APPENDIX A

Theory of Manifolds

In this appendix we assemble most of the definitions and theorems about manifolds and maps that are needed in the course of the book. Our main purpose is to establish notations and terminology, and to collect useful examples. On the whole, therefore, arguments are omitted or summarized when they are already available in easily accessible texts†. We do not make substantial use of infinite dimensional manifolds in the book, apart from the trivial example of an open subset of an infinite dimensional Banach space. Nevertheless, we shall give the infinite dimensional theory here, since it is useful for further reading in the subject, and the generalization requires only a modicum of care.

I. TOPOLOGICAL MANIFOLDS

Let $\mathbf{E}$ be a real Banach space, finite or infinite dimensional. A *topological manifold modelled* on $\mathbf{E}$ is a topological space X such that, for each $x \in X$, there is an open neighbourhood U of x and a homeomorphism $\xi: U \to U'$ onto an open subset U' of $\mathbf{E}$. It is usual to place further restrictions on the topology of X in order to avoid pathological examples. *We shall always assume that X is Hausdorff with a countable basis of open sets.* Together with the manifold hypothesis, these conditions imply that X is normal, metrizable and paracompact. *In addition we shall suppose, unless otherwise stated, that X is connected.*

†For example, Hirsch [1] is an excellent reference for differential topology, backed up by Lang [1] for infinite dimensional manifolds and Helgason [1] for Riemannian metrics. There are also very readable introductions to differential topology by Guillemin and Pollack [1] and Chillingworth [1]. Chillingworth's book is particularly relevant since it contains a very nice account of dynamical systems theory.

The homeomorphism ξ is called a *chart* at x. A set of charts whose domains cover X is called an *atlas* for X. If $\mathbf{E}$ is n-dimensional ($n < \infty$) then X is said to have *dimension n*, and to be a (topological) *n-manifold*. Since two finite dimensional real normed vector spaces are homeomorphic if and only if their dimensions are the same (see, for example, § 1 of Chapter 4 of Hu [2]), the dimension of X is uniquely defined (with the trivial exception of X = empty set) and we may take $\mathbf{E} = \mathbf{R}^n$. If $\mathbf{E}$ is infinite dimensional, we say that X is infinite dimensional. The term *Banach manifold* commonly carries the connotation of infinite dimension.

Examples

(A.0) The topological space consisting of a single point is a 0-dimensional manifold.

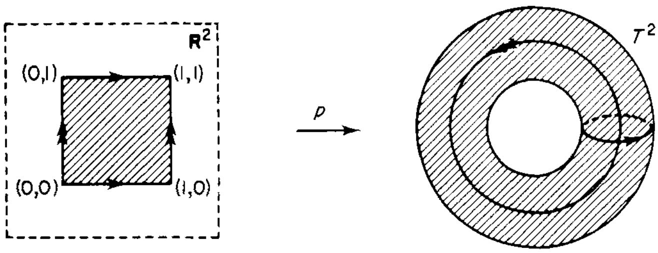
(A.1) Any connected open subset of $\mathbf{E}$ is a manifold modelled on $\mathbf{E}$. The inclusion is a chart, and the single chart is an atlas. Any connected open subset V of a manifold X modelled on $\mathbf{E}$ is a manifold modelled on $\mathbf{E}$. If $\xi: U \to U'$ is a chart on X, $\xi|U \cap V: U \cap V \to \xi(U \cap V)$ is a chart on V.

(A.2) Let $L(\mathbf{E}, \mathbf{F})$ be the Banach space of continuous linear maps from the Banach space $\mathbf{E}$ to the Banach space $\mathbf{F}$. We write $L(\mathbf{E})$ for $L(\mathbf{E}, \mathbf{E})$, and denote by $GL(\mathbf{E})$ the set of (topological) linear automorphisms of $\mathbf{E}$. (Here *topological* means that the automorphisms and their inverses are continuous.) Then $GL(\mathbf{E})$, which is called the general linear group of $\mathbf{E}$, is open in $L(\mathbf{E})$ (Lemma 7.6.1 of Dunford and Schwartz [1]), and hence is a (not necessarily connected) manifold modelled on $L(\mathbf{E})$. For example, $GL(\mathbf{R}^n)$ has two components (automorphisms with positive determinant and automorphisms with negative determinant) which are n^2-manifolds.

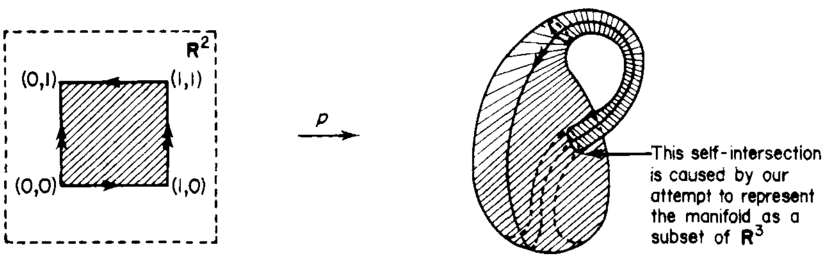
(A.3) If $p: \mathbf{E} \to X$ is a covering map then, by definition, every $x \in X$ has an open neighbourhood U such that $p^{-1}(U)$ is a disjoint union of subsets U_i' each of which is mapped homeomorphically onto U by p. Thus, for any i, $(p|U_i')^{-1}$ is a chart at x, and so X is a manifold modelled on $\mathbf{E}$. In many simple applications, X is the quotient $\mathbf{E}/G$ of a discrete group G acting on $\mathbf{E}$, and p is the quotient map. For example, the *circle* S^1 is the quotient $\mathbf{R}/\mathbf{Z}$ where the action is given by $n \,.\, x = x + n$. Points of S^1 are equivalence classes $[x]$ under the equivalence relation on $\mathbf{R}$ given by $x \sim x'$ if and only if $x - x' \in \mathbf{Z}$, and p is defined by $p(x) = [x]$. We shall give an alternative definition of S^1 shortly. Some examples of 2-manifolds covered by $\mathbf{R}^2$ are (i) the *torus* $T = \mathbf{R}^2/\mathbf{Z}^2$ where $(m, n) \,.\, (x, y) = (x + m, y + n)$, (ii) the *Klein bottle* $\mathbf{R}^2/\mathbf{Z}^2$ where $(m, n) \,.\, (x, y) = ((-1)^n x + m, y + n)$ and (iii) the *Möbius band* $\mathbf{R}^2/\mathbf{Z}$ where $n \,.\, (x, y) = (x + n, (-1)^n y)$. See Figure A.3.

Similarly, if $p: X \to Y$ is any covering of a space Y by a manifold X modelled on $\mathbf{E}$, then Y is a manifold modelled on $\mathbf{E}$.

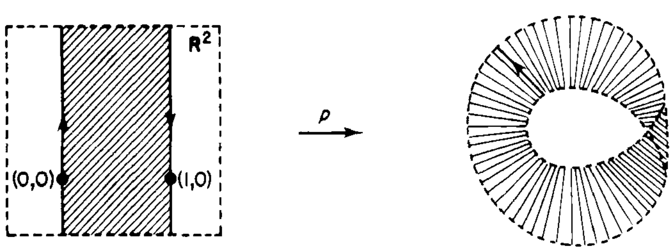
(A.4) The *connected sum* $X \# Y$ of manifolds X and Y modelled on $\mathbf{E}$ is obtained by removing a ball from each and gluing the remnants together along their spherical boundaries. More precisely, let $B_r(0) = \{x \in \mathbf{E}: |x| \leq r\}$ and $S_r(0) = \{x \in \mathbf{E}: |x| = r\}$. Take charts $\xi: U \to U'$ on X and $\eta: V \to V'$ on Y



(i) Torus



(ii) Klein bottle



(iii) Möbius band

FIGURE A.3

with $B_r(0) \subset U' \cap V'$. Let

$$X^- = X \backslash \xi^{-1}(\text{int } B_{r/2}(0)) \quad \text{and} \quad Y^- = Y \backslash \eta^{-1}(\text{int } B_{r/2}(0))$$

and put

$$X \# Y = X^- \cup Y^- / (\xi^{-1}(x) = \eta^{-1}(x) : x \in S_{r/2}(0)).$$

The definition is, up to homeomorphism, independent of choice of charts. For example, if $X = Y = T^2$ then $X \# Y$ is the *pretzel*, or sphere with two handles (see Figure A.4).

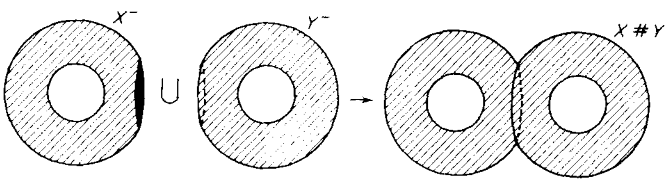


FIGURE A.4

(A.5) If X and Y are manifolds modelled on $\mathbf{E}$ and $\mathbf{F}$ respectively, then $X \times Y$ is a manifold modelled on $\mathbf{E} \times \mathbf{F}$. If ξ and η are charts on X and Y respectively, then $\xi \times \eta$ is a chart on $X \times Y$. For example, the torus T^2 as above defined may be identified with $S^1 \times S^1$ in the obvious way. Precisely, there is a homeomorphism taking $[(x, y)]$ to $([x], [y])$.

We say that a linear subspace $\mathbf{F}$ of $\mathbf{E}$ *splits* $\mathbf{E}$ if it is closed in $\mathbf{E}$, and, for some closed linear subspace $\mathbf{G}$ of $\mathbf{E}$, $\mathbf{E} = \mathbf{F} \oplus \mathbf{G}$. We call dim G the *codimension* of F; it may be ∞. Splitting is guaranteed to occur if $\mathbf{F}$ is finite dimensional. Let X be a manifold modelled on $\mathbf{E}$ and let Y be a connected subset of X. We say that Y is a (*locally flat*) *submanifold* of X if there is a subspace $\mathbf{F}$ of $\mathbf{E}$ that splits $\mathbf{E}$ and, for each $y \in Y$, a chart $\xi: U \to U'$ at y such that ξ maps $U \cap Y$ onto $U' \cap \mathbf{F}$ (see Figure A.6). The *codimension* of Y is the codimension of $\mathbf{F}$. Then obviously the submanifold Y is (with the induced topology) a manifold modelled on $\mathbf{F}$.

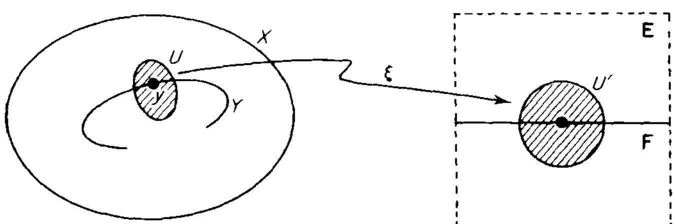


FIGURE A.6

(A.7) Submanifolds of $\mathbf{E}$ often arise in the following way. Suppose that V is an open subset of $\mathbf{E}$ and that $f\colon V \to \mathbf{H}$ is a C^r map ($r \geq 1$) to a Banach space $\mathbf{H}$. Fix on some value $a \in \mathbf{H}$ and consider the level set $f^{-1}(a)$. If, for all $x \in f^{-1}(a)$, the differential $Df(x) \in L(\mathbf{E}, \mathbf{H})$ is surjective (*split surjective* in the infinite dimensional case, which means that ker $Df(x)$ splits $\mathbf{E}$) we say that a is a *regular value* of f, and, in this case, it is a consequence of the inverse mapping theorem that $f^{-1}(a)$ is a disjoint union of submanifolds of V (see Exercise C.13 of Appendix C). For example, if $\mathbf{E} = \mathbf{R}^{n+1}$, $\mathbf{H} = \mathbf{R}$ and $f(x_1, \ldots, x_{n+1}) = x_1^2 + \cdots + x_{n+1}^2$, then 1 is a regular value and $f^{-1}(1)$ is the sphere $x_1^2 + \cdots + x_{n+1}^2 = 1$, the *unit n-sphere* in $\mathbf{R}^{n+1}$, denoted S^n. If $n = 1$, we get the *unit circle* S^1, which is homeomorphic to the manifold defined in A.3.

