Chapter 1

Infinite-dimensional Manifolds

1.1. A global setting for nonlinear DEs

Linear differential equations are often fruitfully studied via techniques from linear functional analysis, including the theory of Banach and Hilbert spaces. In contrast, the proper setting for an important class of nonlinear differential equations is a nonlinear version of functional analysis, based upon infinite-dimensional manifolds modeled on Banach or Hilbert spaces. The theory of such manifolds was developed by Eells, Palais and Smale among others in the 1950s and 1960s, and has proven to be extremely useful for understanding many of the nonlinear differential equations which arise in geometry, such as:

- (1) the equation for geodesics in a Riemannian manifold,
- (2) the equation for harmonic maps from surfaces into a Riemannian manifold, or for minimal surfaces in a Riemannian manifold,
- (3) the Yang-Mills equations for connections in principal bundles over a Riemannian manifold,
- (4) the equations for pseudoholomorphic or $J$-holomorphic curves in a symplectic manifold,
- (5) and the Seiberg-Witten equations.

In all of these examples, the solutions can be expressed as critical points of a real-valued function (often called the action or the energy) defined on an infinite-dimensional manifold, such as the function space manifolds from one finite-dimensional manifold to another, as described in the following
pages. In favorable cases, a gradient (or “pseudogradient”) of the action or energy can then be used to locate critical points (solutions to the differential equations) via what is often called the method of steepest descent.

For this procedure to work, the topology on the function space must be strong enough for the action or energy to be differentiable, yet weak enough to force convergence of a subsequence of a sequence which is tending toward a minimum (or to a minimax solution for a given constraint). The two conflicting conditions often select a unique acceptable topology for the space of functions. In the most favorable circumstances, the topology is strong enough so that it lies within the space of continuous functions, a space which has been studied extensively by topologists. The function space is then often homotopy equivalent, that is equivalent in the sense of homotopy theory, to the space of continuous functions. In favorable cases, this makes existence questions within the theory of nonlinear differential equations accessible by topological methods. An example of this is the Morse theory of geodesics, which became a foundation for interaction between geometry and topology during the twentieth century.

In the case of the Yang-Mills equations or the Seiberg-Witten equations over a four-dimensional base manifold, it is useful to reverse the logic. Instead of topology shedding light on existence of solutions to partial differential equations, it is the space of solutions to these equations that enables one to distinguish between different smooth structures on a given topological four-manifold. This illustrates that at a very fundamental level, topology and nonlinear partial differential equations are closely related, and underscores the importance of developing a general theory of partial differential equations founded upon infinite-dimensional manifolds.

1.2. Infinite-dimensional calculus

Our first topic is the theory of infinite-dimensional manifolds. We refer to the excellent presentations of Lang [Lan95], Smale [Abr63] or Abraham, Marsden and Ratiu [AMR88] for further elaboration of topics only briefly introduced in the following pages.

It was pointed out by Eells, Smale, Palais, Abraham, Lang and others in the 1960s that several variable calculus could be developed not just in finite-dimensional Euclidean spaces, but also with very little extra work within the context of infinite-dimensional Hilbert and Banach spaces, and that this should have important implications for the solution of nonlinear differential equations, particularly those equations coming from the calculus of variations. Gradually, infinite-dimensional calculus became part of the standard tool chest for global analysis. Most of the proofs in the Banach
space setting are straightforward modifications of the proofs in $\mathbb{R}^n$, so we will go very rapidly over these proofs.

**Definition.** A *pre-Hilbert space* is a real vector space $E$ together with a function $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ which satisfies the following inner product axioms:

1. $\langle x, y \rangle = \langle y, x \rangle$, for $x, y \in E$.
2. $\langle ax, y \rangle = a \langle x, y \rangle$, for $a \in \mathbb{R}$ and $x, y \in E$.
3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$, for $x, y, z \in E$.
4. $\langle x, x \rangle \geq 0$, for $x \in E$, equality holding only when $x = 0$.

These axioms simply state that $\langle x, y \rangle$ is a positive-definite symmetric bilinear form on $E$.

Given such a pre-Hilbert space, we can define a map $\| \cdot \| : E \to \mathbb{R}$ by $\| x \| = \sqrt{\langle x, x \rangle}$. We can regard $E$ as a metric space by defining the distance between elements $x$ and $y$ of $E$ to be $d(x, y) = \| x - y \|$.

**Definition.** A *Hilbert space* is a pre-Hilbert space which is complete in terms of the metric $d$.

An example of a finite-dimensional Hilbert spaces is Euclidean space $\mathbb{R}^n$, with the inner product $\langle \cdot, \cdot \rangle$ being defined by $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1y_1 + \cdots + x_ny_n$.

An example of an infinite-dimensional Hilbert space is the space $L^2([0, 1], \mathbb{R})$ of equivalence classes of $L^2$ functions studied in basic analysis courses. To construct it, one starts with defining an inner product $\langle \cdot, \cdot \rangle$ on the space $C^\infty([0, 1], \mathbb{R}) = \{ \text{C}^\infty \text{ functions } f : [0, 1] \to \mathbb{R} \}$ by $\langle \phi, \psi \rangle = \int_0^1 \phi(t) \psi(t) dt$.

We let $L^2([0, 1], \mathbb{R})$ denote the equivalence classes of Cauchy sequences from $C^\infty([0, 1], \mathbb{R})$, two Cauchy sequences $\{ \phi_i \}$ and $\{ \psi_i \}$ being equivalent if for each $\epsilon > 0$ there is a positive integer $N$ such that

$$i, j > N \implies \| \phi_i - \psi_j \| < \epsilon.$$  

(1.1)

Equivalence classes of Cauchy sequences form a vector space, and we define an inner product $\langle \cdot, \cdot \rangle$ on $L^2([0, 1], \mathbb{R})$ by $\langle \{ \phi_i \}, \{ \psi_i \} \rangle = \lim_{i \to \infty} \langle \phi_i, \psi_i \rangle$.

Finally, we check that with this inner product, $L^2([0, 1], \mathbb{R})$ is a Hilbert space.
We say that \( L^2([0, 1], \mathbb{R}) \) is the completion of \( C^\infty([0, 1], \mathbb{R}) \) with respect to this inner product \( \langle \cdot, \cdot \rangle \). This process of completion is virtually the same as that often used to construct the real numbers from the rationals.

**Definition.** A pre-Banach space is a vector space \( E \) together with a function \( \| \cdot \| : E \to \mathbb{R} \) which satisfies the following axioms:

1. \( \| ax \| = |a| \| x \| \), when \( a \in \mathbb{R} \) and \( x \in E \).
2. \( \| x + y \| \leq \| x \| + \| y \| \), when \( x, y \in E \).
3. \( \| x \| \geq 0 \), for \( x \in E \).
4. \( \| x \| = 0 \) only if \( x = 0 \).

A function \( \| \cdot \| : E \to \mathbb{R} \) which satisfies the first three axioms is called a seminorm on \( E \). If, in addition, it satisfies the fourth axiom it is called a norm.

As in the case of Hilbert spaces, we can make \( E \) into a metric space by defining the distance between elements \( x \) and \( y \) of \( E \) to be \( d(x, y) = \| x - y \| \).

**Definition.** A Banach space is a pre-Banach space which is complete in terms of the metric \( d \).

Of course, every Hilbert space is a Banach space with norm \( \| \cdot \| \) defined by \( \| x \| = \sqrt{\langle x, x \rangle} \). Let

\[
C^0([0, 1], \mathbb{R}) = \{ \text{continuous functions } f : [0, 1] \to \mathbb{R} \},
\]

and define

\[
(1.2) \quad \| \cdot \| : C^0([0, 1], \mathbb{R}) \to \mathbb{R} \quad \text{by} \quad \| f \| = \sup \{ |f(t)| : t \in [0, 1] \}.
\]

Then \( \| \cdot \| \) makes \( C^0([0, 1], \mathbb{R}) \) into a Banach space with a norm that does not come from an inner product. More generally, we can consider the space

\[
C^k([0, 1], \mathbb{R}) = \{ \text{functions } f : [0, 1] \to \mathbb{R} \text{ which have continuous derivatives up to order } k \},
\]

a Banach space with respect to the norm

\[
(1.3) \quad \| \cdot \|_k : C^k([0, 1], \mathbb{R}) \to \mathbb{R} \quad \text{defined by} \quad \| f \|_k = \sup \left\{ \sum_{i=0}^{k} |f^{(i)}(t)| : t \in [0, 1] \right\},
\]

where \( f^{(i)}(t) \) denotes the derivative of \( f \) of order \( i \).

When \( 1 \leq p < \infty \) and \( p \neq 2 \), the spaces \( L^p([0, 1], \mathbb{R}) \) studied in basic analysis courses are Banach spaces which are not Hilbert spaces. To construct these spaces, one starts with defining a function \( \| \cdot \| \) on the space
$C^\infty([0, 1], \mathbb{R})$ of $C^\infty$ functions $f : [0, 1] \to \mathbb{R}$ by
\[
\|\phi\| = \left[ \int_0^1 |\phi(t)|^p dt \right]^{1/p},
\]
which is shown to be a norm by means of the Minkowski inequality. This agrees with the norm previously defined on $L^2([0, 1], \mathbb{R})$ when $p = 2$. Just as in the construction of $L^2([0, 1], \mathbb{R})$, one can use this norm to define Cauchy sequences, and let $L^p([0, 1], \mathbb{R})$ be the set of equivalence classes of Cauchy sequences, where two Cauchy sequences $\{\phi_i\}$ and $\{\psi_i\}$ are equivalent if for each $\varepsilon > 0$ there is a positive integer $N$ such that (1.1) holds. The basic properties of $L^p$ spaces are treated in standard references on functional analysis; thus for the Hölder and Minkowski inequalities, for example, one can refer to Theorem III.1 of [RS80].

Each Banach space $E$ has a metric
\[
d : E \times E \to \mathbb{R} \quad \text{defined by} \quad d(e_1, e_2) = \|e_1 - e_2\|,
\]
which makes $E$ into a metric space, and we say that a subset $U$ of $E$ is open if
\[
p \in U \implies B_\varepsilon(p) = \{q \in E : d(p, q) < \varepsilon\} \subset U,
\]
for some $\varepsilon > 0$. A map $T : E_1 \to E_2$ between Banach spaces is continuous if $T^{-1}(U)$ is open for each open $U \subset E_2$, or equivalently, if there is a constant $c > 0$ such that
\[
\|T(e_1)\|_2 \leq c\|e_1\|_1, \quad \text{for all } e_1 \in E_1,
\]
where $\| \cdot \|_1$ and $\| \cdot \|_2$ are the norms on $E_1$ and $E_2$, respectively.

We let $L(E_1, E_2)$ be the space of continuous linear maps $T : E_1 \to E_2$, a Banach space in its own right under the norm
\[
\|T\| = \sup\{\|T(e_1)\| : e_1 \in E_1, \|e_1\| = 1\}.
\]
It is easily shown that $L(E_1, E_2)$ is the same as the space of linear maps from $E_1$ to $E_2$, which are bounded in terms of this norm. In particular, we can define the dual of a Banach space $E$ to be $E^* = L(E, \mathbb{R})$. We say that a Banach space is reflexive if $(E^*)^*$ is isomorphic to $E$. It is proven in analysis courses that $L^p([0, 1], \mathbb{R})$ is reflexive when $1 < p < \infty$ but $L^1([0, 1], \mathbb{R})$ and $C^0([0, 1], \mathbb{R})$ are not.

Banach spaces and continuous linear maps form a category, as do Hilbert spaces and continuous linear maps. The subject known as functional analysis is concerned with the categories of Hilbert spaces, Banach spaces and more general spaces of functions, and is one of the major tools in studying linear partial differential equations. Key theorems from the theory of Banach spaces include the Open Mapping Theorem, the Hahn-Banach Theorem and the Uniform Boundedness Theorem. These theorems are discussed in
The Open Mapping Theorem states that if $T : E_1 \to E_2$ is a continuous surjective map between Banach spaces, it takes open sets to open sets. Thus if $T$ is a continuous bijection, its inverse is continuous. The Hahn-Banach Theorem implies that if $e$ is a nonzero element of a Banach space $E$, then there is a continuous linear function $\lambda : E \to \mathbb{R}$ such that $\lambda(e) \neq 0$. A bilinear map $B : E_1 \times E_2 \to F$ is said to be continuous if there is a constant $c > 0$ such that
\[
\|B(e_1, e_2)\| \leq c\|e_1\|_1\|e_2\|_2, \quad \text{for all } e_1 \in E_1 \text{ and all } e_2 \in E_2.
\]
One of the useful consequences of the Uniform Boundedness Theorem is that such a Bilinear map is continuous if and only if
\[
B(\cdot, e_2) : E_1 \to F \quad \text{and} \quad B(e_1, \cdot) : E_2 \to F
\]
are continuous for each $e_1 \in E_1$ and each $e_2 \in E_2$. The three theorems are almost trivial to prove for finite-dimensional Banach spaces, but the proofs are more subtle for infinite-dimensional Banach spaces and rely on the axiom of choice.

**Definition.** A Banach algebra is a Banach space $E$ together with an associative product $\mu : E \times E \to E$ such that if $x \cdot y = \mu(x, y)$, then
\[
\|x \cdot y\| \leq \|x\|\|y\|, \quad \text{for } x, y \in E.
\]

For example, the Banach space $C^0([0, 1], \mathbb{R})$ with norm given by (1.2) is a commutative Banach algebra with unit when $\mu$ is the ordinary multiplication of functions.

We now turn to the question of how to develop differential calculus for functions defined on Banach spaces. It is actually the topology, or the equivalence class of norms on the Banach space, that is important for calculus, two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a linear space $E$ being equivalent if there is a constant $c > 1$ such that
\[
\frac{1}{c}\|x\|_1 < \|x\|_2 < c\|x\|_1, \quad \text{for } x \in E.
\]

Lang [Lan95] calls a vector space $E$ a Banachable space if it is endowed with an equivalence class of Banach space norms. However, most other authors do not use this term, and we will usually use the simpler term Banach space, it being understood, however, that we may sometimes pass to an equivalent norm in the middle of an argument, when it is the underlying vector space with its topology, the “topological vector space”—not the norm itself—that is important.
**Definition.** Suppose that $E_1$ and $E_2$ are Banach spaces, and that $U$ is an open subset of $E_1$. A continuous map $f : U \to E_2$ is said to be **differentiable** at the point $x_0 \in U$ if there exists a continuous linear map $T : E_1 \to E_2$ such that

$$
\lim_{\|h\| \to 0} \frac{\|f(x_0 + h) - f(x_0) - T(h)\|}{\|h\|} = 0,
$$

where $\| \cdot \|$ denotes both the Banach space norms on $E_1$ and $E_2$. We will call $T$ the (Fréchet) derivative of $f$ at $x_0$ and write $Df(x_0)$.

Just as in ordinary calculus, the derivative $Df(x_0)$ determines the linearization of $f$ near $x_0$, which is the affine function

$$
\tilde{f}(x) = f(x_0) + Df(x_0)(x - x_0)
$$

which most closely approximates $f$ near $x_0$.

If $f$ is differentiable at every $x \in U$, a derivative $Df(x)$ is defined for each $x \in U$ and thus we have a set-theoretic map $Df : U \to L(E_1, E_2)$. If this map $Df$ is continuous, we can also ask whether it is differentiable at $x_0 \in U$. This will be the case if there is a continuous linear map $T : E_1 \to L(E_1, E_2)$ such that

$$
\lim_{\|h\| \to 0} \frac{\|Df(x_0 + h) - Df(x_0) - T(h)\|}{\|h\|} = 0,
$$

in which case we write $D^2f(x_0)$ for $T$ and call $D^2f(x_0)$ the second derivative of $f$ at $x_0$. Note that

$$
D^2f(x_0) \in L(E_1, L(E_1, E_2)) = L^2(E_1, E_2)
$$

$$
= \{ \text{continuous bilinear maps } T : E_1 \times E_1 \to E_2 \},
$$

which can also be made into a Banach space in an obvious way.

