. It is not true in general that $I_{\mathcal{M}} \circ I_{\mathcal{N}} = I_{\mathcal{M}}$.

One final comment concerns a matter of notation from the literature of geometric calculus.

Given a manifold $\mathcal{M}$, the symbol x is used in geometric calculus, in works such as [9] and [20], to denote the identity function on $\mathcal{M}$ or an arbitrary point of $\mathcal{M}$. Its behavior is like that of our concept of the normal identity, and this suggests the use of x as an alternate symbol for $I_{\mathcal{M}}$.

A drawback of this practice is that we now have three ways in which we use this small and innocent-looking symbol:

1. As a chart, $x \colon U \to \mathbb{R}^n$ on $\mathcal{M}$.

2. As a point on the manifold, $x \in \mathcal{M}$.

3. And now as a particular map, $x = I_{\mathcal{M}} \colon V \to V$ acting on neighborhoods of points of $\mathcal{M}$.

There is a possibility of confusion. However in practice, the particular meaning of any given x should be clear, and the resulting applications have an appealing quality of *naturalness*. For example, it seems reasonably sensible to write

$$f \;=\; f \circ I_{\mathcal{M}} \;=\; f(x).$$

Acknowledgment: Special thanks are due to Alan Macdonald for suggesting the writing of this article.

References

[1] Rafał Abłamowicz and Garret Sobczyk, editors. *Lectures on Clifford (Geometric) Algebras and Applications*. Birkhäuser, 2004.

[2] A. Cortzen. Direct construction of Grassmann, Clifford and geometric algebras. arXiv:1011.3698v1, November 2010.

[3] Chris Doran and Anthony Lasenby. *Geometric Algebra for Physicists*. Cambridge University Press, 2003.

[4] Leo Dorst, Daniel Fontijne, and Stephen Mann. *Geometric Algebra for Computer Science (Revised Edition): An Object-Oriented Approach to Geometry*. The Morgan Kaufmann Series in Computer Graphics. Morgan Kaufmann, 2007.

[5] Mikhael Gromov. *Partial Differential Relations*, volume 3. Folge, Bd. 9 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer, 1986.

[6] David Hestenes. A unified language for mathematics and physics. In J. S. R. Chisholm and A. K. Commons, editors, *Clifford Algebras and their Applications in Mathematical Physics*, pages 1–23, Boston, 1986. Dordrecht.

[7] David Hestenes. Mathematical Viruses. In J. Helmstetter A. Micali, R. Boudet, editor, *Clifford Algebras and their Applications in Mathematical Physics*, volume 47 of *Fundamental Theories of Physics*, pages 3–16. Springer, 1992.

[8] David Hestenes. The Shape of Differential Geometry in Geometric Calculus. In L.Dorst and J.Lasenby, editors, *Guide to Geometric Algebra in Practice*, chapter 19, pages 393–410. Springer, 2011.

[9] David Hestenes and Garret Sobczyk. *Clifford Algebra to Geometric Calculus*. D. Reidel, 1984.

[10] Kenneth Hoffman. *Analysis in Euclidean Spaces*. Dover, 2007.

[11] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry*, volume I-II. Interscience, New York, 1963.

[12] Douglas Lindholm and Lars Svennson. Clifford algebra, geometric algebra, and applications. arXiv:0907.5356v1, July 2009.

[13] Alan Macdonald. A Survey of Geometric Algebra and Geometric Calculus. http://faculty.luther.edu/~macdonal., 2010.

[14] Alan Macdonald. *Linear and Geometric Algebra*. Alan Macdonald, 2010.

[15] Alan Macdonald. *Vector and Geometric Calculus*. Alan Macdonald, 2012.

[16] Piotr Mikusiński and Michael D. Taylor. *An Introduction to Multivariable Analysis from Vector to Manifold*. Birkhäuser, Boston, 2002.

[17] Frank Morgan. *Riemannian Geometry, A Beginner's Guide*. A. K. Peters, 1998.

[18] Barrett O'Neill. *Elementary Differential Geometry*. Elsevier, revised second edition edition, 2006.

[19] Qing Han and Jia-Xing Hong. *Isometric Embedding of Riemannian Manifolds in Euclidean Spaces*, volume 130 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2006.

[20] Garret Sobczyk. *New Foundations in Mathematics, The Geometric Concept of Number*. Birkhäuser, 2013.

[21] Michael Spivak. *A Comprehensive Introduction to Differential Geometry*, volume one. Publish or Perish, Inc., 1970.