(A.8) Real n-dimensional projective space $\mathbf{R}P^n$ is defined to be the set of lines through the origin in $\mathbf{R}^{n+1}$. The topology of the space is given by a metric, the distance between two such lines being the distance between their closest points of intersection with S^n. There is a map $p\colon S^n \to \mathbf{R}P^n$ which takes any point to the line joining it to the origin. Since this map is a (double) covering, $\mathbf{R}P^n$ is a manifold, by A.3.

Exercises

(A.9) Let Σ^n be the unit sphere $\{x \in \mathbf{R}^{n+1}\colon \|x\| = 1\}$ in $\mathbf{R}^{n+1}$ with respect to any norm $\|\ \|$. Show that Σ^n is a manifold homeomorphic to S^n.

(A.10) Show that when M and N are submanifolds of X neither $M \cup N$ nor $M \cap N$ need be a submanifold of X.

(A.11) Show that if X and Y are manifolds then, for all $y \in Y$, $X \times \{y\}$ is a submanifold of $X \times Y$.

II. SMOOTH MANIFOLDS AND MAPS

If U is an open subset of a Banach space $\mathbf{E}$, and $\mathbf{F}$ is another Banach space, one has the notion of a C^r ($= r$ times continuously differentiable) map $f\colon U \to \mathbf{F}$. We shall assume familiarity with the basic theory of such maps: some definitions appear in Appendix B below. We shall not be concerned much with C^ω ($=$ analytic) maps. The term *smooth* is used to mean C^r for some r with $1 \leq r \leq \infty$.

We wish to extend the theory of differentiable maps from the context of Banach spaces to the context of manifolds. That is to say, we have to extend the notion of differentiability of a map from the local, chart level to the whole manifold. We do this by restricting ourselves to charts which are smoothly related where they overlap. By doing so we get what is known as a *smooth structure* on the manifold.

Let $\mathscr{A} = \{\xi_i : i \in I\}$ be an atlas for a topological manifold X, where I is some indexing set and $\xi_i : U_i \to U_i' \subset \mathbf{E}$ are charts. We say that $\mathscr{A}$ is a C *atlas* if, for all $i, j \in I$, the coordinate change map $\xi_j \xi_i^{-1} : \xi_i(U_i \cap U_j) \to \xi_j(U_i \cap U_j)$ is C^r (see Figure A.12). A chart ξ on X is C^r-*compatible* with $\mathscr{A}$ if $\mathscr{A} \cup \{\xi\}$ is a C^r

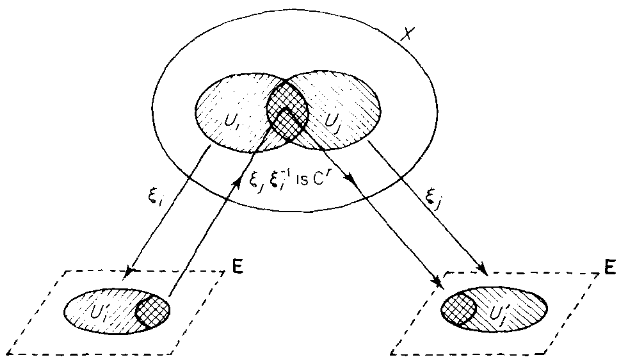


FIGURE A.12

atlas. The set of all charts C^r-compatible with $\mathscr{A}$ is the C^r *structure* generated by $\mathscr{A}$. A C^r *manifold* is a topological manifold X together with a C^r structure $\mathscr{A}$ on X. Any chart in $\mathscr{A}$ is said to be an *admissible* chart on the C^r manifold. *In the text of the book, and from now on in the appendices, "manifold" always means "C^r manifold for some $r \geq 1$" and "chart" means "admissible chart." We often use the symbol X for a C^r-manifold rather than the cumbersome* $(X, \mathscr{A})$.

All the above examples and constructions for topological manifolds may be modified to fit into this new setting. For example the inclusion of an open set U of $\mathbf{E}$ in $\mathbf{E}$ is, by itself, a C^∞ atlas for U, since we have no overlaps to worry about beyond the trivial one of U with itself. If $h: X \to Y$ is a homeomorphism from a C^r manifold X to a topological space Y then h induces a C^r structure on Y, with a chart $\xi h^{-1}: h(U) \to U'$ corresponding to each chart $\xi: U \to U'$ on X. This also works at a local level, so that any connected open subset of a C^r manifold has an induced C^r structure, and so does any space with covering a C^r manifold. We may define C^r *submanifold* by making the charts in the above definition of submanifold admissible charts of a C^r manifold. The level surfaces of the C^r map $f: V \to H$ of Example A.7 turn out to be unions of C^r submanifolds at regular values.

Thus, for example, S^n, and hence $\mathbf{R}P^n$, are C^∞ manifolds. (The C^∞ structure of S^n may be defined, equivalently, by the two charts $\xi: S^n\backslash\{N\} \to \mathbf{R}^n$ and $\eta: S^n\backslash\{S\} \to \mathbf{R}^n$, where N and S are the north and south poles and ξ and η are stereographic projection from those poles.) Products of C^r manifolds are C^r manifolds. The only construction that needs some care is the connected sum construction; it takes a little ingenuity to make $X \# Y$ smooth at the seam.

Now let X and Y be C^r manifolds and let $f: X \to Y$ be any continuous map. We say that f is C^r if, for all admissible charts $\xi: U \to U' \subset \mathbf{E}$ on X and $\eta: V \to V' \subset \mathbf{F}$ on Y, the *local representative*

$$\eta f \xi^{-1}: \xi(U \cap f^{-1}(V)) \to V'$$

is C^r (see Figure A.13). Obviously it is sufficient to check this for all charts in a pair of admissible atlases, one for X and one for Y. If f is C^r and has a C^r inverse $f^{-1}: Y \to X$, we call it a C^r *diffeomorphism* $(r \geqslant 1)$. Any C^r map $f: X \to \mathbf{R}$, where $\mathbf{R}$ has its standard C^∞ structure, is called a C^r *function* on X.

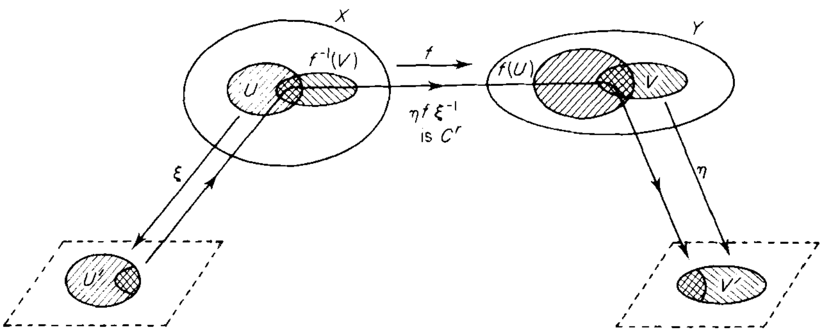


FIGURE A.13

Examples

(A.14) For any C^r manifold X the identity function $id: X \to X$ is a C^r diffeomorphism, and any constant map $c: X \to Y$ onto a point of a C^r manifold Y is a C^r map.

(A.15) Any C^r map $f: U \to V$ of open subsets of Banach spaces is a C^r map when U and V are given their standard C^∞ manifold structures (induced by inclusion).

(A.16) Let $X = Y \times Y$, where Y is a C^r manifold, and let $\delta: Y \to X$, $\tau: X \to X$ and $\pi: X \to Y$ be defined by $\delta(y) = (y, y)$, $\tau(y, z) = (z, y)$ and $\pi(y, z) = y$. Then δ and π are C^r maps, and τ is a C^r diffeomorphism.

(A.17) If Y has the C^r structure induced by a homeomorphism $h: X \to Y$ from a C^r structure on X, then h is a C^r diffeomorphism.

(A.18) Any admissible chart $\xi: U \to U'$ of a C^r manifold X is a C^r diffeomorphism, where U and U' inherit their structures from X and $\mathbf{E}$ via the inclusions.

Exercises

(A.19) Prove that the composite $gf: X \to Z$ of C^r maps $f: X \to Y$ and $g: Y \to Z$ is a C^r map.

(A.20) For any positive integer n, let X_n be the real line $\mathbf{R}$ with the C^∞ structure given by the chart $\xi: \mathbf{R} \to \mathbf{R}$ defined by $\xi(x) = x^{2n-1}$. Show that the map $f_{nm}: X_n \to X_m$ defined by $f_{nm}(x) = x^p$, where $p = (2n-1)/(2m-1)$, is a C^∞ diffeomorphism, but that the identity map $id: \mathbf{R} \to \mathbf{R}$ is a C^1 diffeomorphism if and only if $m = n$.

Any C^s atlas on a topological manifold is trivially a C^r atlas for all r with $r \leqslant s$, and so any C^s structure may be extended to a C^r structure, by adding all C^r charts C^r-compatible with it. Conversely, but non-trivially,

(A.21) Theorem. *Any C^r structure $(r \geqslant 1)$ on a finite dimensional topological manifold contains a C^s structure for all s with $r \leqslant s \leqslant \infty$. Any two such C^s structures are C^s-diffeomorphic.*

For a proof, see Theorem 2.2.9 of Hirsch [1]. Two C^s structures $\mathscr{A}$ and $\mathscr{B}$ on a topological manifold X are *C^s-diffeomorphic* if there is a map $f: X \to X$ that is a C^s diffeomorphism from $(X, \mathscr{A})$ to $(X, \mathscr{B})$. Theorem A.21 encourages us *to restrict our attention to C^∞ manifolds, and we do this in the text of the book.* Note, however, that there are topological manifolds which admit no differentiable structure at all (Theorem A.20 needed $r \geqslant 1$). Kervaire [1] and Smale [1] discovered compact 8-dimensional examples.

As we have seen in Exercise A.20, it is not hard to find different smooth structures on the same topological manifold; to say there that $id: \mathbf{R} \to \mathbf{R}$ is not a diffeomorphism is equivalent to saying that the structures are different. However, it is very much harder, and correspondingly more interesting, to find non-diffeomorphic smooth structures on the same manifold. Nevertheless they do exist. For example there are 28 C^∞ structures on the topological 7-sphere S^7, no two of which are C^1-diffeomorphic (see Milnor [1]).

If $f: X \to Y$ is a C^s map of C^s manifolds, and we extend the C^s structures of the manifolds to C^r structures, for $r < s$, then, of course, we decrease the degree of smoothness of f to C^r. If, conversely, we have a C^r map $f: X \to Y$ of C^r manifolds, and we restrict the C^r structures to C^s structures (as we may by Theorem A.21) we do not usually increase the smoothness of f to C^s; it would be a fluke if we did so. The best we can say about C^r maps $f: X \to Y$ of C^s manifolds is that (in finite dimensions) they may be approximated arbitrarily C^s-closely by C^s maps. See Theorem 2.6 of Hirsch [1]; for an exact statement, one needs to discuss the topology of map spaces, as in Appendix B and Hirsch [1].