We say that a function $f : U \to E_2$, where $U$ is an open subset of $E_1$, is

(1) $C^0$ if it is continuous.
(2) $C^1$ if it is continuous and differentiable at every $x \in U$, and $Df : U \to L(E_1, E_2)$ is continuous.
(3) $C^k$ for $k \geq 2$ if it is $C^1$ and $Df : U \to L(E_1, E_2)$ is $C^{k-1}$.
(4) $C^\infty$ or smooth if it is $C^k$ for every nonnegative integer $k$.

With the above definition of differentiation, many of the arguments in several variable calculus can be transported without difficulty to the Banach space setting, as carried out in detail in [Lan95] or [AMR88]. For example, the Leibniz rule for differentiating a product carries over immediately to the infinite-dimensional setting.
Proposition 1.2.1. Suppose that $B : F_1 \times F_2 \to G$ is a continuous bilinear map between Banach spaces, that $U$ is an open subsets of a Banach space $E$ and
\[ f : U_1 \to F_1, \quad f_2 : U_2 \to F_2 \]
are $C^1$ maps. Then $e \mapsto g(e) = B(f_1(e), f_2(e))$ is a $C^1$ map, and
\[ Dg(x_0)h = B(Df_1(x_0)h, f_2(e)) + B(f_1(x_0)Df_2(x_0)h). \]

Sketch of proof. To simplify notation, we write
\[ g(x) = B(f_1(x), f_2(x)) = f_1(x) \cdot f_2(x). \]
Then
\[ g(x + h) - g(x) = (f_1(x + h) - f_1(x))f_2(x + h) + f_1(x)(f_2(x + h) - f_2(x)), \]
and hence
\[ \frac{g(x + h) - g(x)}{\|h\|} = \frac{f_1(x + h) - f_1(x)}{\|h\|} f_2(x + h) + f_1(x) \frac{f_2(x + h) - f_2(x)}{\|h\|}. \]
The Proposition follows by taking the limit as $\|h\| \to 0$. QED

The following lemma is called the chain rule.

Proposition 1.2.2. If $U_1$ and $U_2$ are open subsets of Banach spaces $E_1$ and $E_2$ and
\[ f : U_1 \to U_2, \quad g : U_2 \to E_3 \]
are $C^1$ maps, then so is $g \circ f : U_1 \to E_3$, and
\[ D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0), \quad \text{for } x_0 \in U_1. \]

Sketch of proof. We have
\[ f(x_0 + h) = f(x_0) + Df(x_0)h + o(h), \]
where the symbol $o(h)$ stands for an element in $E_2$ such that $o(h)/\|h\| \to 0$ as $\|h\| \to 0$. Similarly,
\[ g(f(x_0) + k) = g(f(x_0)) + Dg(f(x_0))(k) + o(k). \]
Setting $k = Df(x_0)h + o(h)$ yields
\[ g(f(x_0 + h)) = g(f(x_0)) + Dg(f(x_0))(Df(x_0)h + o(h)) + o(k). \]
One checks without difficulty using continuity of $g$ that an $o(k)$ expression, where $k$ is a bounded linear function of $h$, is also $o(h)$, and hence
\[ g(f(x_0 + h)) = g(f(x_0)) + Dg(f(x_0))Df(x_0)h + o(h), \]
which gives the desired conclusion. QED
By induction, one immediately shows that the composition of \( C^k \) maps is \( C^k \) and the composition of \( C^\infty \) maps is \( C^\infty \).

Another familiar theorem from several variable calculus in finite dimension is the “equality of mixed partials”. To state the infinite-dimensional version, we let \( E \) and \( F \) be Banach spaces, \( U \) an open subset of \( E \). Suppose that \( f : U \to F \) is a \( C^2 \) map, so that for \( x_0 \in U \),
\[
D^2 f(x_0) \in L(E, L(E, F)) = L^2(E, F).
\]
Of course, a very important case is the one where \( F = \mathbb{R} \), the base field of real numbers.

**Proposition 1.2.3.** \( D^2 f(x_0) \) is symmetric; in other words,
\[
D^2 f(x_0)(h, k) = D^2 f(x_0)(k, h), \quad \text{for } h, k \in E.
\]

**Sketch of proof.** First note that by the Hahn-Banach theorem, it suffices to show that if \( \lambda : F \to \mathbb{R} \) is any continuous linear functional, then \( D^2(\lambda \circ f)(x_0) = \lambda \circ D^2 f(x_0) \) is symmetric, because if \( D^2 f(x_0) \) is not symmetric, the same will be true of \( \lambda \circ D^2 f(x_0) \), for some linear function \( \lambda \). This reduces the proof to the case where \( F = \mathbb{R} \). Next note that via the chain rule,
\[
f(x + h) - f(x) = \int_0^1 (Df)(x + th) h dt,
\]
and by iteration,
\[
f(x + h + k) - f(x + k) - f(x + h) + f(x) = \int_0^1 (Df)(x + th + k) h dt
\]
\[
= \int_0^1 \int_0^1 (D(Df)(x + th + sk)(k)) h ds dt.
\]
Interchanging \( h \) and \( k \) yields
\[
f(x + h + k) - f(x + h) - f(x + k) + f(x) = \int_0^1 (Df)(x + tk + h) k dt
\]
\[
= \int_0^1 \int_0^1 (D(Df)(x + sk + th)(h)) k ds dt.
\]
Since the left-hand sides of the last two expressions are equal, so are the right-hand sides, and hence
\[
\int_0^1 \int_0^1 [(D(Df)(x + th + sk)(k)) h - (D(Df)(x + th + sk)(h)) k] ds dt = 0.
\]
Since \( D(Df) \) is a continuous function, this can only happen if \( D^2 f(x)(k, h) = D^2 f(x)(h, k) \), for all \( x, h \) and \( k \), which is exactly what we needed to prove. QED
More generally, if \( f : U \to F \) is a \( C^k \)-map, then
\[
D^k f(x_0) = D(D^{k-1} f)(x_0) \in L(E, L^{k-1}(E, F)) = L^k(E, F),
\]
and by an induction based on the previous lemma, we see that in fact
\[
D^k f(x_0) \in L^k_s(E, F) = \{ T \in L^k(E, F) : T \text{ is symmetric} \}.
\]
By symmetric we mean that
\[
T(h_{\sigma(1)}, \ldots, h_{\sigma(k)}) = T(h_1, \ldots, h_k),
\]
for all permutations \( \sigma \) in the symmetric group \( S_k \) on \( k \) letters.

It is often useful to have an explicit formula for the higher derivatives of a composition. The following such formula comes from §3 of [Abr63]. If \( p \) and \( k \) are positive integers with \( k \leq p \) and \( (i_1, \ldots, i_k) \) is a \( k \)-tuple of positive integers such that \( i_1 + \cdots + i_k = p \) and \( i_1 \leq i_2 \leq \cdots \leq i_k \), we define integers \( \sigma^p_k(i_1, \ldots, i_k) \) recursively by requiring that \( \sigma^1_1(1) = 1 \) and
\[
\sigma^p_k(i_1, \ldots, i_k) = \delta_{i_1}^1 \sigma^{p-1}_{k-1}(i_2, \ldots, i_k) + \sum_{j=1}^k \sigma^{p-1}_k(i_1, \ldots, i_j + 1, \ldots, i_k),
\]
where \( \delta_{i_1}^1 \) is 1 if \( i_1 = 1 \) and 0 otherwise.

**Proposition 1.2.4.** Suppose that \( U \) and \( V \) are open subsets of Banach spaces \( E \) and \( F \), respectively, and that \( f : U \to V \) and \( g : V \to G \) are \( C^p \) maps, where \( G \) is a third Banach space. Then
\[
D^p(g \circ f) = \sum_{k=1}^p (D^k g \circ f) P^p_k(f),
\]
where
\[
P^p_k(f) = \sum \sigma^p_k(i_1, \ldots, i_k)(D^{i_1} f, \ldots, D^{i_k} f),
\]
the sum being taken over all \( p \)-tuples of positive integers such that \( i_1 + \cdots + i_k = p \) and \( i_1 \leq i_2 \leq \cdots \leq i_k \).

The proof of Proposition 1.2.5 is by induction on \( p \) starting with the case \( p = 1 \), which is an immediate application of the chain rule. For \( p \geq 2 \), one obtains the formula for \( \sigma^p_k \) and the expression for \( D^p(g \circ f) \) by applying the chain rule and the Leibniz rule for differentiating a product.

**Definition of integration along a path.** In order to put Taylor’s theorem in the Banach space setting, we need to define the integral of a continuous map \( \gamma : [0, 1] \to E \) into a Banach space \( E \). The definition is particularly easy if \( E \) is a reflexive Banach space; in this case, we just set
\[
(1.4) \quad \int_0^1 \gamma(t) dt = e, \quad \text{where} \quad \lambda(e) = \int_0^1 \lambda \circ \gamma(t) dt, \quad \text{for} \quad \lambda \in E^*.
\]
A definition of the integral for general Banach spaces can be found in §1.4 of [Lan95].

Let $U$ be an open subset of a Banach space $E$. Following [AMR88], we define a thickening of $U$ to be an open subset $\tilde{U} \subset E \times E$ such that

1. $U \times \{0\} \subset \tilde{U}$, and
2. $(x, h) \in \tilde{U} \Rightarrow x + th \in U$, for $t \in [0, 1]$.

**Proposition 1.2.5 (Taylor’s Theorem).** If a map $f : U \to F$ is of class $C^r$, there exist continuous maps

$$
\phi_k : U \to L^k_s(E, F), \quad \text{for } 1 \leq k \leq r \quad \text{and} \quad R : \tilde{U} \to L^r_s(E, F),
$$

where $\tilde{U}$ is a thickening of $U$, such that $R(x, 0) = 0$ and

$$
f(x + h) = f(x) + \phi_1(x)h + \frac{1}{2!}\phi_2(x)(h, h) + \cdots + \frac{1}{r!}\phi_r(x)(h, \ldots, h)
+ R(x, h)(h, \ldots, h).
$$

Here $\phi_k(x) = (D^k f)(x)$, for $1 \leq k \leq r$.

**Sketch of proof.** We first reduce to the case where $F = \mathbb{R}$ by applying the Hahn-Banach Theorem. One then uses the fundamental theorem of calculus to establish the first order Taylor formula

$$
(1.5) \quad f(x + h) = f(x) + \int_0^1 (Df)(x + th)h dt
= f(x) + (Df)(x)h + \int_0^1 [(Df)(x + th) - (Df)(x)]h dt.
$$

Using the fundamental theorem of calculus again, one next obtains

$$
f(x + h) = f(x) + (Df)(x)h + \int_0^1 \int_0^1 [(D^2f)(x + s th)](h, h)tdsdt
= f(x) + (Df)(x)h + \frac{1}{2!}(D^2f)(x)(h, h)
+ \int_0^1 \int_0^1 [(D^2f)(x + s th) - D^2f(x)](h, h)tdsdt.
$$

Continuing in the same manner, we find that

$$
f(x + h) = f(x) + (Df)(x)h + \frac{1}{2!}(D^2f)(x)(h, h)
+ \cdots + \frac{1}{k!}(D^k f)(x)(h, \ldots, h) + R(x, h)(h, \ldots, h),
$$

where $R(x, h) \in L^k_s(E, F)$ depends continuously on $x$ and $h$ and $R(x, 0) = 0$.

QED
Example 1.2.6. We suppose that the domain of the function is the Banach space $E = L^p(S^1, \mathbb{R}^N)$, the completion of the space $C^\infty(S^1, \mathbb{R}^N)$ of smooth $\mathbb{R}^N$-valued functions on $S^1$ with respect to the $L^p$-norm

$$\|\phi\|_{L^p} = \left[\int_{S^1} |\phi(t)|^p dt\right]^{1/p}, \quad \text{for } \phi \in L^p(S^1, \mathbb{R}^N).$$

Here $S^1$ is regarded as the quotient of the interval $[0, 1]$ obtained by identifying the points 0 and 1, and possesses the standard measure $dt$ with respect to which $S^1$ has measure one. A useful tool for dealing with functions in the $L^p$ spaces is the H"older inequality which states: if $\phi \in L^p$, $\psi \in L^q$ and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ where $p, q, r \geq 1$, then $\phi \psi \in L^r$ and $\|\phi \psi\|_{L^r} \leq \|\phi\|_{L^p} \|\psi\|_{L^q}$.

Using this inequality and the chain rule, it is not difficult to show that when $p \geq 2$, the function $f : E \rightarrow \mathbb{R}$ defined by $f(\phi) = \int_{S^1} (1 + |\phi(t)|^2)^{p/2} dt$

is $C^2$ and that its first and second derivatives are given by the formulae

$$D f(\phi)(\psi) = p \int_{S^1} (1 + |\phi(t)|^2)^{p/2 - 1} \phi(t) \cdot \psi(t) dt$$

and

$$(D^2 f)(\phi)(\psi_1, \psi_2) = p \int_{S^1} (1 + |\phi(t)|^2)^{p/2 - 1} \psi_1(t) \cdot \psi_2(t) dt + p(p - 2) \int_{S^1} (1 + |\phi(t)|^2)^{p/2 - 2} (\phi(t) \cdot \psi_1(t))(\phi(t) \cdot \psi_2(t)) dt.$$ 

To verify these formulae, we apply the chain rule to $f = h \circ g$, where $h$ is integration over $S^1$ and $g : E \rightarrow L^1(S^1, \mathbb{R}^N)$ by $g(\phi)(t) = (1 + |\phi(t)|^2)^{p/2}$. To check that $g$ is $C^1$, we let $h : \mathbb{R}^N \rightarrow \mathbb{R}$ by $h(u) = (1 + u \cdot u)^{p/2}$ and apply Taylor’s formula (1.5) to obtain, for a fixed choice of $t$,

$$h(\phi(t) + \psi(t)) = h(\phi(t)) + (Dh)(\phi(t))\psi(t) + \int_0^1 [(Dh)(\phi(t) + s\psi(t)) - (Dh)(\phi(t))] \psi(t) ds,$$

where

$$Dh(\phi(t)) = p(1 + |\phi(t)|^2)^{p/2 - 1} \phi(t).$$
1.2. Infinite-dimensional calculus

It then follows from (1.6) that
\[
g(\phi + \psi)(t) = g(\phi)(t) + T(\phi)(t)\psi(t)
+ \int_0^1 [T(\phi + s\psi) - T(\phi)](t)\psi(t)ds,
\]
where \( T(\phi)(t) = Dh(\phi(t)) \). If \((1/p) + (1/q) = 1\), one can show that the Hölder inequality provides a continuous map
\[
T : L^p(S^1, \mathbb{R}^N) \to L^q(S^1, \mathbb{R}^N),
\]
and hence given any \( \varepsilon > 0 \), we can make
\[
\| [T(\phi + s\psi) - T(\phi)] \|_{L^q} < \varepsilon
\]
by choosing \( \| \psi \|_{L^p} \) is sufficiently small. Thus the integral term in (1.7) is \( o(\psi) \) and the derivative of \( g \) is
\[
Dg(\phi)(\psi) = T(\phi)\psi = (1 + |\phi|^2)^{p/2 - 1}\phi \cdot \psi.
\]
Moreover, continuity of \( T \) implies that \( g : E \to L^1 \) and \( f : E \to \mathbb{R} \) are \( C^1 \). A similar argument using the second order Taylor expansion shows that \( f \) is \( C^2 \).