III. SMOOTH VECTOR BUNDLES

This section is designed to provide a simple framework for the description of the tangent bundle of a smooth manifold and other associated concepts. The definition follows a pattern that can be used to specify many important structures on a manifold. The essential idea is to consider special types of chart, and to restrict the coordinate change maps to some particular type. We have already done this in defining smooth structures on a manifold.

Let $\mathbf{E}$ and $\mathbf{F}$ be Banach spaces. Let U be open in $\mathbf{E}$ and let $\pi_1: U \times \mathbf{F} \to U$ be projection onto the first factor (i.e. $\pi_1(x, y) = x$). We call π_1 (or sometimes $U \times \mathbf{F}$) a *local vector bundle*. To describe the coordinate change maps that interest us, we need the notion of a *C^r local vector bundle map*. This is a C^r map $\tilde{f}: U \times \mathbf{F} \to V \times \mathbf{H}$ (for V open in $\mathbf{G}$, where $\mathbf{G}$ and $\mathbf{H}$ are Banach spaces) which

(i) covers a C^r map $f: U \to V$, in the sense that the diagram

$$\begin{tikzcd} U \times \mathbf{F} \arrow[r, "\tilde{f}"] \arrow[d, "\pi_1"'] & V \times \mathbf{H} \arrow[d, "\pi_1"] \\ U \arrow[r, "f"] & V \end{tikzcd}$$

commutes, and

(ii) is a C^r-smoothly varying linear map of fibres. That is to say, for all $(x, y) \in U \times \mathbf{F}$,

$$\tilde{f}(x, y) = (f(x), T_x(y)),$$

where $T_x \in L(\mathbf{F}, \mathbf{H})$ and the map $x \mapsto T_x$ from U to $L(\mathbf{F}, \mathbf{H})$ is C^r.

Note that in finite dimensions $\tilde{f}$ C^r automatically implies that $x \mapsto T_x$ is C^r.

Now let X be a C^r manifold modelled on E, let B be a topological space and let $\pi: B \to X$ be a continuous surjection. A *C^r vector bundle chart* (or C^r

local trivialization) on π is, for some open subset U of X a homeomorphism $\tilde{\xi}$ from $\tilde{U}=\pi^{-1}(U)$ to a local vector bundle $U'\times\mathbf{F}$ that covers some admissible chart ξ: $U\to U'$ on X. That is to say the diagram

$$\begin{tikzcd} \tilde{U} \arrow[r, "\tilde{\xi}"] \arrow[d, "\pi"'] & U'\times\mathbf{F} \arrow[d, "\pi_1"] \\ U \arrow[r, "\xi"] & U' \end{tikzcd}$$

commutes. A C^r *vector bundle atlas* for π is, for some indexing set I, a set $\tilde{\mathscr{A}}=\{\tilde{\xi}_i\colon i\in I\}$ of C^r vector bundle charts $\tilde{\xi}_i$: $\tilde{U}_i=\pi^{-1}(U_i)\to U_i'\times\mathbf{F}$ satisfying

(i) $\{U_i\colon i\in I\}$ is a covering of X, and

(ii) for all $i, j\in I$ the map

$$\tilde{\xi}_j\tilde{\xi}_i^{-1}:\tilde{\xi}_i(\tilde{U}_i\cap\tilde{U}_j)\to\tilde{\xi}_j(\tilde{U}_i\cap\tilde{U}_j)$$

is a C^r local vector bundle map. We now proceed as when defining smooth structures on a manifold and say that a C^r vector bundle chart $\tilde{\xi}$ on π is C^r*-compatible* with $\tilde{\mathscr{A}}$ if $\tilde{\mathscr{A}}\cup\{\tilde{\xi}\}$ is a C^r vector bundle atlas. A C^r *vector bundle structure* on π is a maximal C^r vector bundle atlas on π, and a C^r *vector bundle* $(\pi,\tilde{\mathscr{A}})$ is π together with a given C^r *vector bundle structure* $\tilde{\mathscr{A}}$. As usual we abbreviate $(\pi,\tilde{\mathscr{A}})$ by π if the structure $\tilde{\mathscr{A}}$ is unambiguous, and we refer to the elements of $\tilde{\mathscr{A}}$ as (*admissible vector bundle*) *charts* on π. Moreover a further common abuse of language which we sometimes follow is to refer to "the vector bundle B" if the map $\pi\colon B\to X$ is unambiguous. We call π, B, X and $\mathbf{F}$ the *projection*, *total space*, *base* (*space*) and *fibre* of the vector bundle, and say that π is *over* X.

(A.22) Example. If X is any C^r manifold and $\mathbf{F}$ is any Banach space, then $\pi_1\colon X\times\mathbf{F}\to X$ is a C^r vector bundle with base X, fibre $\mathbf{F}$ and an atlas given by charts of the form $\xi\times id\colon U\times\mathbf{F}\to U'\times\mathbf{F}$ where $\xi\in$ an atlas of X. Such a bundle is called the *product* or *trivial* bundle.

(A.23) Example. We may regard any Banach space $\mathbf{F}$ as a smooth vector bundle with base $\{0\}$ and fibre $\mathbf{F}$.

(A.24) Example. The Möbius band $M=\mathbf{R}^2/\mathbf{Z}$ defined in Example A.3 (iii) as a topological 2-manifold is the total space of a non-trivial C^∞ vector bundle with base S^1 and fibre $\mathbf{R}$. The projection $\pi\colon M\to S^1$ takes $[(x, y)]$ to $[x]$, where [ ] denotes the equivalence class under the relations defining $M=\mathbf{R}^2/\mathbf{Z}$ and $S^1=\mathbf{R}/\mathbf{Z}$. A C^∞ vector bundle atlas giving the structure consists of the two vector bundle charts $\tilde{\xi}\colon\tilde{U}\to U'\times\mathbf{R}$ and $\tilde{\eta}\colon\tilde{V}\to V'\times\mathbf{R}$, where $U' =]0, 1[$, $V' =]\frac{1}{2}, \frac{3}{2}[$, $\tilde{U} = \pi^{-1}(U) = [U' \times \mathbf{R}]$, $\tilde{V} = [V' \times \mathbf{R}]$, and $\tilde{\xi}$ and $\tilde{\eta}$ are the maps $[(s, t)] \to (s, t)$. The coordinate change map is $\theta : \tilde{\xi}(\tilde{W}) \to \tilde{\eta}(\tilde{W})$, where $\tilde{W} = \tilde{U} \cap \tilde{V}$, $\theta(s, t) = (s, t)$ for $s \in]\frac{1}{2}, 1[$ and $\theta(s, t) = (1 + s, -t)$ for $s \in]0, \frac{1}{2}[$. It is clear that the bundle is not trivial, since M is not orientable†.

(A.25) Remark. In the case $r = 0$, the condition that the first factor V of a local vector bundle is an open subset of a Banach space is inappropriate and unnecessarily restrictive. The definition works perfectly well with V any topological space. With this modification, we may define a C^0 vector bundle over any topological space as base. Such bundles occur naturally in the theory of dynamical systems (for example the tangent bundle over an exotic basic set X of a dynamical system on a manifold).

Any C^r vector bundle structure $\tilde{\mathscr{A}}$ for $\pi : B \to X$ is certainly a C^r manifold atlas on B, and thus determines a C^r manifold structure on B. (*As usual we restrict attention to B Hausdorff with a countable basis of open sets.*) We always regard B as furnished with this structure. With respect to it, π is a C^r map of manifolds. For all $\tilde{\xi}_i : \tilde{U}_i \to U_i' \times \mathbf{F}$ in $\tilde{\mathscr{A}}$, let Z_i denote the subset $\tilde{\xi}_i^{-1}(U' \times \{0\})$ of B. Since local vector bundle maps are linear on the second factor, it is clear that $Z_i \cap \tilde{U}_j = Z_j \cap \tilde{U}_i$ for all $i, j \in I$. Thus the union of all Z_i, $i \in I$, forms a C^r submanifold, Z say, of B which is termed, rather loosely, the *zero section* of B (see Figure A.26). Clearly π maps Z C^r diffeomorphically onto

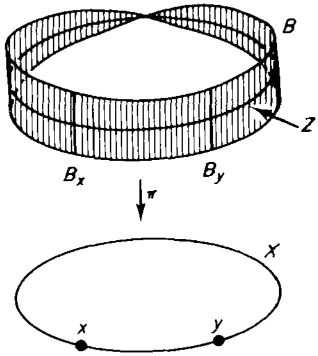


FIGURE A.26

the base X. For each point $x \in X$, the subset $B_x = \pi^{-1}(x)$ is a C^r submanifold of B, C^r diffeomorphic to $\mathbf{F}$, called the *fibre over x*. It inherits a linear structure from $\mathbf{F}$. That is to say, we may perform addition and multiplication by scalars in B_x by mapping to $\{x\} \times \mathbf{F}$ by some chart ξ, performing the corresponding operations in $\{x\} \times \mathbf{F}$ (identified with $\mathbf{F}$) and then mapping

† A finite dimensional manifold is *orientable* if it has an atlas all of whose coordinate change maps have differentials with positive determinant at all points (see Hirsch [1]).

back by ξ^{-1}. By definition of local vector bundle maps the net result does not depend on ξ. Thus B is indeed "a bundle of vector spaces", whence, of course, the name. Notice that although each fibre B_x also inherits via charts a topology from that of $\mathbf{F}$, it does *not* generally inherit a norm from $\mathbf{F}$, since a coordinate change map is not generally an isometry of fibres.

(A.27) Example. Let $\pi: B \to X$ and $\rho: C \to X$ be C^r vector bundles over a common base X, with fibres $\mathbf{F}$ and $\mathbf{G}$ respectively. As we have seen, for each $x \in X$ the fibres B_x and C_x inherit topologies, and so we may consider the topological vector space $L(B_x, C_x)$ of continuous linear maps from B_x to C_x. Let $L(B, C) = \bigcup_{x \in X} L(B_x, C_x)$. Then $L(B, C)$ is the total space of a C^r vector bundle, which we denote $L(\pi, \rho)$, with fibre $L(B_x, C_x)$ over x, defined as follows. Let $\tilde{\xi}: \tilde{U} \to U' \times \mathbf{F}$ and $\tilde{\eta}: \tilde{V} \to V' \times \mathbf{G}$ be charts on π and ρ covering the same chart $\xi: U \to U'$ (i.e. $\xi = \eta$, $U = V$, $U' = V'$). We have an associated chart $\tilde{\zeta}: L(\tilde{U}, \tilde{V}) \to L(U' \times \mathbf{F}, U' \times \mathbf{G}) = U' \times L(\mathbf{F}, \mathbf{G})$ defined, for all $T \in L(B_x, C_x)$ and for all $v \in B_x$, by $\tilde{\zeta}(T)(\tilde{\xi}(v)) = \tilde{\eta}(T(v))$. We may take C^r atlases $\{\tilde{\xi}_i : i \in I\}$ for π and $\{\tilde{\eta}_i : i \in I\}$ for ρ such that, for each $i \in I$, $\tilde{\xi}_i$ and $\tilde{\eta}_i$ have the above property.Tthen the associated atlas $\{\tilde{\zeta}_i : i \in I\}$ is a C^r atlas for $L(\pi, \rho)$. We call $L(\pi, \rho)$ the *linear map bundle* from π to ρ.