On the other hand, if \( f \) were \( C^3 \) in the case \( N = 1 \), one could verify that the third derivative of \( g \) would be given by the formula
\[
(D^3g)(\psi_1, \psi_2, \psi_3) = 2p(p - 2)(1 + |\phi|^2)^{p/2 - 2}\phi\psi_1\psi_2\psi_3
+ p(p - 2)(p - 4)(1 + |\phi|^2)^{p/2 - 3}\phi^3\psi_1\psi_2\psi_3,
\]
and if \( p < 3 \), for a suitable smooth choice of \( \phi \), we could choose \( \psi_1, \psi_2 \) and \( \psi_3 \) in \( L^p \) such that the product \( \psi_1\psi_2\psi_3 \) is not in \( L^1 \). This implies that \( f \) cannot be \( C^3 \) when \( 2 < p < 3 \).

We have seen that many of the basic results of differential calculus of several variables extend with little change to the infinite-dimensional context. The following theorem is somewhat deeper:

**Theorem 1.2.7 (Inverse Function Theorem).** If \( U_1 \) and \( U_2 \) are open subsets of Banach spaces \( E_1 \) and \( E_2 \) with \( x_0 \in U_1 \), and \( f : U_1 \to U_2 \) is a \( C^\infty \) map such that \( Df(x_0) \in L(E_1, E_2) \) is invertible, then there are open neighborhoods \( V_1 \) of \( x_0 \) and \( V_2 \) of \( f(x_0) \), and a \( C^\infty \) map \( g : V_2 \to V_1 \), such that
\[
f \circ g = \text{id}_{V_2} \quad \text{and} \quad g \circ f = \text{id}_{V_1}.
\]
Moreover, \( Dg(f(x)) = [Df(x)]^{-1} \), for \( x \in V_1 \).

**Sketch of proof.** We can assume without loss of generality that \( x_0 = 0 \in U_1 \) and \( f(0) = 0 \in U_2 \). We can assume, moreover, that \( E_1 = E_2 \) and \( Df(0) \) is the identity map by replacing \( f \) by \( Df(0)^{-1} \circ f \). Define \( h : U_1 \to E_1 \) by
\[ h(x) = x - f(x). \] Then \( Dh(0) = 0 \), and by continuity of \( Dh \) there exists \( \delta > 0 \) such that
\[
\| x \| \leq \delta \quad \Rightarrow \quad x \in U_1 \quad \text{and} \quad \| dh(x) \| < \frac{1}{2}.
\]
If \( \| x \|, \| y \| < \delta \), then it follows from the chain rule that
\[
\| h(x) - h(y) \| = \left\| \int_0^1 \frac{d}{dt}(h(t x + (1-t)y))dt \right\|
= \left\| \int_0^1 Dh(t x + (1-t)y)(x-y)dt \right\|
\leq \left[ \int_0^1 \| Dh(t x + (1-t)y) \| dt \right] \| x - y \| < \frac{1}{2} \| x - y \|,
\]
so \( h \) decreases distances.

More generally, if \( y \in E_1 \) and \( \| y \| < \delta/2 \), we can define the map \( h_y : U_1 \rightarrow E_1 \) by \( h_y(x) = h(x) + y \). Then
\[
\| x \| \leq \delta \quad \Rightarrow \quad \| h_y(x) \| \leq \| y \| + \| h(x) \| < \frac{\delta}{2} + \frac{1}{2} \| x \| \leq \delta,
\]
so \( h_y \) takes the closed ball \( \overline{B_\delta(0)} \) to the open ball \( B_\delta(0) \), and
\[
\| h_y(x) - h_y(x') \| = \| h(x) - h(x') \| < \frac{1}{2} \| x - x' \|,
\]
so \( h_y \) is a contraction. Thus by the well-known Contraction Lemma, given \( y \) with \( \| y \| < \delta/2 \), there is a unique fixed point \( x \) of \( h_y \); that is, there is a unique \( x \) such that \( \| x \| \leq \delta \) and
\[
h_y(x) = x \quad \Rightarrow \quad x - f(x) + y = x \quad \Rightarrow \quad f(x) = y.
\]

Let
\[
V_2 = \{ y \in E_1 : \| y \| < \delta/2 \} \quad \text{and} \quad V_1 = \{ x \in E_1 : \| x \| < \delta, f(x) \in V_2 \}
\]
and define
\[
g : V_2 \rightarrow V_1 \quad \text{by} \quad g(y) = x \in V_1 \quad \Leftrightarrow \quad f(x) = y.
\]
Then \( g \) is a set-theoretic inverse to \( f : V_1 \rightarrow V_2 \).

To show that \( g \) is continuous, it suffices to show that \( |x - x'| \leq 2|f(x) - f(x')| \). But
\[
| x - x' | \leq |(x - f(x)) - (x' - f(x'))| + |f(x) - f(x')|
\leq |h(x) - h(x')| + |f(x) - f(x')| \leq \frac{1}{2} |x - x'| + |f(x) - f(x')|,
\]
which clearly implies the desired result.
To see that that $g$ is $C^1$, we note first that if $x_0 \in V_1$,
\[ f(x) - f(x_0) = Df(x_0)(x - x_0) + o(|x - x_0|). \]
If $y_0 = f(x_0)$ and $y = f(x)$, we can rewrite this equation as
\[ y - y_0 = Df(x_0)(g(y) - g(y_0)) + o(|x - x_0|), \]
or
\[ g(y) - g(y_0) = [Df(x_0)]^{-1}(y - y_0 + o(|x - x_0|)). \]
The continuity argument shows that $[Df(x_0)]^{-1}(o(|x - x_0|))$ is $o(|y - y_0|)$, so $g$ is $C^1$ with derivative $Dg(y) = (Df)^{-1}(g(y))$.

Finally, one uses “bootstrapping”:
\[ g \in C^1 \Rightarrow (Df)^{-1} \circ g \in C^1 \Rightarrow Dg \in C^1 \Rightarrow g \in C^2 \Rightarrow \cdots \]
to conclude that $g$ is $C^\infty$, and the theorem is proven. Later, we will see that the technique of bootstrapping used here is a fundamental tool for the study of nonlinear PDEs. QED

Before stating an important corollary of the Inverse Function Theorem, we point out one of the difficulties in dealing with Banach spaces. A closed linear subspace $E_1$ of a Banach space $E$ is a Banach space in its own right, but there may not exist a complementary closed subspace $E_2$ such that $E$ is linearly homeomorphic to $E_1 \oplus E_2$. We say that a subspace $E_1$ of a Banach space $E$ is split if there does exist such a complement $E_2$. For example, any finite-dimensional subspace of a Banach space is split as is any closed subspace of finite codimension. Moreover, any closed subspace of a Hilbert space is split, because the inner product can be used to define the orthogonal complement.

**Corollary 1.2.8.** If $U$ is an open subset of the Banach space $E$ with $x \in U$, and $f : U \to F$ is a $C^\infty$ map such that $Df(x) \in L(E, F)$ is surjective with split kernel, then there exists an open subset $V \subset U$ and a diffeomorphism $\phi : V_1 \times V_2 \to V$, where $V_1$ and $V_2$ are open subsets of Banach spaces $E_1$ and $E_2$ with $E = E_1 \oplus E_2$ and $E_2 \cong F$, such that $f \circ \phi$ is the projection on the second factor.

**Sketch of proof.** We can assume without loss of generality that $x = 0$ and $f(0) = 0$. Let $E_1$ be the kernel of $Df(0)$ and since it splits, let $E_2$ be a complement such that $E = E_1 \oplus E_2$. Note that $Df(0)$ establishes an isomorphism from $E_2$ to $F$. Let
\[ g : E = E_1 \oplus E_2 \to E = E_1 \oplus E_2 \quad \text{by} \quad g(x_1, x_2) = (x_1, f(x_1, x_2)). \]
One easily checks that $Dg(0)$ is invertible. Now apply the inverse function theorem to construct a smooth map 

$$\phi : V_1 \times V_2 \longrightarrow V \subset U$$

such that $g \circ \phi = \text{id}$, the identity map. The projection on $E_2$ is just $f \circ \phi$. QED

**Digression.** Notably absent from our examples of Banach spaces is the space $C^\infty([0,1],\mathbb{R}) = \{ \text{functions } f : [0,1] \to \mathbb{R} \text{ which have continuous derivatives of all orders} \}$, which one suspects should have importance for the theory of nonlinear partial differential equations. Unfortunately, the natural topology to use on this space is not defined by a single norm or seminorm, but by a countable collection of norms $\{ \| \cdot \|_k : k \in \mathbb{Z}, k \geq 0 \}$, where $\| \cdot \|_k$ is defined by (1.3).

**Definition.** A *Fréchet space* is a vector space $E$ together with a countable collection of seminorms $\{ \| \cdot \|_k : k \in \mathbb{Z}, k \geq 0 \}$ satisfying the following axioms:

1. $\| x \|_k = 0$ for all $k$ only if $x = 0$.
2. Suppose that $\{ x_i \}$ is a sequence of elements from $E$. If for each $\epsilon > 0$, there is a $N$ such that $i, j \geq N$ implies that $\| x_i - x_j \|_k < \epsilon$ for all $k$, then there is an element $x \in E$ such that $\| x_i - x \|_k$ converges to zero for all $k$.

Of course every Banach space is a Fréchet space, but $C^\infty([0,1],\mathbb{R})$ with the collection of norms defined above is a Fréchet space which is not a Banach space. Given a Fréchet space $E$ with seminorms $\{ \| \cdot \|_k : k \in \mathbb{Z}, k \geq 0 \}$, we can then define a bounded metric

$$d : E \times E \longrightarrow \mathbb{R} \quad \text{by} \quad d(x,y) = \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{\| x - y \|_k}{1 + \| x - y \|_k},$$

which one can check is complete. This metric defines a topology, so we can talk of continuous maps $f : E_1 \to E_2$ from the Fréchet space $E_1$ to the Fréchet space $E_2$, and we have a category consisting of Fréchet spaces and continuous linear maps.

**Definition.** Suppose that $E_1$ and $E_2$ are Fréchet spaces, and that $U$ is an open subset of $E_1$. A continuous map $f : U \to E_2$ is said to be *continuously differentiable* on $U$ if there exists a continuous map $Df : U \times E_1 \to E_2$ such that

$$Df(x)y = \lim_{t \to 0} \frac{f(x + ty) - f(x)}{t},$$
where $t$ ranges throughout $\mathbb{R} - \{0\}$, it being understood that the limit on the right-hand side exists for all $x \in U$ and all $y \in E_1$. It is proven in Hamilton’s survey article [Ham82], Part I, 3.2.3 and 3.2.5, that if $f : U \to E_2$ is continuously differentiable, the map $y \mapsto Df(x)y$ is linear. Thus, if $E_1$ and $E_2$ are Banach spaces, the above definition agrees with the definition previously given.

We could develop much of the infinite-dimensional calculus and the theory of infinite-dimensional manifolds in the category of Fréchet spaces, and in fact this is carried out in [Ham82], but the Inverse Function Theorem does not hold for Fréchet spaces. Moreover, in the existence theory for solutions to nonlinear partial differential equations, it is often convenient to first prove existence in a given Banach space and then prove regularity using bootstrapping, just as we did in the proof of the Inverse Function Theorem. This technique seems particularly well-adapted to Banach spaces. For these reasons, we prefer to think of $C^\infty([0,1], \mathbb{R})$ as the intersection of a “chain” of Banach spaces,

$$
\cdots \subseteq C^{k+1}([0,1], \mathbb{R}) \subseteq C^k([0,1], \mathbb{R}) \subseteq C^{k-1}([0,1], \mathbb{R}) \subseteq \cdots \subseteq C^0([0,1], \mathbb{R}).
$$

For proving theorems, we will usually work in a suitable Banach spaces so that we can use theorems like the Inverse Function Theorem. However, the statements of theorems are sometimes more elegant when phrased in terms of a Fréchet space, or a chain of Banach spaces.

1.3. Manifolds modeled on Banach spaces

Definition. Let $E$ be a separable Banach space. A connected smooth manifold modeled on $E$ is a connected second countable Hausdorff space $M$ together with a collection $A = \{(U_\alpha, \phi_\alpha) : \alpha \in A\}$, where each $U_\alpha$ is an open subset of $M$ and each

$$
\phi_\alpha : U_\alpha \to \phi_\alpha(U_\alpha) \subseteq E
$$

is a homeomorphism onto an open subset $\phi_\alpha(U_\alpha)$ of $E$ such that:

1. $\bigcup\{U_\alpha : \alpha \in A\} = M$.
2. $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is $C^\infty$, for all $\alpha, \beta \in A$.

A nonconnected Hausdorff space is called a smooth manifold or a Banach manifold if each component is a connected smooth manifold modeled on some Banach space. (We allow different components to be modeled on different Banach spaces.) The smooth manifold is called a Hilbert manifold if each component is modeled on a Hilbert space.
We say that \( A = \{(U_\alpha, \phi_\alpha) : \alpha \in A\} \) is the atlas defining the smooth structure on \( M \), and each \((U_\alpha, \phi_\alpha)\) is one of the charts in the atlas.

**Remark.** We could define Fréchet manifolds by simply replacing the phrase “Banach space” by “Fréchet space” in the above definition.

Let \( M \) and \( N \) be smooth manifolds modeled on Banach spaces \( E \) and \( F \), respectively. Suppose that \( M \) and \( N \) have atlases \( A = \{(U_\alpha, \phi_\alpha) : \alpha \in A\} \) and \( B = \{(V_\beta, \psi_\beta) : \beta \in B\} \). Then a continuous map \( F : M \to N \) is said to be smooth if \( \psi_\beta \circ F \circ \phi_\alpha^{-1} \) is \( C^\infty \), where defined, for \( \alpha \in A \) and \( \beta \in B \). It follows from the chain rule that the composition of smooth maps is smooth. In this way we obtain a category whose objects are smooth manifolds modeled on Banach spaces and whose morphisms are smooth maps between such manifolds.

As in the case of finite-dimensional manifolds, a diffeomorphism is a smooth map between manifolds with smooth inverse. We will often identify two smooth manifolds if there is a diffeomorphism from one to the other. Later we will construct invariants (such as de Rham cohomology) that will often enable us to determine that two infinite-dimensional manifolds cannot be diffeomorphic.

Of course, the simplest example of a smooth manifold modeled on \( E \) is an open subset \( U \) of \( E \) in which the atlas is \( \{(U, \text{id}_U)\} \). However, the examples of most interest to us will be function spaces.

**Key Example.** Suppose that \( M^n \) is a complete Riemannian manifold of finite dimension \( n \), which we can assume is isometrically imbedded in Euclidean space \( \mathbb{R}^N \) in accordance with the celebrated Nash imbedding theorem [N56]. Suppose that \( \Sigma \) is a compact smooth manifold of finite dimension \( m \). Then

\[
C^0(\Sigma, \mathbb{R}^N) = \{\text{continuous maps } f : \Sigma \to \mathbb{R}^N\}
\]

is a Banach space, and we claim that the subspace

\[
C^0(\Sigma, M) = \{\text{continuous maps } f : \Sigma \to M\} \subseteq C^0(\Sigma, \mathbb{R}^N)
\]

is an infinite-dimensional Banach manifold, which is modeled on the space of continuous sections of the pullback \( f^*TM \) of the tangent bundle of \( M \) to \( \Sigma \). To define the model space, we let

\[
f^*TM = \{(p, v) \in \Sigma \times TM : f(p) = \pi(v)\},
\]

where \( \pi : TM \to M \) is the usual projection, which is the total space of a vector bundle over \( \Sigma \), which we call the pullback bundle. For simplicity, we often denote this bundle by \( f^*TM \). A continuous section \( X \) of \( f^*TM \) can
be identified with a continuous map

\[ X : \Sigma \to TM \quad \text{such that} \quad X(p) \in T_pM, \quad \text{for} \ p \in \Sigma. \]

**Theorem 1.3.1.** If \( \Sigma \) and \( M \) are finite-dimensional smooth connected manifolds, with \( \Sigma \) compact and \( M \) having a complete Riemannian metric, then \( C^0(\Sigma, M) \) is a smooth manifold, the various components modeled on the Banach spaces

\[ C^0(f^*TM) = \{ \text{continuous sections of} \ f^*TM \}, \]

for choice of smooth map \( f : \Sigma \to M \). Here the norm on \( C^0(f^*TM) \) is

\[ \|X\|_{C^0} = \sup\{|X(p)| : p \in \Sigma\}, \]

the \(|\cdot|\) on the right being length as defined by the Riemannian metric on \( M \).

To construct the charts on \( C^0(\Sigma, M) \) we use the exponential map of \( M \):

**Definition of exponential chart.** For given choices of a smooth map \( f : \Sigma \to M \) and \( \varepsilon > 0 \), we set

\[ V_{f,\varepsilon} = \{ X \in C^0(f^*TM) : \|X\|_{C^0} < \varepsilon \}, \]

and

\[ U_{f,\varepsilon} = \{ g \in C^0(\Sigma, M) : d_M(g(p), f(p)) < \varepsilon \text{ for all } p \in \Sigma \}, \]

where \( d_M : M \times M \to \mathbb{R} \) is the distance function on \( M \) defined by the Riemannian metric. We choose \( \varepsilon > 0 \) sufficiently small that for each \( p \in \Sigma \),

\[ \exp_{f(p)} : \{ v \in T_{f(p)}M : |v| < \varepsilon \} \to \{ q \in M : d_M(q, f(p)) < \varepsilon \} \]

is a diffeomorphism. We then define a bijection \( \psi_{f,\varepsilon} : V_{f,\varepsilon} \to U_{f,\varepsilon} \) by

\[ \psi_{f,\varepsilon}(X)(p) = \exp_{f(p)}(X(p)), \]

and its inverse \( \phi_{f,\varepsilon} : U_{f,\varepsilon} \to V_{f,\varepsilon} \) by

\[ \phi_{f}(g)(p) = \exp_{f(p)}^{-1}(g(p)). \]

We say that \( (U_{f,\varepsilon}, \phi_{f,\varepsilon}) \) is an exponential chart of radius \( \varepsilon \) for \( C^0(\Sigma, M) \) centered at the smooth map \( f : \Sigma \to M \).

**Proof of Theorem 1.3.1.** (Our proof follows the presentation in Abraham [Abr63].) First note that since \( C^0(\Sigma, \mathbb{R}^N) \) is second countable and Hausdorff, so is its subspace \( C^0(\Sigma, M) \). As atlas we take

\[ \mathcal{A} = \{ (U_{f,\varepsilon}, \phi_{f,\varepsilon}) : (U_{f,\varepsilon}, \phi_{f,\varepsilon}) \text{ is an exponential chart} \}. \]

Since smooth maps \( f : \Sigma \to M \) are dense in \( C^0(\Sigma, M) \), the union of the elements of \( \mathcal{A} \) cover \( C^0(\Sigma, M) \). Thus to verify that \( C^0(\Sigma, M) \) is a smooth
manifold, we need only check that \( \phi_{f_2,\varepsilon_2} \circ (\phi_{f_1,\varepsilon_1})^{-1} \) is smooth where defined, when
\[
(U_{f_1,\varepsilon_1}, \phi_{f_1,\varepsilon_1}), (U_{f_2,\varepsilon_2}, \phi_{f_2,\varepsilon_2}) \in \mathcal{A}.
\]

Let \( U \) be the open subset of the total space of \( f_1^*TM \) on which the composition \( (\exp_{f_2(p)})^{-1} \circ \exp_{f_1(p)} \) is defined, and note that \( U \cap T_{f_1(p)}M \) is convex with compact closure for each \( p \in \Sigma \). Define
\[
g : U \to (\text{total space of } f_2^*TM)
\]
by
\[
g(p, v) = (p, (\exp_{f_2(p)})^{-1} \circ \exp_{f_1(p)}(v)), \quad \text{for } p \in \Sigma, \ v \in T_{f_1(p)}M.
\]
We can think of \( g \) as a family of smooth maps
\[
p \mapsto \tilde{g}_p : T_{f_1(p)}M \cap U \to T_{f_2(p)}M,
\]
and since \( \Sigma \) is compact and \( U \) has compact closure in each fiber, all derivatives \( D^k(\tilde{g}_p) \) are bounded on \( U \). Noting that \( U \) is an open neighborhood of the image of the zero-section, we set
\[
\tilde{U} = \{ X \in C^0(f_1^*TM) : X(\Sigma) \subseteq U \}
\]
and define \( \omega_g : \tilde{U} \to C^0(f_2^*TM) \) by \( \omega_g(X) = g \circ X \). To finish the proof that \( C^0(\Sigma, M) \) is a smooth manifold, it suffices to show that \( \omega_g \) is smooth. This is a special case of what is sometimes called the “\( \omega \)-lemma”:

**Lemma 1.3.2.** Suppose that \( E_1 \) and \( E_2 \) are finite-dimensional vector bundles over the compact smooth manifold \( \Sigma \) and that \( U \) is a bounded open neighborhood of the image of the zero section in \( E_1 \) whose restriction to each fiber of \( E_1 \) is convex. If \( g : U \to (\text{total space of } E_2) \) is a smooth map which takes the fiber of \( E_1 \) over \( p \) to the fiber of \( E_2 \) over \( p \) (for each \( p \in \Sigma \)), and
\[
\tilde{U} = \{ \sigma \in C^0(E_1) : \sigma(\Sigma) \subseteq U \},
\]
then the map \( \omega_g : \tilde{U} \to C^0(E_2) \), defined by \( \omega_g(\sigma)(p) = g(\sigma(p)) \) for \( p \in \Sigma \), is also smooth.

As in the example above, we can think of \( g \) as providing a smooth collection of maps
\[
p \mapsto \tilde{g}_p : (E_1)_p \cap U \to (E_2)_p,
\]
where \( (E_1)_p \) and \( (E_2)_p \) are the fibers of \( E_1 \) and \( E_2 \) at \( p \in \Sigma \).

**Proof.** It is straightforward to show that \( \omega_g \) is continuous. The idea for proving that \( \omega_g \) is \( C^1 \) is to use the first-order pointwise Taylor expansion for each \( \tilde{g}_p \) to determine a corresponding Taylor expansion on sections. Suppose
that \( \sigma \) and \( \eta \) are \( C^0 \) sections of \( E_1 \) with \( \sigma \) and \( \sigma + \eta \) lying in \( \tilde{U} \). It follows from Taylor’s theorem for any given \( p \in \Sigma \) that

\[
\tilde{g}_p((\sigma + \eta)(p)) = \tilde{g}_p(\sigma(p)) + D\tilde{g}_p(\sigma(p))\eta(p) + R(\sigma, \eta)(p)(\eta(p)),
\]

where

\[
R(\sigma, \eta)(p)(\eta(p)) = \left[ \int_0^1 [D\tilde{g}_p((\sigma + t\eta)(p)) - D\tilde{g}_p(\sigma(p))]dt \right] \eta(p).
\]

We can write this on the section level as

\[
\omega_g(\sigma + \eta) = \omega_g(\sigma) + T(\sigma)(\eta) + R(\sigma, \eta)\eta,
\]

where

\[
T(\sigma)(\eta)(p) = D\tilde{g}_p(\sigma(p))\eta(p)
\]

and \( R(\sigma, \eta) \) is the remainder term given by (1.8). Note that

\[
\|T(\sigma)(\eta)\|_{C^0} = \sup_{p \in \Sigma} |D\tilde{g}_p(\sigma(p))\eta(p)| \leq \sup_{p \in \Sigma} |D\tilde{g}_p(\sigma(p))| \|\eta\|_{C^0},
\]

so \( T(\sigma) \) is a continuous linear map from a neighborhood of 0 in \( C^0(E_1) \) to \( C^0(E_2) \), and it extends to a linear map from \( C^0(E_1) \) to \( C^0(E_2) \). To see that \( \omega_g \) has a derivative at \( \sigma \) given by

\[
(1.9) \quad D\omega_g(\sigma) = T(\sigma), \quad D\omega_g(\sigma)(\eta)(p) = D\tilde{g}_p(\sigma(p))\eta(p),
\]

we need to estimate the error term \( R(\sigma, \eta) \). But uniform continuity of \( Dg \) over \( \tilde{U} \) shows that

\[
\sup_{p \in \Sigma} \left| \int_0^1 [D\tilde{g}_p((\sigma + t\eta)(p)) - D\tilde{g}_p(\sigma(p))]dt \right| \leq (\text{constant})\|\eta\|_{C^0},
\]

so we get the needed estimate

\[
\|R(\sigma, \eta)\|_{C^0} \leq (\text{constant})\|\eta\|_{C^0}^2 = o(\eta),
\]

so \( T(\sigma) \) is indeed the derivative. Since \( D\tilde{g}_p \) is uniformly continuous on \( \tilde{U} \), it is relatively straightforward to show that \( D\omega_g \) is continuous, establishing that \( \omega_g \) is \( C^1 \).

We would like to extend this argument to higher derivatives, and for this we factor the derivative given by (1.9): Recalling that \( g : U \to (\text{total space of } E_2) \), we define a corresponding map

\[
\tilde{D}g : U \to (\text{total space of } L(E_1, E_2))
\]

by setting \( \tilde{D}g(v) = D\tilde{g}_p(v) \), for \( v \in (E_1)_p \cap U \).

(We can regard \( \tilde{D}g \) as a “partial derivative” of \( g \) in which the point \( p \in \Sigma \) is held fixed.) We can then define

\[
\omega_{\tilde{D}g} : \tilde{U} \to C^0(L(E_1, E_2)) \quad \text{by} \quad \omega_{\tilde{D}g}(\sigma)(p) = D\tilde{g}_p(\sigma(p)).
\]
Using the fact that \( C^0(\Sigma, \mathbb{R}) \) is a Banach algebra, we can show that the map

\[
A : C^0(L(E_1, E_2)) \to L(C^0(E_1), C^0(E_2))
\]

defined by \( A(T)(\sigma)(p) = T(p) \cdot \sigma(p) \),

is continuous linear, and this provides the desired factorization,

\[
D\omega_g = A \cdot \omega_{\tilde{D}g}.
\]

The first-order Taylor expansion argument presented in the preceding paragraph now shows that since \( \tilde{D}g \) is \( C^\infty \), \( \omega_{\tilde{D}g} \) is \( C^1 \), from which we conclude that \( D\omega_g \) is \( C^1 \), and hence \( \omega_g \) is \( C^2 \).

Now we use an induction on \( k \in \mathbb{N} \):

\[
\omega_{\tilde{D}g} \text{ is } C^k \quad \rightarrow \quad D\omega_g \text{ is } C^k \quad \rightarrow \quad \omega_g \text{ is } C^{k+1}.
\]

This then implies that \( \omega_g \) is \( C^\infty \), and proves Lemma 1.3.2 and Theorem 1.3.1.

QED

The above lemma also has the following consequence:

**Theorem 1.3.3.** If \( g : M \to N \) is a smooth map, then the map

\[
\omega_g : C^0(\Sigma, M) \to C^0(\Sigma, N) \quad \text{defined by} \quad \omega_g(f) = g \circ f,
\]

is also smooth.

To prove the “moreover” part of the theorem, we need to show that if \( f_1 : \Sigma \to M \) and \( f_2 : \Sigma \to N \) are smooth, then

\[
\phi_{f_2,\varepsilon_2} \circ \omega_g \circ \phi_{f_1,\varepsilon_1}^{-1} \quad \text{is smooth where defined.}
\]

The proof of this is a straightforward application of Lemma 1.3.2.

Unfortunately, the manifolds \( C^0(\Sigma, M) \) are not sufficient for constructing a global theory of partial differential equations. We need to be able to differentiate elements in our function spaces. Thus we need to start off with a somewhat stronger Banach space than the space \( C^0(\Sigma, \mathbb{R}) \) of continuous real-valued functions on \( \Sigma \). Thus for \( k \in \mathbb{N} \), we are led to consider the space \( C^k(\Sigma, \mathbb{R}) \) of real-valued functions on \( \Sigma \) which have continuous derivatives up to order \( k \), a Banach space with respect to the norm

\[
\|f\|_{C^k} = \sup \{ \|f\|(p) + \|Df\|(p) + \cdots + \|D^k f\|(p) : p \in \Sigma \}.
\]

In fact, it is easily checked that if \( f, g \in C^k(\Sigma, \mathbb{R}) \), then

\[
\|fg\|_{C^k} \leq (\text{constant})\|f\|_{C^k}\|g\|_{C^k},
\]

where (constant) denotes a positive constant, so after possibly replacing \( \| \cdot \|_{C^k} \) with an equivalent norm, \( C^k(\Sigma, \mathbb{R}) \) is in fact a Banach algebra.
1.3. Manifolds modeled on Banach spaces

More generally, we can consider the space \( C^k(\Sigma, \mathbb{R}^N) \) of \( \mathbb{R}^N \)-valued functions on \( \Sigma \) which have continuous derivatives up to order \( k \), which is also a Banach space with norm defined by (1.10). The Banach algebra condition (1.11) ensures that we can define a continuous multiplication

\[
\mu : C^k(\Sigma, L(\mathbb{R}^N, \mathbb{R}^M)) \times C^k(\Sigma, \mathbb{R}^N) \to C^k(\Sigma, \mathbb{R}^M)
\]

by simply multiplying in the range.

If \( M \) is an \( n \)-dimensional Riemannian manifold which is isometrically imbedded in \( \mathbb{R}^N \), we let \( C^k(\Sigma, M) = \{ f \in C^k(\Sigma, \mathbb{R}^N) : f(p) \in M \text{ for all } p \in \Sigma \} \).

We claim that \( C^k(\Sigma, M) \) is a smooth manifold.

**Theorem 1.3.4.** If \( \Sigma \) is a compact smooth manifold and \( M \) is a complete Riemannian manifold, both connected and of finite dimension, then \( C^k(\Sigma, M) \) is a smooth manifold, the various components modeled on the Banach spaces

\[
C^k(f^*TM) = \{ C^k \text{ sections of } f^*TM \},
\]

where \( f : \Sigma \to M \) is a smooth map. Here the norm on \( C^k(f^*TM) \) is

\[
\|X\|_{C^k} = \sup \{|X(p)| + |\nabla X| + \cdots + |\nabla^k X| : p \in \Sigma\},
\]

where \( \nabla \) is the Levi-Civita connection and the \( |\cdot| \) is length defined by the Riemannian metric on \( M \).

To prove this, we first note that since \( C^k(\Sigma, \mathbb{R}^N) \) is second countable and Hausdorff, so is its subspace \( C^k(\Sigma, M) \). The charts are just restrictions of exponential charts for \( C^0(\Sigma, M) \): Suppose that \( f : \Sigma \to M \) is \( C^\infty \) and \( (U_{f,\varepsilon}, \phi_{f,\varepsilon}) \) is an exponential chart on \( C^k(\Sigma, M) \). We then set

\[
V_{f,\varepsilon}^k = \{ X \in C^k(f^*TM) : \|X\|_{C^0} < \varepsilon \} = V_{f,\varepsilon} \cap C^k(f^*TM),
\]

and

\[
U_{f,\varepsilon}^k = \{ g \in C^k(\Sigma, M) : d_M(g(p), f(p)) < \varepsilon \text{ for all } p \in \Sigma \} = U_{f,\varepsilon} \cap C^k(\Sigma, M).
\]

We define \( \psi_{f,\varepsilon}^k : V_{f,\varepsilon}^k \to U_{f,\varepsilon}^k \) by

\[
\psi_{f,\varepsilon}^k(X)(p) = \exp_{f(p)}(X(p))
\]

and its inverse \( \phi_{f,\varepsilon}^k : U_{f,\varepsilon}^k \to V_{f,\varepsilon}^k \) by

\[
\phi_f(g)(p) = \exp_{f(p)}^{-1}(g(p)).
\]

This gives an exponential chart \( (U_{f,\varepsilon}^k, \phi_{f,\varepsilon}^k) \) for \( C^k(\Sigma, M) \).