(A.28) Example. (*Products and sums of vector bundles*) Let $\pi: B \to X$ and $\rho: C \to Y$ be C^r vector bundles with fibres $\mathbf{E}$ and $\mathbf{F}$ respectively. Then the product $\pi \times \rho: B \times C \to X \times Y$ has a C^r vector bundle structure, with fibre $\mathbf{E} \times \mathbf{F}$, defined as follows. Let $\tilde{\xi}: \tilde{U} \to U' \times \mathbf{E}$ and $\tilde{\eta}: \tilde{V} \to V' \times \mathbf{F}$ be C^r vector bundle charts for π and ρ respectively. Then

$$\tilde{\xi} \times \tilde{\eta}: \tilde{U} \times \tilde{V} \to (U' \times \mathbf{E}) \times (V' \times \mathbf{F}) = (U' \times V') \times (\mathbf{E} \times \mathbf{F})$$

is a vector bundle chart on $\pi \times \rho$, and the set of all such product maps $\tilde{\xi} \times \tilde{\eta}$ forms a C^r vector bundle atlas for $\pi \times \rho$. If $X = Y$, and one restricts $\pi \times \rho$ to the fibres over the diagonal $\{(x, x): x \in X\}$, one has a bundle over X (identified with the diagonal by $x = (x, x)$) called the (*Whitney*) *sum* $\pi \oplus \rho$ of the bundles π and ρ. Its fibre is still, of course, $\mathbf{E} \times \mathbf{F}$, which is canonically isomorphic to $\mathbf{E} \oplus \mathbf{F}$.

Let $\pi: B \to X$ and $\rho: C \to Y$ be C^r vector bundles. A map $\tilde{f}: B \to C$ is a C^r *vector bundle map* if, for all admissible charts $\tilde{\xi}; \tilde{U} \to U' \times \mathbf{E}$ on π and $\tilde{\eta}: \tilde{V} \to V' \times \mathbf{F}$ on ρ, the induced map (or *local representative*)

$$\tilde{\eta}\tilde{f}\tilde{\xi}^{-1}: \tilde{\xi}(\tilde{f}^{-1}(\tilde{V}) \cap \tilde{U}) \to V' \times \mathbf{F}$$

is a C^r local vector bundle map. Thus $\tilde{f}$ is a C^r map of C^r manifolds which maps fibres of π linearly onto fibres of ρ. Note that $\tilde{f}$ maps the zero section Z of π into the zero section T of ρ, and thus induces (via the diffeomorphisms $\pi|Z: Z \to X$ and $\rho|T: T \to Y$) a C^r map $f: X \to Y$ such that the diagram over the page commutes. We say that $\tilde{f}$ *covers*, or is *over*, f.

$$\begin{array}{ccc} B & \xrightarrow{\tilde{f}} & C \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \rho} \\ X & \xrightarrow{f} & Y \end{array}$$

If $\tilde{f}: B \to C$ is a bijective C^r vector bundle map, and $\tilde{f}^{-1}: C \to B$ is also a C^r vector bundle map, then $\tilde{f}$ is said to be a *C^r vector bundle isomorphism*, and π is *C^r vector bundle isomorphic* to ρ.

(A.29) Exercise. Prove that the composite of two C^r vector bundle maps is another C^r vector bundle map.

(A.30) Exercise. Prove that there are, up to C^r vector bundle isomorphism, precisely two C^r vector bundles with base S^1 and fibre $\mathbf{R}^n$ $(n \geq 1)$.

IV. THE TANGENT BUNDLE

Suppose that we are given a C^r manifold X $(r \geq 1)$. We shall associate with each point x of X a vector space T_xX which we may think of as the set of all possible velocities at x of a particle moving on X (see the section on vector fields in Chapter 3). This gives us a vector bundle, with base space X, which is called the *tangent bundle* of X. There is more than one way of constructing T_xX. We do it via charts. Any chart ξ at x takes the path of a moving particle on X to the path of a moving particle in the model space $\mathbf{E}$. We could define the velocity of the first particle at x to be the velocity of the second particle at $\xi(x)$, but this latter depends on the chart ξ. We get round this problem by the ingenious use of an equivalence relation.

Let $\mathscr{A} = \{\xi_i: U_i \to U_i': i \in I\}$ be the C^r structure of X. Consider the subset A of $X \times \mathbf{E} \times I$ given by $A = \{(x, p, i): x \in U_i\}$ and define an equivalence relation $\sim$ on A putting $(x, p, i) \sim (y, q, j)$ if and only if $x = y$ and $q = D\chi(\xi_i(x))(p)$, where $\chi: \xi_i(U_i \cap U_j) \to \xi_j(U_i \cap U_j)$ is the coordinate change map, and $D\chi: \xi_i(U_i \cap U_j) \to L(\mathbf{E})$ is its derivative (see Appendix B), which is C^{r-1}. Let $TX = A/\sim$, and denote the $\sim$ class of (x, p, i) by $[x, p, i]$. Then there is a map $\pi_X: TX \to X$ given by $\pi_X([x, p, i]) = x$. Let $\tilde{U}_i = \pi_X^{-1}(U_i)$, and consider, for all $i \in I$, the set of maps $\tilde{\xi}_i: \tilde{U}_i \to U_i' \times \mathbf{E}$ given by $\tilde{\xi}_i([x, p, i)] = (\xi_i(x), p)$. It follows trivially from the definition of $\sim$ that $\{\tilde{\xi}_i: i \in I\}$ is a C^{r-1} vector bundle atlas for π_X. This determines a C^{r-1} vector bundle structure for π_X with fibre $\mathbf{E}$. With this structure, π_X (or TX) is called the *tangent bundle* of X.

The fibre $(TX)_x$ is called the *tangent space* to X at x. It is usually written as T_xX, or X_x (by abuse of notation). Its points are called *tangent vectors* (*to* X)

at x. It is sometimes useful to think of the base space X of π_X as identified with the zero section of π_X, the point x with the zero vector 0_x of T_xX.

(A.31) Example. If U is an open subset of $\mathbf{E}$, with its usual C^∞ manifold structure, then TU is C^∞ vector bundle isomorphic to the trivial bundle $U\times\mathbf{E}$. An explicit isomorphism is given as follows. Let $\xi_i: U\to\mathbf{E}$ be the inclusion, where $i\in I$. Then any point $v\in TU$ can be written uniquely as $[x, p, i]$ where $x\in U$ and $p\in\mathbf{E}$. Define $\tilde{f}$: $TU\to U\times\mathbf{E}$ by $f(v)=(x,p)$. Then $\tilde{f}$ is a C^∞ vector bundle isomorphism. It is very common to identify TU with $U\times\mathbf{E}$ by this isomorphism.

(A.32) Exercise. Construct a C^∞ vector bundle isomorphism from TS^1 to the trivial bundle $S^1\times\mathbf{R}$.

(A.33) Exercise. Let X and Y be C^r manifolds modelled on $\mathbf{E}$ and $\mathbf{F}$ respectively. By Example A.28, $\pi_X\times\pi_Y$ is a C^{r-1} vector bundle over $X\times Y$. Construct a C^{r-1} vector bundle isomorphism from $\pi_X\times\pi_Y$ to $\pi_{X\times Y}$, the tangent bundle of the product manifold $X\times Y$.

(A.34) Example. (*Parallelizable manifolds*) *A* C^r manifold X modelled on a Banach space $\mathbf{E}$ is said to be *parallelizable* if there is a C^{r-1} vector bundle isomorphism from π_X to the trivial bundle $X\times\mathbf{E}$. Such an isomorphism is called a *trivialization*. Thus any open subset of $\mathbf{E}$ is parallelizable, and so is S^1. The product of two parallelizable manifolds is parallelizable, by Exercise A.33. As we commented in the introduction to the book in connection with the spherical pendulum, S^2 is not parallelizable.

The concept of *derivative*, or *linear approximation map*, is basic to differential calculus, and, when we work on smooth manifolds, it makes its appearance as the *tangent map*. Let X and Y be C^r manifolds modelled on $\mathbf{E}$ and $\mathbf{F}$ respectively, and let $f: X\to Y$ be a C^r map $(r\geqslant 1)$. Let $v\in T_xX$ and let $\xi_i: U_i\to U_i'$ and $\eta_j: V_j\to V_j'$ be charts at x and $f(x)$ respectively. Then we have the local representative $\phi: W\to V_j'$, where $W=\xi_i(U_i\cap f^{-1}(V_j))$ and $\phi\xi_i(x)=\eta_j f(x)$ for all $x\in\xi_i^{-1}(W)$. If $v=[x,p,i]$, we define an element $Tf(v)$ of TY by $Tf(v)=[f(x), q, j]$, where $q=D\phi(\xi_i(x))(p)$. One must check that $Tf(v)$ is independent of choice of charts ξ_i and η_j. This is a routine exercise, and we leave it to the reader, together with the proof of the following result:

(A.35) Proposition. *The map* $Tf: TX\to TY$ *is a* C^{r-1} *bundle morphism. If* $g: Y\to Z$ *is another* C^r *map of manifolds, then* $T(gf)=(Tg)(Tf)$. *For any* X, $T(id_X)=id_{TX}$. □

In the language of category theory, we have constructed a *covariant functor* from the category of C^r manifolds and C^r maps to the category of C^{r-1} vector bundles and C^{r-1} vector bundle maps. We call T the tangent functor. We denote by T_xf the restriction $(Tf)_x: T_xX\to T_{f(x)}Y$ of Tf to a single fibre. It is a continuous linear map of topological vector spaces.

(A.36) Example. If X and Y are open subspaces of Banach spaces $\mathbf{E}$ and $\mathbf{F}$ then we may identify TX and TY with $X \times \mathbf{E}$ and $Y \times \mathbf{E}$ as explained in Example A.31. The derivative $Df: X \to L(\mathbf{E}, \mathbf{F})$ and the tangent map $Tf: TX \to TY$ are related by $Tf(x, p) = (f(x), Df(x)(p))$. The *double tangent map* $T^2f = T(Tf): T(TX) \to T(TY)$ is given by $T^2f((x, p), (u, v)) = ((f(x), Df(x)(p)), (Df(x)(u), D^2f(x)(p, u) + Df(x)(v)))$, where $T(TX)$ is identified with $(X \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E})$.

If U is an open subspace of a C^r manifold X and $\iota: U \to X$ is the inclusion, then the C^{r-1} vector bundle map $T\iota: TU \to TX$ maps TU bijectively onto the open subset $\pi_X^{-1}(U)$ of TX. One customarily identifies TU with its image under $T\iota$. If $\xi_i: U_i \to U_i'$ is any C^r admissible chart on X then the map $T\xi_i: TU_i \to TU_i' = U_i' \times \mathbf{E}$ is a C^{r-1} admissible chart on TX. Modulo the identifications, it is precisely the chart $\tilde{\xi}_i$ in the definition of TX.