As smooth atlas for \( C^k(\Sigma, M) \), we take

\[
\mathcal{A} = \left\{ (U_{f,\varepsilon}^k, \phi_{f,\varepsilon}^k) : (U_{f,\varepsilon}^k, \phi_{f,\varepsilon}^k) \text{ is an exponential chart on } C^k(\Sigma, M) \right\}.
\]
Since smooth maps \( f : \Sigma \to M \) are dense in \( C^k(\Sigma, M) \), the union of the elements of \( \mathcal{A} \) cover \( C^k(\Sigma, M) \), and to verify that it is a smooth manifold, we need only check that \( \phi^k_{f_2,\varepsilon_2} \circ (\phi^k_{f_1,\varepsilon_1})^{-1} \) is smooth where defined, when \( (U^k_{f_1,\varepsilon_1}, \phi^k_{f_1,\varepsilon_1}), (U^k_{f_2,\varepsilon_2}, \phi^k_{f_2,\varepsilon_2}) \in \mathcal{A} \).

But this follows from a corresponding \( \omega \)-lemma:

**Lemma 1.3.5.** Suppose that \( E_1 \) and \( E_2 \) are finite-dimensional vector bundles over the compact smooth manifold \( \Sigma \) and that \( U \) is a bounded open neighborhood of the zero section of \( E_1 \) whose restriction to each fiber of \( E_1 \) is convex. If \( g : U \to (\text{total space of } E_2) \) is a smooth map which takes the fiber of \( E_1 \) over \( p \) to the fiber of \( E_2 \) over \( p \) (for each \( p \in \Sigma \)), and

\[
\hat{U} = \{ \sigma \in C^k(E_1) : \sigma(\Sigma) \subseteq U \},
\]

then the map \( \omega_g : \hat{U} \to C^k(E_2) \), defined by \( \omega_g(\sigma) = g(\sigma) \), is also smooth.

The proof is virtually identical to the proof given for Lemma 1.3.2. The proof extends to \( C^k \) maps because the space \( C^k(\Sigma, \mathbb{R}) \) has two key properties:

1. it is a Banach algebra, and
2. there is a continuous inclusion from \( C^k(\Sigma, \mathbb{R}) \) into the Banach algebra \( C^0(\Sigma, \mathbb{R}) \) of continuous functions.

Moreover, as before, we have the following:

**Theorem 1.3.6.** If \( g : M \to N \) is a smooth map, then the map

\[
\omega_g : C^k(\Sigma, M) \to C^k(\Sigma, N) \quad \text{defined by} \quad \omega_g(f) = g \circ f,
\]

is also smooth.

To summarize, for each pair \( (\Sigma, M) \) of finite-dimensional smooth manifolds, with \( \Sigma \) compact, we have a chain of Banach manifolds,

\[
\cdots \subseteq C^{k+1}(\Sigma, M) \subseteq C^k(\Sigma, M) \subseteq C^{k-1}(\Sigma, M) \subseteq \cdots \subseteq C^0(\Sigma, M).
\]

The intersection of these manifolds is the space \( C^\infty(\Sigma, M) \) of \( C^\infty \) maps from \( \Sigma \) to \( M \), which could be made into a Fréchet manifold, but we will not describe the details of that construction.

We can formulate many problems from the calculus of variations in terms of the infinite-dimensional manifolds that we have constructed. Thus, for example, we can define the **action function**

\[ J : C^1(S^1, M) \to \mathbb{R} \quad \text{by} \quad J(\gamma) = \frac{1}{2} \int_{S^1} |\gamma'(t)|^2 dt, \]
and check without much difficulty that \( J \) is a smooth map. As we learned in elementary differential geometry courses, the “critical points” for \( J \) are the smooth closed geodesics on \( M \).

Suppose that \( \Sigma \) is a compact two-dimensional Riemann surface. Thus we can imagine that \( \Sigma \) has a Riemannian metric, but we forget about the metric except for its conformal equivalence class, which we denote by \( \omega \). Suppose that

\[
\{(U_\alpha, (x_\alpha, y_\alpha)) : \alpha \in A\}
\]

is an atlas of isothermal coordinate charts on \( \Sigma \), and let \( \{\psi_\alpha : \alpha \in A\} \) be a partition of unity subordinate to \( \{U_\alpha : \alpha \in A\} \). We can then define the Dirichlet integral \( E_\omega : C^1(\Sigma, M) \to \mathbb{R} \) by

\[
E_\omega(f) = \frac{1}{2} \int_\Sigma \sum_{\alpha \in A} \psi_\alpha \left( \frac{\partial f}{\partial x_\alpha} \right)^2 + \left( \frac{\partial f}{\partial y_\alpha} \right)^2 \right) dx_\alpha dy_\alpha.
\]

Once again it is relatively easy to check that \( E_\omega \) is a smooth map on the infinite-dimensional manifold \( C^1(\Sigma, M) \). Later we will call “critical points” for \( E_\omega \) harmonic maps.

More generally, suppose that \( \Sigma \) is an \( m \)-dimensional Riemannian manifold with Riemannian metric expressed in local coordinates \((x^1, \ldots, x^m)\) on \( \Sigma \) by

\[
\sum_{a,b=1}^m \eta_{ab} dx^a dx^b.
\]

If \( f : \Sigma \to M \subseteq \mathbb{R}^N \) is a smooth map and \((\eta^{ab})\) denotes the matrix inverse to \((\eta_{ab})\), we set

\[
|df|^2 = \sum_{a,b=1}^m \eta^{ab} \frac{\partial f}{\partial x^a} \cdot \frac{\partial f}{\partial x^b} \quad \text{and} \quad dA = \sqrt{\det(\eta_{ab})} dx^1 \cdots dx^m.
\]

We can then define the Dirichlet integral

\[
E : C^1(\Sigma, M) \to \mathbb{R} \quad \text{by} \quad E(f) = \frac{1}{2} \int_\Sigma |df|^2 dA,
\]

which is once again a smooth real-valued function on the infinite-dimensional manifold \( C^1(\Sigma, M) \). In the case where the domain is two-dimensional, choice of isothermal parameters leads to exactly the same integrand as before, so this generalizes the previous energy functions to higher dimensional domains.

1.4. The basic mapping spaces

For the study of geodesics, harmonic and minimal surfaces and pseudoholomorphic curves, as well as other nonlinear partial differential equations, we need a collection of function spaces with weak enough topologies that we
can prove convergence of a sequence which is tending toward an infimum of energy on a given component. The infinite-dimensional manifolds that have proven to be most useful in this regard are those modeled on Sobolev spaces. In this section, we describe the simplest of these spaces.

If $\Sigma$ is a compact Riemannian manifold, we can define an inner product $\langle \cdot, \cdot \rangle$ on the space $C^\infty(\Sigma, \mathbb{R})$ of smooth real-valued maps by

$$
(f, g) = \int_\Sigma (fg + \langle Df, Dg \rangle) dA,
$$

where the inner product $\langle \cdot, \cdot \rangle$ on the right is the usual inner product in the cotangent space defined by the Riemannian metric on $\Sigma$, and $dA$ denotes the area or volume form on $\Sigma$. The inner product $\langle \cdot, \cdot \rangle$ makes the space $C^\infty(\Sigma, \mathbb{R})$ of smooth functions into a pre-Hilbert space. Any pre-Hilbert space has a Hilbert space completion, the set of equivalence classes of Cauchy sequences, as described at the beginning of §1.2. The Hilbert space completion in our case is denoted by $L^2_1(\Sigma, \mathbb{R})$, and is called the Sobolev space of $L^2_1$-functions on $\Sigma$.

A second important Sobolev space generalizes the $L^p$ spaces studied in real analysis when $1 \leq p < \infty$. We start by defining a norm $\| \cdot \|_{L^p_1}$ on $C^\infty(\Sigma, \mathbb{R})$ by

$$
\left( \| f \|_{L^p_1} \right)^p = \int_\Sigma (|f|^p + |Df|^p) dA,
$$

where $|Df|$ is the length calculated with respect to the Riemannian metric on $\Sigma$. This norm makes $C^\infty(\Sigma, \mathbb{R})$ into a pre-Banach space and we can construct the Banach space completion as before. This Banach space completion is denoted by $L^p_1(\Sigma, \mathbb{R})$ and is called the Sobolev space of $L^p_1$-functions on $\Sigma$.

Of course, when $p = 2$ this reduces to the previous example.

If we change the Riemannian metric on $\Sigma$ we change the norm to an equivalent norm, but remember that equivalent norms determine the same open sets and yield the same completion. It is actually “Banachable spaces”, topological vector spaces whose topology comes from a Banach space norm, that are the objects of interest.

We can also define versions of these Sobolev spaces for higher order derivatives. Thus we can define the Sobolev norm $\| \cdot \|_{L^p_k}$ on $C^\infty(\Sigma, \mathbb{R})$ by

$$
\left( \| f \|_{L^p_k} \right)^p = \int_\Sigma (|f|^p + |Df|^p + \cdots + |D^k f|^p) dA,
$$

and construct the completion with respect to this norm, which is denoted $L^p_k(\Sigma, \mathbb{R})$. The resulting space is a Banach space, and a Hilbert space when $p = 2$. We thus obtain a chain of Banach spaces,

$$
\cdots \subset L^p_k(\Sigma, \mathbb{R}) \subset \cdots \subset L^p_1(\Sigma, \mathbb{R}) \subset L^p(\Sigma, \mathbb{R}),
$$
and the intersection of all the spaces in the chain is just the space $C^\infty(\Sigma, \mathbb{R})$ of smooth functions. These spaces are essential for the modern theory of partial differential equations and they are compared by means of the Sobolev Lemma:

**Sobolev Lemma.** When $p$ and $k$ are sufficiently large, $L^p_k(\Sigma, \mathbb{R})$ is a subspace of $C^0(\Sigma, \mathbb{R})$:

\[
L^p_k(\Sigma, \mathbb{R}) \subseteq C^0(\Sigma, \mathbb{R}) \quad \text{for} \quad pk > \dim(\Sigma).
\]

**Banach Algebra Lemma.** When $pk > \dim(\Sigma)$, $L^p_k(\Sigma, \mathbb{R})$ is a Banach algebra in an appropriate norm defining the “Banachable” structure.

In this section, we will prove two basic cases of these lemmas. More general cases can be found in standard references on partial differential equations, such as Evans [Eva98]. For the case where $\Sigma = S^1$, where $S^1$ is the unit interval $[0, 1]$ with endpoints identified, the Sobolev Lemma for $L^2_1(S^1, \mathbb{R})$ is relatively easy to establish:

**Lemma 1.4.1.** There is a continuous linear injection $i : L^2_1(S^1, \mathbb{R}) \to C^0(S^1, \mathbb{R})$ which extends the inclusion $C^\infty(S^1, \mathbb{R}) \subset C^0(S^1, \mathbb{R})$.

**Proof.** We begin with the sequence of inequalities:

\[
|f(t)| \leq |f(\tau)| + \int_t^\tau |f'(u)|du \leq |f(\tau)| + \int_{S^1} |f'(u)|du,
\]

for $t \leq \tau$ (which we can always arrange by replacing $\tau$ by an equivalent parameter $\tau$ with $\tau \in [t, t+1]$ if necessary). Averaging over $\tau$ and using the Cauchy-Schwarz inequality yields

\[
|f(t)| \leq \left[ \int_{S^1} |f(u)|du + \int_{S^1} |f'(u)|du \right]^{1/2} \leq \left[ \int_{S^1} |f(u)|^2du \right]^{1/2} + \left[ \int_{S^1} |f'(u)|^2du \right]^{1/2} \leq (f, f)^{1/2},
\]

where $(\cdot, \cdot)$ denotes the $L^2_1$ inner product. Taking the supremum over all $t$ yields

\[
\|f\|_{C^0} \leq \|f\|_{L^2_1}.
\]

Thus a Cauchy sequence with respect to the $L^2_1$ inner product gets taken under the inclusion $C^\infty(S^1, \mathbb{R}) \subset C^0(S^1, \mathbb{R})$ to a Cauchy sequence with respect to the $C^0$-norm. By definition, an element of $L^2_1$ is an equivalence class of
Cauchy sequences, and the map $i$ is defined by sending this equivalence class to the limit of the $C^0$ Cauchy sequence. It is immediate that $i$ is injective. QED

**Lemma 1.4.2.** After possibly replacing $\| \cdot \|_{L^2_1}$ with an equivalent norm, $L^2_1(S^1, \mathbb{R})$ is a Banach algebra, so multiplication of functions is a continuous bilinear map:

$$L^2_1(S^1, \mathbb{R}) \times L^2_1(S^1, \mathbb{R}) \rightarrow L^2_1(S^1, \mathbb{R}).$$

**Proof.** It follows from the Cauchy-Schwarz inequality that

$$\|fg\|^2_{L^2_1} = (fg, fg) = \int_{S^1} [(fg)^2 + (fg')^2] dt = \int_{S^1} [(fg)^2 + (f'g + fg')^2] dt$$

$$= \int_{S^1} [(fg)^2 + (f')^2 g^2 + 2fgf'g' + f^2(g')^2] dt$$

$$\leq \int_{S^1} [(fg)^2 + (f')^2 g^2 + f^2(g')^2] dt + 2\|fg\|_{C^0} \int_{S^1} (f'g') dt$$

$$\leq \|f\|_{C^0}^2 \|g\|^2_{L^2_1} + \|g\|_{C^0}^2 \|f\|^2_{L^2_1} + \|f\|_{C^0}^2 \|g\|^2_{L^2_1} + 2\|fg\|_{C^0} \|f\|_{L^2_1} \|g\|_{L^2_1}.$$  

It then follows from (1.13) that

$$\|fg\|^2_{L^2_1} \leq (\text{constant}) \|f\|^2_{L^2_1} \|g\|^2_{L^2_1};$$

where (constant) denotes a positive constant, finishing the proof of the lemma. QED

It follows from this lemma that the multiplication map

$$L^2_1(S^1, \text{Hom}(\mathbb{R}^m, \mathbb{R}^n)) \times L^2_1(S^1, \mathbb{R}^m) \rightarrow L^2_1(S^1, \mathbb{R}^n)$$

is continuous.

Suppose now that $M$ is a complete connected finite-dimensional Riemannian manifold isometrically imbedded as a proper submanifold of an ambient Euclidean space $\mathbb{R}^N$. We let

$$L^2_1(S^1, M) = \{ \gamma \in L^2_1(S^1, \mathbb{R}^N) : \gamma(t) \in M \text{ for all } t \in S^1 \},$$

which is a closed subspace of $L^2_1(S^1, \mathbb{R}^N)$ by Lemma 1.4.1. We claim that $L^2_1(S^1, M)$ is an infinite-dimensional smooth manifold, the proof being just like the proof for $C^0(S^1, M)$. If $\gamma : S^1 \rightarrow M$ is a smooth curve, we choose $\varepsilon > 0$ so that for every $t \in S^1$,

$$\exp_{\gamma(t)} : \{v \in T_{f(p)}M : |v| < \varepsilon\} \rightarrow \{q \in M : d_M(q, \gamma(t)) < \varepsilon\}$$
1.4. Basic mapping spaces

is a diffeomorphism. We then set

\[ V_{\gamma, \varepsilon} = \{ X \in L^2_1(S^1, \gamma^*TM) : \| X \|_{C^0} < \varepsilon \}, \]
\[ U_{\gamma, \varepsilon} = \{ \lambda \in L^2_1(S^1, M) : d_M(\lambda(t), \gamma(t)) < \varepsilon \text{ for all } t \in S^1 \}, \]

and define

\[ \phi_{\gamma, \varepsilon} : U_{\gamma, \varepsilon} \to V_{\gamma, \varepsilon} \text{ by } \phi_{\gamma, \varepsilon}(\lambda)(t) = \exp_{\gamma(t)}^{-1}(\lambda(t)). \]

Finally, we prove that

\[ A = \{ (U_{\gamma, \varepsilon}, \phi_{\gamma, \varepsilon}) : \gamma : S^1 \to M \text{ is a } C^\infty \text{ map and } \varepsilon > 0 \text{ is small enough that } \phi_{\gamma, \varepsilon} \text{ is a homeomorphism} \}. \]

is a smooth atlas for \( L^2_1(S^1, M) \) by the same argument used to establish Lemma 1.3.2. Noting that \( L^2_1(\gamma^*TM) \) is actually a Hilbert space, we obtain the following:

**Theorem 1.4.3.** If \( M \) is a smooth manifold, then \( L^2_1(S^1, M) \) is a smooth manifold modeled on the Hilbert spaces \( L^2_1(\gamma^*TM) \) for \( \gamma : S^1 \to M \) a smooth map. Moreover, if \( g : M \to N \) is a smooth map, then the map

\[ \omega_g : L^2_1(S^1, M) \to L^2_1(S^1, N) \text{ defined by } \omega_g(\gamma) = g \circ \gamma, \]

is also smooth.

We are also interested in \( L^p_1 \)-maps from a compact oriented surface \( \Sigma \) when \( p > 2 \). It turns out that these are Hölder continuous in accordance with the following definition.

**Definition.** If \( \Sigma \) is a metric space, a map \( f : \Sigma \to \mathbb{R} \) is said to be Hölder continuous with Hölder exponent \( \gamma \in (0, 1] \) if there is a constant \( C > 0 \) such that

\[ |f(p) - f(q)| \leq C \ d(p, q)\gamma \quad \text{for all } p, q \in \Sigma. \]

We let \( C^{0, \gamma}(\Sigma, \mathbb{R}) \) be the space of all functions \( f : \Sigma \to \mathbb{R} \) which are Hölder continuous. If \( f \in C^{0, \gamma}(\Sigma, \mathbb{R}) \), we let

\[ [f]_{C^{0, \gamma}} = \sup \left\{ \frac{f(p) - f(q)}{d(p, q)^\gamma} : p, q \in \Sigma, p \neq q \right\}. \]

The following lemma states that the space \( L^p_1(\Sigma, \mathbb{R}) \) lies within Sobolev range.

**Lemma 1.4.4.** If \( \Sigma \) is a compact oriented surface and \( p > 2 \), there is a continuous linear injection \( i : L^p_1(\Sigma, \mathbb{R}) \to C^{0, \gamma}(\Sigma, \mathbb{R}) \), where \( \gamma = 1 - 2/p \).
which extends the inclusion $C^\infty(\Sigma, \mathbb{R}) \subset C^{0,\gamma}(\Sigma, \mathbb{R})$. Moreover, there is a constant $C > 0$ such that

$$[f]_{C^{0,\gamma}} \leq C \|f\|_{L^p_1}. $$

A complete proof of this is given in Evans [Eva98], §5.6.2. We only prove the weaker result that $L^p_1(\Sigma, \mathbb{R}) \subseteq C^{0,\gamma}(\Sigma, \mathbb{R})$. We begin by assuming that $\Sigma$ is the torus with flat Riemannian metric expressed in terms of suitable conformal coordinates as $ds^2 = dx^2 + dy^2$. We consider a smooth function $f(x,y)$ on the disk $D(p,r_0)$ of radius $r_0$ about $p \in \Sigma$ defined in terms of Euclidean coordinates centered at $p$ by $x^2 + y^2 \leq r_0^2$, then after shifting to polar coordinates $(r,\theta)$ defined by

$$x = r \cos \theta, \quad y = r \sin \theta,$$

we see that

$$|f(s,\theta) - f(p)| = \int_0^s \frac{\partial f}{\partial r}(r,\theta) dr \leq \int_0^s |Df|(r,\theta) dr,$$

and hence

$$\int_0^{2\pi} |f(s,\theta) - f(p)| d\theta \leq \int_0^{2\pi} \int_0^s |Df|(r,\theta) dr d\theta \leq \int_{D(p,r_0)} \frac{|Df|}{r} dxdy.$$

Thus

$$\int_0^{2\pi} \int_0^{r_0} |f(s,\theta) - f(p)| ds d\theta \leq \left[ \int_0^{r_0} r dr \right] \left[ \int_{D(p,r_0)} \frac{|Df|}{r} dxdy \right]$$

and hence

$$\int_{D(p,r_0)} |f(x,y) - f(p)| dxdy \leq \frac{r_0^2}{2} \int_{D(p,r_0)} \frac{|Df|}{r} dxdy.$$

It follows from the Hölder inequality that

$$\frac{r_0^2}{2} \int_{D(p,r_0)} \frac{|Df|}{r} dxdy \leq \frac{r_0^2}{2} \left[ \int_{D(p,r_0)} |Df|^p dxdy \right]^{1/p} \left[ \int_{D(p,r_0)} dxdy \right]^{(p-1)/p}$$

while direct integration yields

$$\int_{D(p,r_0)} \frac{dxdy}{r^{p/(p-1)}} = \int_{D(p,r_0)} r^{p/(1-p)} dxdy$$

$$= \int_0^{2\pi} \int_0^{r_0} r^{1/(1-p)} dr d\theta = \frac{2\pi (p - 1)}{p - 2} r_0^{(p-2)/(p-1)}.$$
Thus
\begin{equation}
\int_{D(p,r_0)} |f(x, y) - f(p)| \, dx \, dy \leq \frac{r_0^2}{2} \left( \frac{2\pi (p - 1)}{p - 2} \right)^{(p-1)/p} r_0^{(p-2)/p} \|Df\|_{L^p}.
\end{equation}

It follows from (1.14) and the Hölder inequality that
\[
\pi r_0^2 |f(0)| \leq \int_{D(p,r_0)} |f(x, y) - f(p)| \, dx \, dy + \int_{D(p,r_0)} |f(x, y)| \, dx \, dy 
\leq (\text{constant}) \|Df\|_{L^p} + (\text{constant}) \|f\|_{L^1} 
\leq (\text{constant}) \|Df\|_{L^p} + (\text{constant}) \|f\|_{L^p} (\text{area of } D(p, r_0))^{(p-1)/p} 
\leq (\text{constant}) \|f\|_{L^p},
\]

which quickly yields the desired result when \( \Sigma \) is the flat torus.

If \( \Sigma \) is a more general Riemann surface, we can use the uniformization theorem to give \( \Sigma \) a Riemannian metric of constant curvature and total volume one. Choose \( r_0 > 0 \) less than the injectivity radius of this metric. A modification of the above argument can then be applied to a normal coordinate ball of radius \( r_0 \) showing that if \( p \in \Sigma \), then
\[
|f(p)| \leq (\text{constant}) \|f\|_{L^p}, \quad \text{and hence} \quad \|f\|_{C^0} \leq (\text{constant}) \|f\|_{L^p}.
\]
(\text{Note that changing the Riemannian metric on } \Sigma \text{ merely replaces the } L^p \text{-norm by an equivalent norm, so adopting the constant curvature metric imposes no restriction.}) Thus if a sequence \( \{f_i\} \) of smooth functions on \( \Sigma \) converges to a limit in \( L^p \)-norm, \( \{f_i\} \) converges also in \( C^0 \) norm to a unique limit function \( f_\infty \in C^0 \). Thus any element of \( L^p(\Sigma, \mathbb{R}) \) can be identified with a continuous function and the Lemma is proven.

**Remark 1.4.5.** The previous Sobolev Lemma for \( L^p \) maps from a surface begins to fail as \( p \) approaches 2 from above. The reason is that the highest order term in the \( L^2 \)-norm is conformally invariant and hence invariant under dilations. In the case where \( D \) is the unit disk centered at the origin in \( \mathbb{R}^2 \) this highest order term is
\[
\int_D |Df|^2 \, dx \, dy.
\]
In contrast, the highest order term in the \( L^p \)-norm is not invariant under dilations. Given a smooth map \( f : D \to \mathbb{R}^N \) which takes the boundary \( \partial D_\epsilon \) to a point, we can define a dilated map \( f_\epsilon : D_\epsilon \to \mathbb{R}^N \), where \( D_\epsilon \) is the ball of radius \( \epsilon > 0 \) centered at the origin in \( \mathbb{R}^2 \), by
\[
f_\epsilon(x, y) = f \left( \frac{x}{\epsilon}, \frac{y}{\epsilon} \right).
\]
1. Infinite-dimensional Manifolds

Then

\[ |Df_{\epsilon}(x, y)| = \frac{1}{\epsilon} \left| Df\left(\frac{x}{\epsilon}, \frac{y}{\epsilon}\right) \right| \]

and, if \( \tilde{x} = x/\epsilon \) and \( \tilde{y} = y/\epsilon \) denote the coordinates corresponding to \( x \) and \( y \) on the unit disk \( D_1 \), then

\[ \int_{D_\epsilon} |Df_{\epsilon}|^p dx dy = \left( \frac{1}{\epsilon} \right)^{(p-2)} \int_{D_1} |Df|^{2\alpha} d\tilde{x} d\tilde{y}. \]

Thus, as \( \epsilon \to 0 \), the \( L^p \) norm of \( f_{\epsilon} \) approaches infinity as long as \( p > 2 \). In particular, a bound on the \( L^2 \) norm does not imply a bound on the \( C^0 \)-norm.

**Lemma 1.4.6.** If \( \Sigma \) is a compact oriented surface and \( p > 2 \), \( L^p_1(\Sigma, \mathbb{R}) \) satisfies

\[ \|fg\|_{L^p_1} \leq 2\|f\|_{L^p_1}\|g\|_{L^p_1}. \]

Thus after passing to an equivalent norm, we can show that \( L^p_1(\Sigma, \mathbb{R}) \) is a Banach algebra.

**Sketch of proof.** By the previous lemma,

\[ \|fg\|_{L^p} \leq \|f\|_{C^0}\|g\|_{L^p} + \|g\|_{C^0}\|f\|_{L^p} \leq \|f\|_{L^p_1}\|g\|_{L^p} + \|g\|_{L^p_1}\|f\|_{L^p} \]

and

\[ \|D(fg)\|_{L^p} \leq \|f\|_{C^0}\|Dg\|_{L^p} + \|g\|_{C^0}\|Df\|_{L^p} \leq \|f\|_{L^p_1}\|Dg\|_{L^p} + \|g\|_{L^p_1}\|Df\|_{L^p}. \]

Adding these two inequalities yields the statement of the lemma. QED

If \( \Sigma \) is a compact surface and \( p > 2 \), we let

\[ L^p_1(\Sigma, M) = \{ f \in L^p_1(\Sigma, \mathbb{R}^N) : f(p) \in M \text{ for all } p \in \Sigma \}, \]

a closed subspace of \( L^p_1(\Sigma, \mathbb{R}^N) \) by Lemma 1.4.3.

If \( \Sigma \) is a compact surface and \( p > 2 \), we claim that \( L^p_1(\Sigma, M) \) is an infinite-dimensional smooth manifold. In this case, when \( f : \Sigma \to M \) is a smooth parametrized surface, we let

\[ V_{\gamma, \epsilon} = \{ X \in L^p_1(\Sigma^*TM) : \|X\|_{C^0} < \epsilon \}, \]

\[ U_{\gamma, \epsilon} = \{ g \in L^p_1(\Sigma, M) : d_M(f(p), g(p)) < \epsilon \text{ for all } p \in \Sigma \}, \]

and if \( \epsilon > 0 \) is sufficiently small, we define

\[ \phi_{\gamma, \epsilon} : V_{\gamma, \epsilon} \to U_{\gamma, \epsilon} \text{ by } \phi_{\gamma, \epsilon}(X)(p) = \exp_{f(p)}^{-1}(X(p)). \]

We then set

\[ \mathcal{A} = \{ (U_{\gamma, \epsilon}, \phi_{\gamma, \epsilon}) : f : \Sigma \to M \text{ is a } C^\infty \text{ map and } \epsilon > 0 \text{ is small enough that } \phi_{\gamma, \epsilon} \text{ is a homeomorphism} \} \]
and once again prove that it is a smooth atlas by the same argument used to establish Lemma 1.3.2. Indeed, the same argument holds for $L^p_k(\Sigma, M)$ so long as $pk > 2$, and we obtain:

**Theorem 1.4.7.** If $\Sigma$ is a compact smooth surface and $M$ is a smooth manifold, then for $pk > 2$, $L^p_k(\Sigma, M)$ is a smooth manifold modeled on the Banach spaces $L^p_k(f^*TM)$ for $f : \Sigma \to M$ a smooth map. Moreover, if $g : M \to N$ is a smooth map, then the map

$$\omega_g : L^p_k(\Sigma, M) \to L^p_k(\Sigma, N) \text{ defined by } \omega_g(\gamma) = g \circ \gamma,$$

is also smooth.

In exactly the same way, we could show that if $\Sigma$ is an $m$-dimensional smooth manifold and $pk > m$, then $L^p_k(\Sigma, M)$ is an infinite-dimensional smooth manifold.

In the case where $\Sigma$ is a compact surface, the general Sobolev Lemma (1.12) provide sequences of Banach manifolds and inclusions

$$\cdots \subseteq C^k(\Sigma, M) \subseteq \cdots \subseteq C^2(\Sigma, M) \subseteq L^p_1(\Sigma, M) \subseteq C^0(\Sigma, M)$$

or

$$\cdots \subseteq L^2_k(\Sigma, M) \subseteq \cdots \subseteq L^2_2(\Sigma, M) \subseteq L^p_1(\Sigma, M) \subseteq C^0(\Sigma, M)$$

when $p > 2$. It is important to note that the charts defining the atlases for all the manifolds in these sequences are simply the restrictions of the exponential charts we defined on $C^0(\Sigma, M)$. These sequences are ideally suited to a technique commonly used in the theory of elliptic partial differential equations: one starts by proving existence of a solution in a space such as $L^p_1(\Sigma, M)$, then uses the method of “elliptic bootstrapping” to show that the solution actually lies in $C^k(\Sigma, M)$ for higher and higher values of $k$. The spaces $L^2_k(\Sigma, M)$ in the second sequence have the added advantage of being Hilbert manifolds when $k$ is large, which are sometimes a little easier to work with.

1.5. Homotopy type of the space of maps

A continuous map $f : X \to Y$ between topological spaces is said to be a homotopy equivalence if there is a continuous map $g : Y \to X$ such that $f \circ g$ and $g \circ f$ are both homotopic to the identity, and if such a homotopy equivalence exists, we say that $X$ and $Y$ have the same homotopy type. The following theorem on homotopy equivalence was proven quite early in the theory of manifolds of maps; see the survey article of Eells [Eel66] for appropriate references.
Theorem 1.5.1. Let \( M \) be a compact connected Riemannian manifold. Then the inclusions
\[
C^k(S^1, M) \subseteq C^0(S^1, M) \quad \text{and} \quad C^k(S^1, M) \subseteq L^2_1(S^1, M)
\]
are homotopy equivalences when \( k \geq 1 \).

The point of this theorem is that from the point of view of homotopy theory, \( L^2_1(S^1, M) \) is essentially the same as the space of continuous maps \( C^0(S^1, M) \) with the compact open topology. This latter space has been extensively studied by topologists and much is known about its homotopy and homology groups, as we will see later.