(A.37) Example. If $I = [a, b]$ is a real interval and $\gamma: I \to X$ is a smooth map†, then γ is called a *curve* on X. The tangent space TI is identified with $I \times \mathbf{R}$, and for all $t \in I$, 1_t denotes the element $(t, 1)$ of TI. The vector $T\gamma(1_t)$ in the tangent space $T_{\gamma(t)}X$ is called the *velocity* of γ at (time) t. We usually abbreviate $T\gamma(1_t)$ to $\gamma'(t)$. If $X = \mathbf{E}$ and we identify $T_{\gamma(t)}X$ with $\mathbf{E}$, this usage fits in with the standard notation of differential calculus.

(A.38) Example. (*The canonical involution*) If $\pi: B \to Y$ is a C^r vector bundle and if $f: B \to C$ is a bijection onto a set C, then f induces on C a C^r vector bundle structure such that f is a C^r vector bundle isomorphism. If C has from the outset its own C^r manifold structure (and in particular if $C = B$) then we are usually interested in maps f for which the induced C^r structure is the original one. Trivially this is so if and only if f is a C^r diffeomorphism with respect to the original structure on C.

An example of this occurs when $B = C = T^2X = T(TX)$, for any C^r manifold X ($r \geq 2$). In forming $T(TX)$ from X, we, in effect, twice take tangents to X; firstly in forming TX, secondly in taking tangents to TX (for we envisage X as embedded in TX as the zero section). Thus T^2X has a built in symmetry (to use the term rather loosely), and there is a C^{r-2} *involution* (= diffeomorphism of period 2) that exchanges the two "tangent spaces to X" at each point x. This map is called the *canonical involution* of T^2X. We now give it a precise description.

We continue with the notations in the definition of TX. Let $\mathscr{B} = \{\tilde{\xi}_i : i \in I\}$ be the atlas of TX corresponding to the atlas $\mathscr{A} = \{\xi_i : i \in I\}$ of X. Points of $T(TX)$ are of the form $[[x, p, i], (q, r), j]$, where $j \in I$, $(q, r) \in \mathbf{E}^2$ and $[x, p, i]$ is a point of TX in the domain of $\tilde{\xi}_j$. The outer brackets denote equivalence class with respect to $\sim$ in the construction of T^2X. We may write any such point as $[[x, p, i], (u, v), i]$ for some $(u, v) \in \mathbf{E}^2$, since certainly $[x, p, i] \in \tilde{U}_i$.

† That is to say, γ can be extended to a smooth map of some open interval J with $I \subset J$.

The canonical involution f takes this point to $[[x, u, i], (p, v), i]$. Checking that f is well defined and a C^{r-2} diffeomorphism is a useful exercise. Now recall that the C^{r-1} map π_X takes $[x, p, i]$ to x. Thus we have two important maps from T^2X to TX, namely π_{TX} and $T(\pi_X)$. The first takes $[[x, p, i], (u, v), i]$ to $[x, p, i]$, and the second, as may easily be verified, takes it to $[x, u, i]$. Thus the diagram

$$\begin{tikzcd} T^2X \arrow[rr, leftrightarrow, "f"] \arrow[dr, "\pi_{TX}"'] & & T^2X \arrow[dl, "T(\pi_X)"] \\ & TX & \end{tikzcd}$$

commutes.

V. IMMERSIONS, EMBEDDINGS AND SUBMERSIONS

Let $f: X \to Y$ be a C^r map of C^r manifolds ($r \geq 1$). For each $x \in X$, the tangent map $T_xf: T_xX \to T_yY$, where $y = f(x)$, is continuous linear. Its kernel ker T_xf and its image im T_xf are linear subspaces of T_xX and T_yY respectively. Ker T_xf is automatically a closed subspace; im T_xf is not necessarily closed unless Y is finite dimensional. We say that f is *immersive* at x if T_xf is injective and im T_xf splits T_yY. Dually, f is *submersive* at x if T_xf is surjective and ker T_xf splits T_xX. We say that f is an *immersion* if it is immersive at x for all $x \in X$, and a *submersion* if it is submersive at x for all $x \in X$.

(A.39) ***Note.*** If im T_xf (resp. ker T_xf) has finite dimension or finite codimension in T_yY (resp. T_xX) then splitting is automatic. This is because, firstly, any finite dimensional subspace or closed finite codimensional subspace splits any Banach space and, secondly, any finite codimensional image of a continuous linear map is closed (see Lang [2]). In particular, one may omit the splitting condition from the above definitions when either X or Y is finite dimensional.

(A.40) Example. The C^∞ map $f: S^1 \to S^1$ given by $f([x]) = [nx]$ for all $[x] \in S^1 = \mathbf{R}^2/\mathbf{Z}$, where n is any integer, is an immersion (and also a submersion) if and only if $n \neq 0$.

(A.41) Example. For any C^r vector bundle $\pi: B \to X$, π is a C^r submersion. In particular, if X is a C^r manifold, π_X is a C^{r-1} submersion for $r \geq 2$.

The image of an injective C^r immersion is sometimes called a C^r *immersed submanifold*. This is an abuse of language, since it need not be a topological submanifold; think of the numeral 6 regarded as the image of a C^∞ immersion of $\mathbf{R}$ in $\mathbf{R}^2$. If $f: X \to Y$ is an injective immersion, we say that im $T_x f$ is the *tangent space to the immersed submanifold* at $f(x)$. An injective C^r immersion $f: X \to Y$ whose image is a C^r submanifold of Y is called a C^r *embedding.* Any C^r immersion is locally a C^r embedding, by Exercise C.12 of Appendix C.

(A.42) Example. In irrational flow on the torus (see Examples 1.25 and 2.9) the map: $\mathbf{R} \to T^2$ taking t to $([x+t],[y+\theta t])$ is an injective immersion but not an embedding.

(A.43) Example. The inclusion of a C^r submanifold of X in X is a C^r embedding.

(A.44) Example. If X and Y are C^r manifolds, and we define $i: X \to X \times Y$ by $i(x) = (x, y)$ for some given $y \in Y$, then i is an embedding. Similarly the diagonal map $x \mapsto (x, x)$ from X to $X \times X$ is an embedding.

(A.45) Exercise. Prove that $f: X \to Y$ is a C^r embedding ($r \geq 1$) if and only if it is a C^r immersion and a *topological embedding* (i.e. maps X homeomorphically onto $f(X)$).

(A.46) Example. (*Foliations and laminations*) Let X be a C^r manifold modelled on $\mathbf{E}$ and let $\mathbf{E} = \mathbf{F} \times \mathbf{G}$ be a splitting of $\mathbf{E}$. A C^r *foliation* is a disjoint decomposition of X into C^r injectively immersed submanifolds, called *leaves*, satisfying the following condition: There is an admissible atlas of charts of the form $\xi: U \to \mathbf{F} \times \mathbf{G}$, called *foliation boxes*, such that, for all $y \in \mathbf{G}$, $\xi^{-1}(\mathbf{F} \times \{y\})$ is contained in a single leaf, and is the image of an open set under the injective immersion giving the leaf (see Figure A.46). The

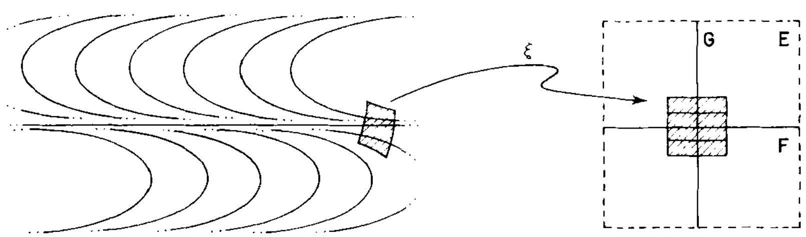


FIGURE A.46

dimensions of $\mathbf{F}$ and $\mathbf{G}$ are called respectively the *dimension* and *codimension* of the foliation. It is also possible to give an equivalent purely local definition of foliation by considering a maximal atlas of C^r admissible charts for which all coordinate change maps χ have the property $\chi_2(x, y) = \chi_2(x', y')$ if and only if $y = y'$, for all (x, y) and (x', y') in $\mathbf{F} \times \mathbf{G}$.

Rational and irrational flows both give 1-dimensional C^∞ foliations of the torus (the leaves being the orbits of the flow). As we have seen in Chapter 6, the stable manifold of a hyperbolic closed orbit of a C^r flow is C^r foliated by the stable manifolds of the individual points of the orbit. This situation is not typical of hyperbolic sets in general. Usually the stable manifolds of the individual points only C^r *laminate* the stable manifold of the set. To define this notion, we regard $\mathbf{E}$ as a trivial vector bundle with base $\mathbf{G}$ and fibre $\mathbf{F}$, and weaken the definition of C^r foliation by relaxing the condition that the foliation boxes ξ are C^r admissible charts and insisting only that they are homeomorphisms with F^r inverses (see Appendix B for F^r maps). Of course they must still form a topological atlas for X and satisfy the foliation box condition.

If $f: X \to Y$ is a C^r map of manifolds, and, for some $y \in Y$, f is submersive at every point of $f^{-1}(y)$, then $f^{-1}(y)$ is a disjoint union of C^r submanifolds of X, with dimension $\dim X - \dim Y$ if this makes sense. This result generalizes the remarks of Example A.6 and is, again, a consequence of the inverse mapping theorem. More generally still, let W be a C^r-immersed submanifold of Y. We say that f is *transverse* to W, written $f \pitchfork W$, if for all $y \in f(X) \cap W$, $T_y(Y) = W_y + \mathrm{im}\, T_x f$, where $y = f(x)$ and W_y is the tangent space to W at y, and $(T_x f)^{-1}(W_y)$ splits $T_x X$. In this case, $f^{-1}(W)$ is a C^r submanifold of X whose codimension equals the codimension of W in Y. If V and W are two C^r immersed submanifolds of Y, we say that V is transverse to W, written $V \pitchfork W$, if some injective C^r immersion f with image V is transverse to W. It follows that $V \cap W$ is a C^r immersed submanifold.

VI. SECTIONS OF VECTOR BUNDLES

Let $\pi: B \to X$ be a C^r vector bundle. A map $\sigma: X \to B$ such that $\pi\sigma: B \to B$ is the identity on B is called a *section* of π. Thus σ is a section if and only if it maps every point x of X into the fibre B_x over x. Figure A.47 illustrates this

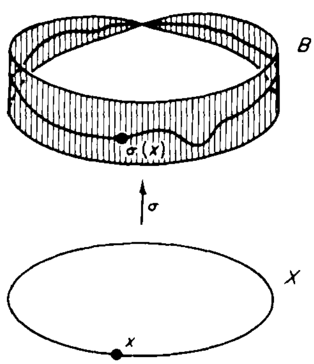


FIGURE A.47

idea in the case when B is the Möbius band (Example A.24). When π is the tangent bundle $\pi_X: TX \to X$, sections are called *vector fields* on X.

In the case of the trivial bundle $\pi_1: X \times \mathbf{F} \to X$, the sections of π_1 are precisely all maps $(id, f): X \to X \times \mathbf{F}$, and, of course, the section (id, f) is C^r if and only if the map $f: X \to \mathbf{F}$ is C^r. The map f is called the *principal part* of the section. For a general vector bundle $\pi: B \to X$ we may trivialize the situation locally, using an admissible chart. Let $\tilde{\xi}: \tilde{U} \to U' \times \mathbf{F}$ be such a chart, so that the diagram

$$\begin{tikzcd} \tilde{U} \arrow[r, "\tilde{\xi}"] \arrow[d, "\pi"'] & U' \times F \arrow[d, "\pi_1"] \\ U \arrow[r, "\xi"] & U' \end{tikzcd}$$

commutes. Then the section σ of π, when restricted to U, induces a section, τ say, of the trivial bundle π_1, defined by commutativity of the diagram

$$\begin{tikzcd} \tilde{U} \arrow[r, "\tilde{\xi}"] & U' \times F \\ U \arrow[u, "\sigma|U"] \arrow[r, "\xi"] & U' \arrow[u, "\tau"'] \end{tikzcd}$$

In this case τ is said to be a *local representative* of σ.

(A.48) Exercise. Prove that any C^r section of a vector bundle is a C^r embedding.

(A.49) Example. (*The spherical pendulum*) In the introduction to the book we discussed the spherical pendulum, and found that its motion could be modelled by a C^∞ vector field v on the C^∞ manifold $X = TS^2$ of dimension 4. If U is the complement in S^2 of a single meridian of longitude, then there is an admissible chart $\xi: U \to U'$ on S^2 given by $\xi(y) = (\theta, \phi)$ where θ and ϕ are the Euler angles. Correspondingly there are admissible charts $\tilde{\xi}: \tilde{U} \to U' \times \mathbf{R}^2$ for X and $\tilde{\tilde{\xi}}: \tilde{\tilde{U}} \to (U' \times \mathbf{R}^2) \times R^4$ for TX, where $\tilde{U} = \pi_{S^2}^{-1}(U)$ and $\tilde{\tilde{U}} = \pi_X^{-1}(\tilde{U})$. In terms of these local coordinates, the vector field $v: X \to TX$ has local representative

$$(\theta, \phi, \lambda, \mu) \to ((\theta, \phi, \lambda, \mu), p)$$

where $(\theta, \phi, \lambda, \mu) \in U' \times \mathbf{R}^2$ and $p \in \mathbf{R}^4$ is the point

$$(\lambda, \mu, \mu^2 \sin\theta \cos\theta + g \sin\theta, -2\lambda\mu \cot\theta).$$

(A.50) Example. (*Second order equations and sprays*) Notice that, for the vector field v of Example A.49, the first two coordinates of the principal part of $v(x)$ are the same as the last two coordinates of X, namely (λ, μ). This came about because we originally converted a system of second order equations on U' into a system of first order equations on $U' \times \mathbf{R}^2$ by the substitution $\theta' = \lambda$, $\phi' = \mu$. Since, as we commented in Note 3.17, this is a standard procedure, we ought to analyse the situation a bit further.

Let $X = TM$, where M is a C^r manifold $(r \geq 2)$. *A first order ordinary differential equation* on M is just a vector field on M. A *second order ordinary differential equation* on M is a vector field v on X satisfying $T(\pi_M)v = id_X$. This is precisely the coordinate free generalization of the condition in the last paragraph. If $\xi: U \to U'$ is a chart on M, with corresponding charts $\tilde{\xi}: \tilde{U} \to U' \times \mathbf{E}$ on X and $\tilde{\tilde{\xi}}: \tilde{\tilde{U}} \to (U' \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E})$ on TX, and if $f: U' \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}$ is the principal part of the local representative of v, then the condition says that the first coordinate of $f(x, y)$ is y. Now let $\gamma: I \to X$ be any integral curve of v, and let $\delta: I \to M$ be the projection of γ onto M by π_M. Thus $\delta = \pi_M \gamma$. We call δ a *solution of the second order equation*. Differentiating at $t \in I$, we deduce that

$$\delta'(t) = T\pi_M(\gamma'(t)) = T\pi_M(v\gamma(t)) = \gamma(t).$$

That is to say, the velocity of the curve δ at t is the value of the curve γ at t. See Figure A.50. There is no reason why δ should not have self intersections.

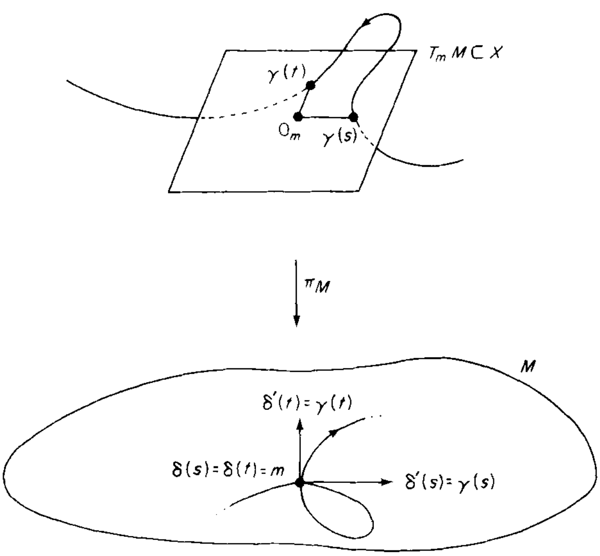


FIGURE A.50

Uniquenesss of the integral curve of v at a given point of X corresponds to uniqueness of the solution δ of the second order equation through a given point of M with a given velocity at $t = 0$.

Concentrating on a particular point $m \in M$ for a moment, we have infinitely many solutions δ_x starting at m at $t = 0$, one for each possible velocity x at m. In fact we have infinitely many starting off in a given direction with various speeds, and there is no reason why these should be in any way related (see Figure A.51(i)). However, given $x \in T_mM$ and $a \in \mathbf{R}$, there is a very natural way of obtaining from δ_x a curve with velocity ax at m at time $t = 0$, and that is by speeding δ_x up by a factor a. That is to say, the curve δ_{ax} defined by $\delta_{ax}(t) = \delta_x(at)$ has the required property (see Figure A.51(ii)). Note that for $a \neq 0$, δ_x and δ_{ax} have the same image. It is a particularly nice situation when the solution curves δ fit together in this way.

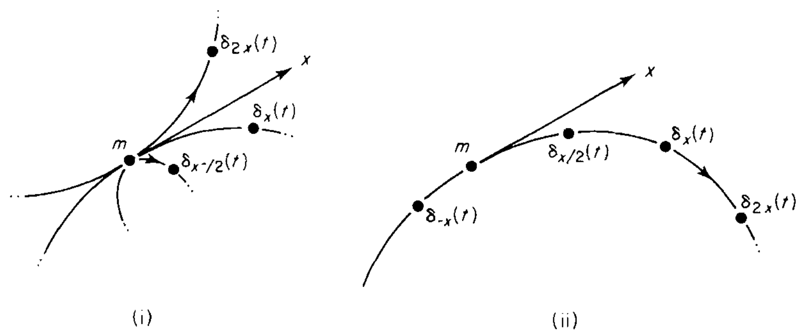


FIGURE A.51

When they do so, we call the second order differential equation a *spray*, and the images of the solution curves *geodesics* of the spray.

Let us find the condition on v that makes it a spray. We require that if $\delta'_{ax}(0) = a\delta'_x(0) = ax$, then $\delta_{ax}(t) = \delta_x(at)$. Differentiating this latter relation gives $\delta'_{ax}(t) = a\delta'_x(t)$. Since v is a second order equation, we may restate this as $\gamma_{ax}(t) = a\gamma_x(at)$. To find the condition on v, we differentiate again. This not completely straightforward, because the outside a is scalar multiplication on the vector bundle TM; it affects the fibre direction but not the "zero section direction" (if we could make sense of this at points of TM which are not on the zero section). However, if we denote by $\mu_a : TM \to TM$ scalar multiplication by a on the fibres, then we can write $\gamma_{ax}(t) = \mu_a\gamma_x(at)$ and obtain $\gamma'_{ax}(t) = T\mu_a(a\gamma'_x(at))$. Putting $t = 0$, we get, for all $x \in TM$ and $a \in \mathbf{R}$,

(A.52) $$v(ax) = aT\mu_a(v(x)).$$

Conversely, if v is a C^1 second order differential equation on M satisfying (A.52) then, by the uniqueness theorem (Theorem 3.34) its integral curves satisfy $\gamma_{ax} = a\gamma_x(at)$, and hence the solution curves satisfy $\delta_{ax}(t) = \delta_x(at)$ as required.

Let v be a C^1 spray on M. By (A.52), 0_m is a zero of v, for all $m \in M$. Thus, by Theorem 3.22, there is some neighbourhood of 0_m in TM such that, for all x in it, integral curves of v at x are defined on $[0, 1]$. Thus, for some neighbourhood N of the zero section of TM, there is a map $\exp: N \to M$ defined by $\exp x = \delta_x(1)$. This is called the *exponential* map of the spray. It is C^r when v is C^r. For all $m \in M$, the restriction $\exp_m$ of exp to $N \cap T_mM$ maps rays through the origin onto geodesics at m, since $\exp_m(tx) = \delta_{tx}(1) = \delta_x(t)$. Moreover, since $d/dt(\exp_m(tx)) = \delta'_x(t) = x$ at $t = 0$, the derivative of $\exp_m$ at 0_m is the identity map on T_mM. More precisely, if we identify $T(T_mM)$ with $T_mM \times T_mM$ in the standard way, then $T\exp_m: T(T_mM) \to TM$ satisfies $T\exp_m(0_m, x) = x$ for all $x \in T_mM$. It follows from the inverse function theorem that $\exp_m$ maps some neighbourhood of 0_m in T_mM diffeomorphically onto a neighbourhood of m in M.

(A.53) Example. If U is an open subset of R^n then the *standard spray* $v: U \times \mathbf{R}^n = TU \to T^2U = (U \times \mathbf{R}^n) \times (\mathbf{R}^n \times \mathbf{R}^n)$ is given by $v(m, x) = (m, x, x, 0)$. The integral curve of v at (m, x) is $t \mapsto (m + tx, x)$ and the solution curve through m with velocity (m, x) is $t \mapsto m + tx$. The exponential map is given by $\exp(m, x) = m + x$.

(A.54) Example. The circle $S^1 = \mathbf{R}/\mathbf{Z}$ inherits its *standard spray* from the standard spray of $\mathbf{R}$ by the covering map $m \to [m]$. Thus $v([m], x) = ([m], x, x, 0)$, where we write TS^1 as $S^1 \times \mathbf{R}$. The exponential map is $\exp([m], x) = [m + x]$. In terms of the standard embedding of S^1 in $\mathbf{R}^2$, exp maps each tangent space round the circle, taking the tangent vector x at p to the point $p\, e^{2\pi i x}$. For higher dimensional spheres S^n, the geodesics are great circle arcs which may be derived from the above example by taking 2-plane sections through $0 \in \mathbf{R}^{n+1}$.

A *zero* of a section $\sigma: X \to B$ is a point $x \in X$ such that $\sigma(x)$ is the zero 0_x of the vector space B_x. The *zero section* is the section defined by $\sigma(x) = 0_x$ for all $x \in X$ (but, as we commented earlier, the phrase is sometimes used loosely for the image of the map). A section without zeros is said to be *nowhere zero*. For many vector bundles, nowhere zero sections do not exist. For example, the reader can soon verify experimentally that the Möbius band has no nowhere zero sections. Neither does the tangent bundle TS^{2n}, by the Poincaré–Hopf index theorem (Theorem 5.79).

The set $C^r(\pi)$ of all C^r sections of π has algebraic structures generated by the linear structure of the fibres. It is a real vector space, with structure defined pointwise by

$$(a\sigma + b\tau)(x) = a\sigma(x) + b\tau(x)$$

for $\sigma, \tau \in C^r(\pi)$, $a, b \in \mathbf{R}$ and $x \in X$. Moreover it is a module over the ring $\mathscr{F}^r(X)$ of C^r functions on X, when we put

$$(f\,.\,\sigma + g\,.\,\tau)(x) = f(x)\,.\,\sigma(x) + g(x)\,.\,\tau(x)$$

for $f, g \in \mathscr{F}^r(X)$. We shall discuss the space $C^r(\pi)$ further in Appendix B, and give it a natural Banach space structure in certain circumstances. When π is the tangent bundle of X, we denote $C^r(\pi)$ by $\Gamma^r(X)$.