The proof of the theorem is an application of the theory of smoothing operators. In preparation, we suppose that \( \phi : \mathbb{R} \to \mathbb{R} \) is a smooth map which vanishes outside \([-1, 1]\). Suppose, moreover, that \( \phi \geq 0 \) and \( \int_{\mathbb{R}} \phi = 1 \).

For \( \varepsilon > 0 \), let \( \phi_\varepsilon(t) = (1/\varepsilon)\phi(t/\varepsilon) \), so that
\[
supp(\phi_\varepsilon) \subset [-\varepsilon, \varepsilon] \quad \text{and} \quad \int_{\mathbb{R}} \phi_\varepsilon = 1.
\]

If \( \gamma \in C^0(S^1, \mathbb{R}^N) \), we can regard \( \gamma \) as an element of \( C^0(\mathbb{R}, \mathbb{R}^N) \) such that \( \gamma(t + 1) = \gamma(t) \) for all \( t \), and we define \( \phi_\varepsilon \ast \gamma \in C^0(\mathbb{R}, \mathbb{R}^N) \) by
\[
(\phi_\varepsilon \ast \gamma)(t) = \int_{\mathbb{R}} \phi_\varepsilon(t - \tau)\gamma(\tau)\,d\tau = \int_{\mathbb{R}} \phi_\varepsilon(s)\gamma(t - s)\,ds.
\]

It is immediately checked that \( \phi_\varepsilon \ast \gamma \) is \( C^\infty \) and
\[
\frac{d^k}{dt^k}(\phi_\varepsilon \ast \gamma)(t) = \frac{d^k}{dt^k}(\phi_\varepsilon) \ast \gamma = \int_{\mathbb{R}} \left( \frac{d^k}{dt^k}\phi_\varepsilon \right)(t - \tau)\gamma(\tau)\,d\tau.
\]

Moreover, \( (\phi_\varepsilon \ast \gamma)(t + 1) = (\phi_\varepsilon \ast \gamma)(t) \), and hence \( \phi_\varepsilon \ast \gamma \in C^\infty(S^1, \mathbb{R}^N) \). We can thus define smoothing operators
\[
S_\varepsilon : C^0(S^1, \mathbb{R}^N) \to C^k(S^1, \mathbb{R}^N), \quad S_\varepsilon : L^2_1(S^1, \mathbb{R}^N) \to C^k(S^1, \mathbb{R}^N)
\]
by \( S_\varepsilon(\gamma) = \phi_\varepsilon \ast \gamma \). It is not difficult to show that the maps
\[
S_\varepsilon : C^0(S^1, \mathbb{R}^N) \to C^k(S^1, \mathbb{R}^N), \quad S_\varepsilon : L^2_1(S^1, \mathbb{R}^N) \to C^k(S^1, \mathbb{R}^N)
\]
are continuous.

Proof of Theorem 1.5.1. Recall that we regard \( M \) as a submanifold of \( \mathbb{R}^N \). Choose \( \delta > 0 \) so small that the exponential map
\[
\exp : NM \to \mathbb{R}^N, \quad \text{defined by} \quad \exp(v) = p + v, \quad \text{for} \ v \in N_pM,
\]
1.5. Homotopy type of the space of maps

maps $NM_\delta = \{v \in NM : |v| < \delta\}$ diffeomorphically onto

$$M(\delta) = \{p \in \mathbb{R}^N : d(p, M) < \delta\}.$$ 

Then the “nearest point projection” map $r : M(\delta) \to M$, defined by $r(p + v) = p$ for $p \in M$, is a strong deformation retraction from $M(\delta)$ to $M$. To see this, we define

$$h : M(\delta) \times [0, 1] \to M(\delta) \quad \text{by} \quad h(p + v, t) = p + (1 - t)v,$$

and check that

1. $h(q, 0) = q$, for $q \in M(\delta)$,
2. $h(q, 1) = r(q) \in M$, for $q \in M(\delta)$, and
3. $h(p, t) = p$, for $p \in M$.

We have a similar strong deformation retraction on the function space level. The $\omega$-Lemma gives us a smooth map:

$$\omega_h : L^2_1(S^1, M(\delta) \times [0, 1]) \to L^2_1(S^1, M(\delta)), \quad \text{defined by} \quad \omega_h(\gamma) = h \circ \gamma.$$

We define

$$j : L^2_1(S^1, M(\delta)) \times [0, 1] \to L^2_1(S^1, M(\delta) \times [0, 1]) \quad \text{by} \quad j(\gamma, \tau)(t) = (\gamma(t), \tau)$$

and let $H = \omega_h \circ j$. Then

1. $H(\gamma, 0) = \gamma$, for $\gamma \in L^2_1(S^1, M(\delta))$,
2. $H(\gamma, 1) = r \circ \gamma \in L^2_1(S^1, M)$, for $\gamma \in L^2_1(S^1, M(\delta))$, and
3. $H(\gamma, t) = \gamma$, for $\gamma \in L^2_1(S^1, M)$.

We can therefore define a strong deformation retraction

$$R : L^2_1(S^1, M(\delta)) \to L^2_1(S^1, M)$$

by $R(\gamma) = \omega_r(\gamma) = H(\gamma, 1)$. In a similar fashion, we can define a strong deformation retraction

$$R : C^k(S^1, M(\delta)) \to C^0(S^1, M),$$

whenever $k \geq 0$.

Let $\varepsilon_k = 2^{-k}$ and let

$$C^0(S^1, M)^{\varepsilon_k} = \{\gamma \in C^0(S^1, M) : \gamma \text{ maps the closed interval}$$

$$[(m - 1)2^{-k}, (m + 1)2^{-k}] \text{ into the open ball}$$

$$B(\gamma(m2^{-k}); \delta) \text{ for each integer } m \text{ such that } 0 \leq m \leq 2^k\},$$

an open set in the compact-open topology. They key point of this set is that when $|s - t| < 2^{-k}$ then the straight line from $\gamma(s)$ to $\gamma(t)$ in $\mathbb{R}^N$ lies entirely
within $M(\delta)$ hence $S_k \ast \gamma$ lies within $M(\delta)$ when $\varepsilon \leq \varepsilon_k$. Note that

$$C^0(S^1, M) = \bigcup_{k=1}^{\infty} C^0(S^1, M)^{\varepsilon_k}, \quad L^2_1(S^1, M)^{\varepsilon_{k+1}} \subset L^2_1(S^1, M)^{\varepsilon_k},$$

and we therefore say that $C^0(S^1, M)$ is a monotone union of the subspaces $C^0(S^1, M)^{\varepsilon_k}$. Similarly, we let

$$L^2_1(S^1, M)^{\varepsilon_k} = L^2_1(S^1, M) \cap C^0(S^1, M)^{\varepsilon_k},$$

$$C^k(S^1, M)^{\varepsilon_k} = C^k(S^1, M) \cap C^0(S^1, M)^{\varepsilon_k}, \quad \text{when } k \geq 1,$$

thereby expressing $L^2_1(S^1, M)$ and $C^k(S^1, M)$ as monotone unions for $k \geq 1$.

By analogous formulae, we define

$$C^0(S^1, M(\delta))^{\varepsilon_k}, \quad L^2_1(S^1, M(\delta))^{\varepsilon_k} \quad \text{and} \quad C^k(S^1, M(\delta))^{\varepsilon_k},$$

for $k \geq 1$. We can then define smoothing operators

$$S_{\varepsilon_k} : L^2_1(S^1, M)^{\varepsilon_k} \to C^k(S^1, M(\delta))^{\varepsilon_k},$$

since

$$\gamma \in L^2_1(S^1, M)^{\varepsilon_k} \Rightarrow S_{\varepsilon_k} \ast \gamma \in C^k(S^1, M(\delta))^{\varepsilon_k}.$$

We define $s$ to be the composition of

$$S_{\varepsilon_k} : L^2_1(S^1, M)^{\varepsilon_k} \to C^k(S^1, M(\delta))^{\varepsilon_k} \quad \text{and} \quad R : C^k(S^1, M(\delta))^{\varepsilon_k} \to C^k(S^1, M)^{\varepsilon_k}.$$

We claim that if $i : C^k(S^1, M)^{\varepsilon_k} \subset L^2_1(S^1, M)^{\varepsilon_k}$ is the inclusion, then

$$s \circ i \quad \text{and} \quad i \circ s$$

are homotopic to the identity. This is easy to verify. To get the homotopy from $s \circ i$ to the identity, we simply define

$$H_1 : C^k(S^1, M)^{\varepsilon_k} \times [0, 1] \to C^k(S^1, M)^{\varepsilon_k}$$

by

$$H_1(\gamma, t) = R \circ (tS_{\varepsilon_k} + (1-t)\text{id}) \circ i(\gamma).$$

Similarly, to get the homotopy from $i \circ s$ to the identity, we define

$$H_2 : L^2_1(S^1, M)^{\varepsilon_k} \times [0, 1] \to L^2_1(S^1, M)^{\varepsilon_k}$$

by

$$H_2(\gamma, t) = i \circ R \circ (tS_{\varepsilon_k} + (1-t)\text{id})(\gamma),$$

This shows that for each $k \in \mathbb{N}$, the inclusion

$$C^k(S^1, M)^{\varepsilon_k} \subset L^2_1(S^1, M)^{\varepsilon_k}$$

is a homotopy equivalence.
To finish the proof, we must take an appropriate limit as \( k \to \infty \). Suppose that the metrizable space \( X \) is a monotone union of open subsets, by which we mean that we have a sequence of spaces
\[
U_1 \subseteq U_2 \subseteq U_3 \subseteq \cdots \text{ such that } X = \bigcup \{U_k : k \in \mathbb{N}\}.
\]
Suppose, moreover, that we let
\[
X^* = (U_1 \times [1, 2]) \cup (U_2 \times [2, 3]) \cup (U_3 \times [3, 4]) \cup \cdots
\]
topologized as a subset of \( X \times \mathbb{R} \). We say that \( X \) is the homotopy direct limit of the subspaces \( \{U_k : k \in \mathbb{N}\} \) if the projection \( p : X^* \to X \) on the first factor is a homotopy equivalence. If the subsets \( U_k \) are open, then the open cover \( \{U_k : k \in \mathbb{N}\} \) has a \( C^0 \) subordinate partition of unity \( \{\psi_k : k \in \mathbb{N}\} \). In this case, the map
\[
f : X \to X^* \text{ defined by } f(x) = \left(x, \sum_{k=1}^{\infty} k\psi_k(x)\right)
\]
is a homotopy inverse to \( p \), showing that \( X \) is a homotopy direct limit in this case. Using this argument, one easily verifies that \( C^k(S^1, M) \) is a homotopy direct limit of its subspaces \( C^k(S^1, M)^{\varepsilon_k} \) and \( L^2_2(S^1, M) \) is a homotopy direct limit of \( L^2_1(S^1, M)^{\varepsilon_k} \).

Now we apply the following lemma, which is just Theorem A from the Appendix to Milnor’s book on Morse theory \([Mil63]\):

**Lemma 1.5.2.** Suppose that \( X \) is the homotopy direct limit of \( \{U_k : k \in \mathbb{N}\} \) and that \( Y \) is the homotopy direct limit of \( \{V_k : k \in \mathbb{N}\} \). If \( f : X \to Y \) is a continuous map such that \( f(U_k) \subseteq V_k \) and the restriction of \( f \) to \( U_k \) is a homotopy equivalence from \( U_k \) to \( V_k \), then \( f \) itself is a homotopy equivalence.

We refer the reader to Milnor for the proof of this lemma. It implies that the inclusion \( C^k(S^1, M) \subset L^2_1(S^1, M) \) is a homotopy equivalence when \( k \geq 1 \). In a similar manner, one verifies that the inclusion \( C^k(S^1, M) \subset C^0(S^1, M) \) is a homotopy equivalence when \( k \geq 1 \), and this completes the proof of Theorem 1.5.1. QED

**Theorem 1.5.3.** Let \( M \) be a compact connected Riemannian manifold, \( \Sigma \) a compact connected Riemann surface. Then the inclusions,
\[
C^k(\Sigma, M) \subseteq C^0(\Sigma, M), \quad C^k(\Sigma, M) \subseteq L_1^p(\Sigma, M) \quad \text{for } p > 2, \quad C^k(\Sigma, M) \subseteq L_2^2(\Sigma, M) \quad \text{for } k \geq 2,
\]
are homotopy equivalences.

The proof is essentially the same as for the previous theorem, with 2 replaced by \( p \), except for the definition of smoothing operators defined on \( \Sigma \). The
construction of such operators is a standard technique in the theory of partial differential equations. We describe only the case $\Sigma = T^2$ here, where $T^2 = \mathbb{C}/\Lambda$, $\Lambda$ being a lattice in $\mathbb{C}$; the construction in this case is particularly transparent. (In the general case, the ideas are the same, but one constructs the smoothing operators by piecing together using a partition of unity on $\Sigma$.)

Note that an element $f \in L^p_1(T^2, \mathbb{R}^N)$ can be regarded as a map $f : \mathbb{C} \to \mathbb{R}^N$ such that $f(z + \lambda) = f(z)$ for $\lambda \in \Lambda$.

Suppose that $\phi : \mathbb{C} \to [0, \infty)$ is a smooth map which vanishes outside $D = \{ z \in \mathbb{C} : |z| \leq 1 \}$ such that

$$\int_{\mathbb{C}} \phi \, dx dy = 1,$$

where $(x, y)$ are the standard coordinates on $\mathbb{C}$. Let $\phi_\varepsilon(z) = (1/\varepsilon^2)\phi(z/\varepsilon)$, so that

$$\text{supp}(\phi_\varepsilon) \subset \{ z \in \mathbb{C} : |z| \leq \varepsilon \} \quad \text{and} \quad \int_{\mathbb{C}} \phi_\varepsilon \, dx dy = 1.$$

If $f : \mathbb{C} \to M$ comes from an element $f \in L^p_1(T^2, \mathbb{R}^N)$, we define $\phi_\varepsilon * f \in C^\infty(\mathbb{C}, \mathbb{R}^N)$ by

$$(\phi_\varepsilon * \gamma)(z) = \int_{\mathbb{C}} \phi_\varepsilon(z - w)\gamma(w) \, dw.$$  

It is immediately checked that $(\phi_\varepsilon * f)(z + \lambda) = (\phi_\varepsilon * f)(z)$ for $\lambda \in \Lambda$, so $(\phi_\varepsilon * f)$ can be identified with an element of $C^\infty(T^2, \mathbb{R}^N)$.

Thus we can define smoothing operators

$$S_\varepsilon : C^0(T^2, \mathbb{R}^N) \to C^k(T^2, \mathbb{R}^N),$$

$$S_\varepsilon : L^p_1(T^2, \mathbb{R}^N) \to C^k(T^2, \mathbb{R}^N)$$

by $S_\varepsilon(f) = \phi_\varepsilon * f$. The proof of the theorem for maps from $\Sigma = T^2$ now continues in exactly the same way as for maps from $S^1$.

Remark 1.5.4. It is interesting to consider Sobolev spaces of maps which do not lie in “Sobolev range”. Thus we can define

$$W^p_1(\Sigma, M) = \{ f \in L^p_1(\Sigma, M) : f(p) \in M \text{ for almost all } p \in \Sigma \} \quad \text{or} \quad$$

$$H^p_{1,S}(\Sigma, M) = (\text{closure of } C^\infty(\Sigma, M) \text{ in } L^p_1(\Sigma, M)), $$

for $p \leq \dim \Sigma$. Although $H^p_{1,S}(\Sigma, M) \subseteq W^p_1(\Sigma, M)$ the inclusion is often strict, and neither space is in general homotopically equivalent to the space $C^0(\Sigma, M)$ of continuous maps. These Sobolev spaces have been studied by Hang and Lin [HL03], among others, and applications to the theory of
harmonic maps are described in the review article of Brezis [Bre03]. When the dimension of $\Sigma$ is at least three, harmonic maps from $\Sigma$ to $M$ can have vastly more complicated singularities than geodesics and harmonic surfaces.