VII. TENSOR BUNDLES

The total space of a C^k vector bundle $\pi: B \to X$ is partitioned into a set of fibres B_x. Each fibre B_x has an associated dual topological vector space $B_x^* = L(B_x, \mathbf{R})$. If we cobble these dual spaces together in a natural way determined by the structure of π, we get a new C^k vector bundle, called the dual π^* of π. Similarly we can, for each fibre B_x, form the space $L_r^s(B_x, \mathbf{R})$ of all real $(r+s)$-linear functions on $(B_x)^r \times (B_x^*)^s$, which we call *tensors of type* (r, s) and we can make these into a C^k vector bundle which we denote by π_r^s. In particular, π^* is π_1^0.

We now describe the structure of π_r^s. We first observe that any (topological) linear automorphism $f: \mathbf{F} \to \mathbf{G}$ of Banach spaces determines a linear automorphism $f_r^s: L_r^s(\mathbf{F}, \mathbf{R}) \to L_r^s(\mathbf{G}, \mathbf{R})$ by the formula

$$f_r^s(T)(p_1, \ldots, p_r, q_1, \ldots, q_s) = T(f^{-1}(p_1), \ldots, f^{-1}(p_r), q_1 f, \ldots, q_s f),$$

where $(p_1, \ldots, p_r, q_1, \ldots, q_s) \in \mathbf{G}^r \times (\mathbf{G}^*)^s$. For any subset U of X, let $L_r^s(\tilde{U}, \mathbf{R})$ denote the disjoint union of $L_r^s(B_x, R)$ for all fibres B_x of π with $x \in U$. Here, as usual, $\tilde{U} = \pi^{-1}(U)$. Let $\pi_r^s: L_r^s(B, \mathbf{R}) \to X$ send $L_r^s(B_x, \mathbf{R})$ to x. Any admissible chart $\tilde{\xi}: \tilde{U} \to U' \times \mathbf{F}$ on π has the form $\tilde{\xi}(u) = (\xi(x), f_x(u))$ where $\pi(u) = x$, and $f_x: B_x \to \mathbf{F}$ is a linear automorphism. We define a map $\tilde{\xi}_r^s: L_r^s(\tilde{U}, \mathbf{R}) \to U' \times L_r^s(\mathbf{F}, \mathbf{R})$ by $\tilde{\xi}_r^s(T) = (\xi(x), (f_x)_r^s(T))$, where $x = \pi_r^s(T)$. We may topologize $L_r^s(B, \mathbf{R})$ by insisting that for all admissible charts $\tilde{\xi}$ on π the maps $\tilde{\xi}_r^s$ are homeomorphisms, and it is easy to check that the $\tilde{\xi}_r^s$ then form a C^k vector bundle atlas for π_r^s. We call π_r^s, with the C^k vector bundle structure determined by this atlas, the *bundle of tensors on* π *of type* (r, s) (*or covariant of order r, contravariant of order s*). A section of π_r^s is called a *tensor field of type* (r, s) *on* π.

(A.55) Example. (*The derivative of a smooth function*) The dual π_X^* of the tangent bundle of a smooth manifold is called the *cotangent bundle* of X. The total space is denoted by T^*X, the fibre T_x^*X (or X_x^*) is called the *cotangent space* at x and its elements are called *cotangent vectors*. Sections of π_X^* are called 1-*forms* on X. If $f: X \to \mathbf{R}$ is a C^r map ($r \geqslant 1$), then $T_x f$ maps $T_x X$ linearly to $T_{f(x)}\mathbf{R} = \{f(x)\} \times \mathbf{R}$. Thus if we identify $\{f(x)\} \times \mathbf{R}$ with $\mathbf{R}$ in the obvious way, $T_x f$ becomes a cotangent vector at x. The C^{r-1} map from X to T^*X taking x to $T_x f$ is a 1-form on X, and we call it the *derivative* Df of f. If X is an open subset of $\mathbf{E}$, and we identify the tangent space $T_x X = \{x\} \times \mathbf{E}$ with $\mathbf{E}$, then Df is the derivative in the usual sense (see Example A.36).

VIII. RIEMANNIAN MANIFOLDS

As we commented earlier, a vector bundle does not usually come equipped with an inner product, or even a norm, on each individual fibre. However these may be given to the bundle as extra structure, and they then give rise to extra theory. In the case of the tangent bundle of a manifold, they enable us to define lengths of curves on the manifold and also to give the manifold a natural metric space structure.

A C^r *Riemannian structure*, or *Riemannian metric*, on a vector bundle $\pi: B \to X$ is a C^r section $\rho: X \to L_2(B, \mathbf{R})$ of the tensor bundle π_2^0 such that the bilinear form $\rho(x)$ on B_x is symmetric and positive definite for all $x \in X$. One usually writes $\rho(x)(p, q)$ as $\langle p, q\rangle_x$ when there is no doubt as to which metric ρ is in use. A C^r Riemannian metric gives rise to a C^r *Finsler*, which is a norm $|\ |_x$ on the fibre B_x that depends on x in a C^r fashion. Of course, $|\ |_x$ is defined by $|p|_x = \sqrt{\langle p, p\rangle_x}$.

(A.56) Example. Any trivial bundle $\pi_1: X \times \mathbf{H} \to X$, where X is a smooth manifold and $\mathbf{H}$ is a Hilbert space, has a Riemannian structure given by the inner product of $\mathbf{H}$. In particular if U is an open subset of $\mathbf{R}^n$, the tangent bundle of U has a Riemannian structure given by the standard inner product on $\mathbf{R}^n$. That is to say

$$\langle (x, y), (x, y')\rangle_x = y_1 y_1' + \cdots + y_n y_n'$$

for all (x, y) and (x, y') in $TU = U \times \mathbf{R}^n$.

A Riemannian metric on the tangent bundle of a smooth manifold X is also said to be a *Riemannian metric* on X. A C^{r+1} *Riemannian manifold* is a C^{r+1} manifold X together with a C^r Riemannian metric on X. The metric of Example A.56 is the *standard Riemannian metric* on an open subset of $\mathbf{R}^n$. One may construct a C^r Riemannian metric on any finite dimensional C^{r+1} manifold X, or, indeed, on any C^{r+1} manifold X modelled on a Hilbert space provided that X admits C^{r+1} partitions of unity.

(A.57) Example. (*Gradient vector fields*) A C^r Riemannian metric on a manifold X enables us to convert C^r 1-forms on X to C^r vector fields on X and vice versa. The former operation is called *sharpening*, the latter *flattening*. By the representation theorem for Hilbert spaces, given any element λ of T_x^*X, there exists a unique element p of T_xX such that $\lambda(q) = \langle p, q\rangle_x$ for all $q \in T_xX$. Conversely any $p \in T_xX$ gives rise to an element λ of T_x^*X defined by the given formula. We thus obtain inverse C^r vector bundle isomorphisms $\rho_\#: T^*X \to TX$ and $\rho_\flat: TX \to T^*X$, and, correspondingly, isomorphisms $\tilde{\rho}_\#: \Omega^r X \to \Gamma^r X$ and $\tilde{\rho}_\flat: \Gamma^r X \to \Omega^r X$, where $\Omega^r X$ is the Banach space of C^r 1-forms on X, and

$$\tilde{\rho}_\#(w) = \rho_\# w, \qquad \tilde{\rho}_\flat(v) = \rho_\flat v.$$

In particular, if $f: X \to \mathbf{R}$ is a C^{r+1} function on X, then the derivative Df is a C^r 1-form on X (see Example A.55) and $\tilde{\rho}_\#(Df)$ is a C^r vector field on X called the *gradient* ∇f of f. For more detail, see Example 3.3 of the text.

(A.58) Exercise. Let $f: X \to \mathbf{R}$ be C^{r+1}. Prove that ∇f is orthogonal to the contours of f, in the sense that, for all $x \in X$ and $p \in \ker T_xf$, $\langle p, \nabla f(x)\rangle_x = 0$.

Let X be a C^{r+1} Riemannian manifold, for sufficiently large r ($r \geq 4$ covers all eventualities). The Riemannian structure on X gives rise to a C^{r-1} spray σ on X, called the *Riemannian spray*, defined as follows. Let $\xi: U \to V$ be an admissible chart on X, so that $T\xi: TU \to TV = V \times \mathbf{E}$ is an admissible chart on TX. We transfer the Riemannian structure to V by the chart. That is to say, we define a function $K: V \times \mathbf{E} \times \mathbf{E} \to \mathbf{R}$ by

$$\textbf{(A.59)} \qquad K(y, p, q) = \langle T\xi^{-1}(y, p), T\xi^{-1}(y, q)\rangle_{\xi^{-1}(y)}$$

so that, for fixed $y \in V$, $K(y, ., .)$ is a symmetric positive definite bilinear function on $\mathbf{E} \times \mathbf{E}$. We wish to define a local representative τ of σ with respect to the chart $T^2\xi: T^2U \to T^2V = (V \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E})$ on T^2X. We do so by the formula

$$\tau(y, p) = (y, p, p, v),$$

where $v \in E$ satisfies, for all $u \in \mathbf{E}$, the formula

$$\textbf{(A.60)} \qquad K(y, u, v) = (\tfrac{1}{2}TK(y, p, p, u, 0, 0) - TK(y, p, u, p, 0, 0))_2,$$

the subscript denoting the second coordinate in $T\mathbf{R} = \mathbf{R} \times \mathbf{R}$. This, by the representation theorem for Hilbert spaces, gives rise to a uniquely defined element $v \in \mathbf{E}$, and it is clear from the formula that the second order equation τ satisfies the condition (A.52) for a spray, using the bilinearity of $K(y, ., .)$. Obviously the left-hand side of (A.60) is connected with the derivative of $\langle\ ,\ \rangle_x$ with respect to x, but the geometrical motivation for its detailed expression is less clear. A very natural, but rather subtle, explanation of it appears in Lang [1]. However, this requires more of the machinery of differential forms than we are prepared to introduce here. We content ourselves by verifying that σ is well defined, in that it does not depend on the choice of chart ξ.

Let $\xi': U' \to V'$ be another chart on X, and let $\chi: \xi(U \cap U') \to \xi'(U \cap U')$ be the coordinate change map. The map K of (A.59) and the corresponding map $K': V' \times \mathbf{E} \times \mathbf{E} \to \mathbf{R}$ induced by ξ' are related on the overlap by

$$K(y, p, q) = K'(y', p', q') \tag{A.61}$$

where $y' = \chi(y)$, $p' = D\chi(y)(p)$ and $q' = D\chi(y)(q)$. We have to show that $T^2\chi$ takes $\tau(y, p)$ to (y', p', p', v') where v' satisfies (A.60)$'$, which is (A.60) with K and all the variables dashed. Now, by Example A.36,

$$T^2\chi(y, p, p, v) = (y', p', p', D^2\chi(y)(p, p) + D\chi(y)(v)).$$

Moreover, differentiating (A.61),

$$TK'(y', p', q', u', 0, 0) = TK(y, p, q, u, -D\chi(y)^{-1}D^2\chi(y)(p, u), \\ -D\chi(y)^{-1}D^2\chi(y)(q, u)),$$

where $u' = D\chi(y)(u)$. Thus

$$\begin{aligned} &\tfrac{1}{2}TK'(y', p', p', u', 0, 0) - TK'(y', p', u', p', 0, 0) \\ &\quad = \tfrac{1}{2}TK(y, p, p, u, -D\chi(y)^{-1}D^2\chi(y)(p, u), -D\chi(y)^{-1}D^2\chi(y)(p, u)) \\ &\qquad - TK(y, p, u, p, -D\chi(y)^{-1}D^2\chi(y)(p, p), -D\chi(y)^{-1}D^2\chi(y)(u, p)). \end{aligned}$$

By (A.60) and the bilinearity of $K(y, ., .)$, the second coordinate of this expression reduces to

$$\begin{aligned} &K(y, u, v) - \tfrac{1}{2}K(y, p, D\chi(y)^{-1}D^2\chi(y)(p, u)) \\ &\qquad\qquad - \tfrac{1}{2}K(y, D\chi(y)^{-1}D^2\chi(y)(p, u), p) \\ &\qquad + K(y, p, D\chi(y)^{-1}D^2\chi(y)(u, p)) + K(y, D\chi(y)^{-1}D^2\chi(y)(p, p), u) \\ &\quad = K(y, u, v + D\chi(y)^{-1}D^2\chi(y)(p, p)) \\ &\quad = K'(y', u', v') \end{aligned}$$

as required.

By the theory of sprays (Example A.50) we now have a C^{r-1} exponential map $\exp: N \to X$ where N is some neighbourhood of the zero section in TX. We call this the *Riemannian exponential map.*

The Riemannian metric on X gives rise to a metric on X in the sense of metric space theory. We denote this by d, and always call it the *Riemannian distance function*, to avoid confusion. We define $d(x, x')$ for $x, x' \in X$ as follows. Since X is connected, we may join x to x' by *piecewise C^1 curve* $\gamma: I \to X$ (that is to say, I is a real interval $[a, b]$, γ is a continuous map and is C^1 on each subinterval $[a_r, a_{r+1}]$ of some subdivision

$$a = a_0 < a_1 < \cdots < a_n = b).$$

Since $|\gamma'(t)|_{\gamma(t)}$ has at worst a finite number of jump discontinuities on $[a, b]$, the length of γ, $\int_a^b |\gamma'(t)|_{\gamma(t)}\, dt$, exists finite, and we define $d(x, x')$ to be the infimum of the lengths of γ as γ ranges over all piecewise C^1 curves joining x to x'. Trivially d is symmetric and satisfies the triangle inequality, and $d(x, x') = 0$. To complete a proof that (X, d) is a metric space, we need to show that $d(x, x) = 0$ implies that $x = x'$. This is an easy consequence of our final theorem, Theorem A.62 below, which says that, near x, distances from x in X correspond under the exponential map $\exp_x$ to distances from 0_x in T_xX, and shortest paths from x correspond to rays through 0_x (and hence are geodesics through x). Moreover, since small open balls with centre 0_x correspond to open balls with centre x with respect to d, and since $\exp_x$ is, near 0_x, a homeomorphism with respect to the original topology on X, we deduce that the metric topology of d is the original topology.

Theorem A.62 has other important corollaries. One, which we used in the proof of the generalized stable manifold theorem (Theorem 6.21), is that if Λ is a compact subset of X then we may choose a number $a > 0$ such that, for all $x \in \Lambda$, $\exp_x$ maps the ball in T_xX with centre 0_x and radius a with respect to $|\quad|_x$ onto the ball in X with centre x and radius a with respect to d.

(A.62) Theorem. *Let $x \in X$, and let B be any open ball with centre 0_x in T_xX small enough for $\exp_x$ to map B diffeomorphically onto its image. Then, for all $y \in B$, $d(x, \exp y) = |y|_x$. Moreover $|y|_x$ is the length of the curve $\gamma: [0, 1] \to X$ defined by $\gamma(t) = \exp ty$.*

Proof. We first comment that any C^3 diffeomorphism $h: X \to Y$ from X to a smooth manifold Y induces a Riemannian structure on Y, defined for all $p, q \in T_yY$ by $\langle p, q\rangle_y = \langle Th^{-1}(p), Th^{-1}(q)\rangle_{h^{-1}(y)}$. Correspondingly, the Riemannian spray $\sigma: TX \to T^2X$ induces a C^1 Riemannian spray $\tau: TY \to T^2Y$, and these are related by the commutative diagram

$$\begin{tikzcd} T^2X \arrow[r, "T^2h"] & T^2Y \\ TX \arrow[u, "\sigma"] \arrow[r, "Th"] & TY \arrow[u, "\tau"'] \end{tikzcd}$$

Each solution curve ε of τ at $y \in Y$ is of the form $h\delta$ where δ is a solution curve of σ at $h^{-1}(y)$. The length of a curve γ in X is equal to the length of the curve $h\gamma$ in Y, and thus the distance from x to x' in X equals the distance from $h(x)$ to $h(x')$ in Y.

We apply this idea with $X = \exp_x B$, $Y = B$ and h the inverse of the restriction $\exp_x: B \to \exp_x B$. To prove the theorem, we have to show that, with respect to the distance function of the new Riemannian metric induced on B by h, the distance from 0_x to y is $|y|_x$, for all $y \in B$, and that this is the length of the straight line segment $\gamma:[0, 1] \to B$ joining 0_x to y (i.e. $\gamma(t) = ty$). We denote the new Riemannian metric on B by $\langle\!\langle\ \rangle\!\rangle_y$ and the associated Finsler by $\|\ \|_y$. Since T_xX is a Banach space, we identify TB with $B \times T_xX$ in the usual way, and write $\langle\!\langle p, q\rangle\!\rangle_y$ or $K(y, p, q)$ for $\langle\!\langle (y, p), (y, q)\rangle\!\rangle_y$ and $\|p\|_y$ for $\|(y, p)\|_y$. Thus K is a real function on $B \times T_xX \times T_xX$. Perhaps we ought to emphasize that B already has a Riemannian metric induced from the inner product $\langle\ \rangle_x$ on T_xX. We denote this by $\langle\ \rangle_x$ as well; thus $\langle (y, p), (y, q)\rangle_x = \langle p, q\rangle_x$. The associated Finsler is denoted by $|\ |_x$.

Now T_xX may be identified with $\mathbf{E}$ (for example, by the isomorphism $T_x\xi$ for any chart ξ at x) and so we may think of h as being a chart. (In the literature h is usually called a *normal* chart at x.) Thus the Riemannian spray $\tau: B \times \mathbf{E} \to (\mathbf{B} \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E})$ is defined by the formulae for the local representative of the Riemannian spray σ on X (see (A.60)). We know that solution curves of τ come under h from solution curves of σ, and thus are curves of the form $\delta_p(t) = tp$. Differentiating each curve δ_p gives an integral curve of τ of the form $\gamma_p(t) = (tp, p)$. Differentiating again, we deduce that, for all $p \in B$ and for all $t \in \mathbf{R}$ such that $tp \in B$, $\tau(tp, p) = (tp, p, p, 0)$. Thus, by (A.60).

(A.63) $$(\tfrac{1}{2}TK(tp, p, p, u, 0, 0) - TK(tp, p, u, p, 0, 0))_2 = 0$$

for all $u \in \mathbf{E}$, and in particular $(TK(tp, p, p, p, 0, 0))_2 = 0$. This says that $\|p\|_z$ is constant for z on the line joining 0_x and p, and hence that $\|p\|_{tp} = \|p\|_{0_x} = |p|_x$. It is now clear that the line segment γ joining 0_x to y has length $|y|_x$ with respect to $\|\ \|$. Note that we have, in fact, shown that the curve $t \mapsto \exp ty$ in X has constant speed $|y|_x$ at every point.

The above property of $\|p\|_x$ holds, more generally, for $\langle\!\langle p, q\rangle\!\rangle_x$, for all $q \in \mathbf{E}$. We assume this for the time being:

(A.64) Lemma. *For all $p, q \in \mathbf{E}$ and for all $t \in \mathbf{R}$ with $tp \in B$,*

$$K(tp, p, q) = \langle p, q\rangle_x,$$

and complete the proof of the theorem. Let $\delta:[a, b] \to B$ be any piecewise C^1 curve joining 0_x to y. We wish to show that δ has length $\geqslant |y|_x$. We may assume that $\delta(t) \neq 0_x$ for $t > a$, for otherwise we may shorten δ. Then, for all but finitely many values of t, the Schwarz inequality gives

$$\|\delta'(t)\|_{\delta(t)} \geqslant |\langle\!\langle \delta(t), \delta'(t)\rangle\!\rangle_{\delta(t)}| / \|\delta(t)\|_{\delta(t)}$$

and by Lemma A.64 the right-hand side equals $|\langle \delta(t), \delta'(t)\rangle_x| / |\delta(t)|_x$. Now consider the curve $\varepsilon:[a, b] \to B$ defined by $\varepsilon(t) = |\delta(t)|_x y/|y|_x$. This parametrizes the line segment $[0_x, y]$, possibly covering it more than once in places, and so its length is at least $|y|_x$. But the formula for its length is $\int_a^b \|\varepsilon'(t)\|_{\varepsilon(t)}\,dt$, which reduces to $\int_a^b (|\langle\delta(t), \delta'(t)\rangle_x|/|\delta(t)|_x)\,dt$. Thus the length of δ is $\geqslant$ the length of ε. □

Proof of Lemma A.64. We first show that, for all $u \in \mathbf{E}$ with $\langle p, u\rangle_x = 0$, $K(tp, p, u) = 0$. We may assume that p and u have unit length with respect to $|\ |_x$, so that, for all $s \in \mathbf{R}$, $tp + su$ has length $\sqrt{t^2+s^2}$. For all $t > 0$, we differentiate the relation

$$K(tp+su, (tp+su)/\sqrt{t^2+s^2}, (tp+su)/\sqrt{t^2+s^2}) = 1$$

with respect to s at $s = 0$, and obtain

$$(TK(tp, p, p, u, u/t, u/t))_2 = 0$$

or, equivalently,

$$(TK(tp, p, p, u, 0, 0))_2 + (2/t)K(tp, p, u) = 0,$$

using the symmetry and bilinearity of $K(y, ., .)$. But, by (A.63),

$$(TK(tp, p, p, u, 0, 0))_2 = 2(TK(tp, p, u, p, 0, 0))_2.$$

Thus if the real function ψ is defined by $\psi(t) = K(tp, p, u)$ then it satisfies the differential equation $t\psi'(t) = -\psi(t)$ for $t > 0$. Hence $t\psi(t) =$ constant. Since ψ is continuous and $\psi(0) = 0$, the constant is zero, and hence $K(tp, p, u) = 0$, as required. Finally, for all $q \in \mathbf{E}$,

$$\langle p, q - \langle p, q\rangle_x p/\langle p, p\rangle_x\rangle_x = 0$$

and hence

$$\begin{aligned} 0 &= K(tp, p, q - \langle p, q\rangle_x p/\langle p, p\rangle_x) \\ &= K(tp, p, q) - (\langle p, q\rangle_x/\langle p, p\rangle_x)K(tp, p, p) \\ &= K(tp, p, q) - \langle p, q\rangle_x. \end{aligned}$$

□