1.6. The $\alpha$- and $\omega$-Lemmas

In this section, we mention a few facts about maps between function spaces which are often useful. For a given choice of compact smooth manifold $\Sigma$, we have constructed a covariant functor from finite-dimensional smooth manifolds and smooth maps to infinite-dimensional smooth manifolds and smooth maps,

$$M \mapsto C^k(\Sigma, M), \quad g : M \to N \mapsto \omega_g : C^k(\Sigma, M) \to C^k(\Sigma, N),$$

where $\omega_g(f) = g \circ f$. If one studies the proof, one finds that in order to get a $C^r$ map $\omega_g$ one needs to assume that $g$ is $C^{k+r}$:

$\omega$-Lemma 1.6.1. If $M$ is a fixed smooth manifold, a $C^{k+r}$ map $g : M \to N$ induces a $C^r$ map

$$\omega_g : C^k(\Sigma, M) \to C^k(\Sigma, N)$$

for $k, r \in \{0\} \cup \mathbb{N}$.

One might hope that when $M$ is a fixed smooth manifold, there is a similar contravariant functor from compact manifolds $\Sigma$ to infinite-dimensional smooth manifolds $C^k(\Sigma, M)$ in which

$$f : \Sigma_1 \to \Sigma_2 \mapsto \alpha_f : C^k(\Sigma_2, M) \to C^k(\Sigma_1, M),$$

where $\alpha_f(g) = g \circ f$. Toward this end, one finds the following fact proven as Theorem 11.2 of [Abr63]:

$\alpha$-Lemma 1.6.2. If $M$ is a fixed smooth manifold, a $C^k$ map $f : \Sigma_1 \to \Sigma_2$ between smooth compact manifolds induces a $C^k$ map

$$\alpha_f : C^k(\Sigma_2, M) \to C^k(\Sigma_1, M), \quad \alpha_f(g) = g \circ f,$$

for $k \in \mathbb{N}$.

A combination of these lemmas (as presented in the survey by Eells [Eel66]) is often useful:

Theorem 1.6.3. If $S$, $\Sigma$ and $M$ are smooth finite-dimensional manifolds with $S$ and $\Sigma$ compact, then

$$\Phi : C^{k+r}(\Sigma, M) \times C^k(S, \Sigma) \to C^k(S, M), \quad \Phi(g, f) = g \circ f,$$

is $C^r$. 

By taking $S$ to be a point, we can derive the following:

**Corollary 1.6.4.** The evaluation map

\[(1.15) \quad ev : C^r(\Sigma, M) \times \Sigma \to M \quad \text{defined by} \quad ev(f, p) = f(p)\]

is $C^r$.

This corollary can also be proven directly as in [AR67] or [AMR88], page 99.

Versions of the $\alpha$- and $\omega$-Lemmas for Hilbert manifolds were later developed by Ebin [Ebi70], and they were used in [EM70] to study Euler’s equations for fluid flow. One can ask whether the group of $L^2_k$ diffeomorphisms might be made into a Banach Lie group. Let $C^1D(\Sigma)$ denote the topological group of $C^1$ diffeomorphisms of a compact manifold $\Sigma$, and let

\[D^k(\Sigma) = C^1D(\Sigma) \cap L^2_k(\Sigma, \Sigma), \quad \text{for} \quad k \in \{0\} \cup \mathbb{N},\]

which can be shown to be a topological group. The following theorem [Ebi70] can be thought of as a combination of the $\alpha$- and $\omega$-Lemmas:

**Theorem 1.6.5.** If $\Sigma$ and $M$ are smooth finite-dimensional manifolds with $\Sigma$ compact, then

\[\Phi : L^2_{k+r}(\Sigma, M) \times D^k(\Sigma) \to L^2_k(\Sigma, M), \quad \Phi(g, f) = g \circ f,\]

is $C^r$ for $k, r \in \{0\} \cup \mathbb{N}$.

The loss of derivatives in the statement of Theorem 1.6.3 prevents $D^k(\Sigma)$ from being a Banach Lie group under composition. On the other hand, suppose that $G$ is a finite-dimensional Lie subgroup of the group of $C^\infty$ automorphisms which acts on the compact surface $\Sigma$. Then Theorem 1.6.2 implies that the map

\[L^2_{k+1}(\Sigma, M) \times G \subseteq L^2_{k+1}(\Sigma, M) \times D^k(\Sigma) \to L^2_k(\Sigma, M)\]

is $C^1$. Thus one gets a $C^1$ action while losing one derivative in the space of functions.

**1.7. The tangent and cotangent bundles**

Many constructions from the theory of finite-dimensional manifolds are relatively easy to extend to infinite-dimensional Banach or Hilbert manifolds. These include tensors of various ranks, vector fields and differential equations, connections, Riemannian metrics on Hilbert manifolds, Finsler metrics
on Banach manifolds, differential forms and de Rham cohomology. Many
of these constructions are carried out in great detail in the Lang's book
[**Lan95**] on infinite-dimensional manifolds. We provide a brief summary
here.

We first extend familiar definitions of tangent and cotangent bundles
to the infinite-dimensional context. Let $\mathcal{M}$ be an infinite-dimensional smooth
manifold modeled on a Banach space $E$ with smooth atlas $\{(U_\alpha, \phi_\alpha) : \alpha \in \mathcal{A}\}$. Consider the collection of triples $(\alpha, p, v)$, where $\alpha \in \mathcal{A}$, $p \in U_\alpha$ and $v \in E$. On this collection of triples we define an equivalence relation $\sim$ by

$$(\alpha, p, v) \sim (\beta, q, w) \iff p = q \text{ and } w = D(\phi_\beta \circ \phi^{-1}_\alpha)(\phi_\alpha(p))v.$$ 

The set of equivalence classes is called the tangent bundle of $\mathcal{M}$ and is
denoted by $T\mathcal{M}$.

Let $[\alpha, p, v]$ denote the equivalence class of $(\alpha, p, v)$ and define

$$\pi : T\mathcal{M} \to \mathcal{M} \text{ by } \pi([\alpha, p, v]) = p.$$ 

Let $U_\alpha = \{[\alpha, p, v] ; p \in U_\alpha, v \in E\}$, and define

$$\tilde{\phi_\alpha} : U_\alpha \to E \oplus E \text{ by } \tilde{\phi_\alpha}([\alpha, p, v]) = (\phi_\alpha(p), v).$$ 

Then $\{([U_\alpha, \tilde{\phi_\alpha}) : \alpha \in \mathcal{A}\}$ is a smooth atlas on $T\mathcal{M}$ making $T\mathcal{M}$ into a
smooth manifold modeled on the Banach space $E \oplus E$. If $p \in \mathcal{M}$, we let $T_p\mathcal{M} = \pi^{-1}(p)$, the fiber of the tangent bundle over $p$, and call $T_p\mathcal{M}$ the
tangent space to $\mathcal{M}$ at $p$. Just as in the finite-dimensional case, elements of
$T_p\mathcal{M}$ are called tangent vectors.

If $\gamma : (a, b) \to \mathcal{M}$ is a smooth curve, $t \in (a, b)$ and $\gamma(t) \in U_\alpha$, we define

$$\gamma'(t) \in T_p\mathcal{M} \text{ by } \gamma'(t) = [\alpha, \gamma(t), D(\phi_\alpha \circ \gamma)(t) \cdot 1],$$

a tangent vector called the velocity vector to $\gamma$ at $t$.

If $F : \mathcal{M} \to \mathcal{N}$ is a smooth map between manifolds with atlases $\{(U_\alpha, \phi_\alpha) : \alpha \in \mathcal{A}\}$ and $\{(V_\beta, \psi_\beta) : \beta \in \mathcal{B}\}$ and $p \in \mathcal{M}$, we can define the differential

$(F)_p : T_p\mathcal{M} \to T_{F(p)}\mathcal{N}$ by

$$(F)_p([\alpha, p, v]) = [\beta, F(p), (D(\psi_\beta \circ F \circ \phi^{-1}_\alpha)(\phi_\alpha(p)))(v)],$$

where $p \in U_\alpha$ and $F(p) \in V_\beta$. Note that if $\gamma : (a, b) \to \mathcal{M}$ is a $C^1$ curve,

$$(F)_p(\gamma'(t)) = (F \circ \gamma)'(t), \quad \text{for } t \in (a, b).$$

The differentials fit together to form a map of tangent bundles $F_* : T\mathcal{M} \to T\mathcal{N}$.

In a very similar way, we can also describe the cotangent bundle of $\mathcal{M}$. We consider a similar collection of triples $(\alpha, p, v^*)$, where $\alpha \in \mathcal{A}$, $p \in U_\alpha$ and
$v^* \in E^*$, where $E^*$ is the Banach space dual to $E$, the space of continuous linear maps $T : E \to \mathbb{R}$. This time we choose the equivalence relation

$$(\alpha, p, v^*) \sim (\beta, q, w^*) \iff p = q \text{ and } v^* = [D(\phi_\beta \circ \phi_\alpha^{-1})(\phi_\alpha(p))] w^*,$$

where $(\cdot)^*$ denotes transpose map defined by

$$([D(\phi_\beta \circ \phi_\alpha^{-1})(\phi_\alpha(p))]^* w^*)(v) = w^*(D(\phi_\beta \circ \phi_\alpha^{-1})(\phi_\alpha(p))(v),$$

and we let $[\alpha, p, v^*]$ denote the equivalence class of $(\alpha, p, v^*)$. The cotangent bundle $T^*M$ is the set of these equivalence classes, and it is a smooth manifold modeled on the Banach space $E \oplus E^*$. Once again, we have a projection

$$\pi : TM \to M \text{ defined by } \pi([\alpha, p, v]) = p.$$ If $p \in M$, the fiber $T^*_p M$ of the cotangent bundle over $p$ is called the cotangent space to $M$ at $p$, and it can be regarded as the dual space to $T_p M$. As a morphism $F : M \to N$ induces a map in the opposite direction $(F^*)_p : T^*_p N \to T^*_p M$ by

$$(F^*)_p([\beta, F(p), w^*]) = [\alpha, p, (D(\psi_\beta \circ F \circ \phi_\alpha^{-1})(\phi_\alpha(p))^*(w^*)].$$

We can also define various tensor bundles, such as the $k$-th tensor power and the $k$-th exterior power of the cotangent bundle. For the first of these, we utilize the Banach space $L^k(E, \mathbb{R})$ of continuous maps

$$T : E \times E \times \cdots \times E \to \mathbb{R}$$

which are linear in each variable, the so-called space of continuous $k$-linear maps or continuous $k$-multilinear maps. For the second, we use its subspace of continuous alternating $k$-linear maps $L^k_a(E, \mathbb{R})$. An element $T \in L^k(E, \mathbb{R})$ is said to be alternating if

$$T(h_{\sigma(1)}, \ldots, h_{\sigma(k)}) = (\text{sgn}(\sigma))T(h_1, \ldots, h_k), \quad \text{for all } \sigma \in S_k,$$

where $S_k$ denotes the symmetric group on $k$ letters and sgn($\sigma$) denotes the sign of the permutation $\sigma \in S_k$. A continuous linear map $\Phi : E \to F$ between Banach spaces induces continuous linear maps

$$\Phi^* : L^k(F, \mathbb{R}) \to L^k(E, \mathbb{R}), \quad \Phi^* : L^k_a(F, \mathbb{R}) \to L^k_a(E, \mathbb{R})$$

by means of the formula

$$(\Phi^* T)(v_1, \ldots, v_k) = T(\Phi(v_1), \ldots, \Phi(v_k)).$$

To define the $k$-th tensor power of the cotangent bundle, we start with triples $(\alpha, p, T)$, where $\alpha \in A, p \in U_\alpha$ and $T \in L^k(E, \mathbb{R})$ and the equivalence relation

$$(\alpha, p, T_\alpha) \sim (\beta, q, T_\beta) \iff p = q \text{ and } T_\alpha = \Phi^* T_\beta,$$
1.7. Tangent and cotangent bundles

where $\Phi = D(\phi_\beta \circ \phi^{-1}_\alpha)(\phi_\alpha(p))$. The $k$-th tensor power of the cotangent bundle $\otimes^k T^* M$ is the set of equivalence classes. The $k$-th exterior power is defined the same way, except that $T$ is taken to lie in $L^k(E, \mathbb{R})$. The fibers $\otimes^k T^*_p M$ and $\Lambda^k T^*_p M$ are called the $k$-th tensor power and the exterior power of the cotangent space to $M$ at $p$.

In the case where $M = L^2_1(S^1, M)$, $M$ being oriented, the tangent bundle has another description, namely

$$TL^2_1(S^1, M) = L^2_1(S^1, TM).$$

To see this, recall how we constructed the atlas on $L^2_1(S^1, M)$. If $\gamma$ is a smooth element of $L^2_1(S^1, M)$, we let

$$U_{\gamma, \epsilon} = \{ \lambda \in L^2_1(S^1, M) : d_M(\lambda(t), \gamma(t)) < \epsilon \text{ for all } t \in S^1 \},$$

$$V_{\gamma, \epsilon} = \{ X \in L^2_1(S^1, \mathbb{R}^n) : (X(t), X(t)) < \epsilon \text{ for all } t \in S^1 \},$$

and define

$$\psi_{\gamma, \epsilon} : V_{\gamma, \epsilon} \to U_{\gamma, \epsilon} \quad \text{by} \quad (\psi_{\gamma, \epsilon}(X))(t) = \exp_{\gamma(t)}(X(t)).$$

For $\epsilon$ sufficiently small, $\psi_{\gamma, \epsilon}$ is a bijection with inverse $\phi_{\gamma, \epsilon} : U_{\gamma, \epsilon} \to V_{\gamma, \epsilon}$, and

$$\{(U_{\gamma, \epsilon}, \phi_{\gamma, \epsilon}) : \gamma \text{ is smooth and } \epsilon \text{ is sufficiently small}\}$$

is a smooth atlas for $L^2_1(S^1, M)$.

Now for each smooth $\gamma$, we can construct a lift $\tilde{\gamma} \in L^2_1(S^1, TM)$ by setting

$$\tilde{\gamma}(t) = 0_{\gamma(t)} \in T_{\gamma(t)} M,$$

and construct a corresponding chart on $L^2_1(S^1, TM)$. The remarkable fact is that we can choose the chart to be valid over all of

$$\tilde{U}_{\gamma, \epsilon} = \Omega_\pi^{-1}(U_{\gamma, \epsilon}) = \{ X \in L^2_1(S^1, TM) : \pi \circ X \in U_{\gamma, \epsilon} \}.$$

Indeed, we can set

$$\tilde{V}_{\gamma, \epsilon} = V_{\gamma, \epsilon} \times \{ L^2_1\text{-sections of } \gamma^* TM \}$$

and define $\tilde{\psi}_{\gamma, \epsilon} : \tilde{V}_{\gamma, \epsilon} \to \tilde{U}_{\gamma, \epsilon}$ by

$$\tilde{\psi}_{\gamma, \epsilon}(X, Y) = (\exp_{\gamma(t)} X(t), (d(\exp_{\gamma(t)})) X(t) Y(t)).$$

Finally, define $\tilde{\phi}_{\gamma, \epsilon} : \tilde{U}_{\gamma, \epsilon} \to \tilde{V}_{\gamma, \epsilon}$ by $\tilde{\phi}_{\gamma, \epsilon} = \tilde{\psi}_{\gamma, \epsilon}^{-1}$. Then

$$(\tilde{\phi}_{\gamma, \epsilon} \circ \tilde{\phi}_{\gamma, \epsilon}^{-1})(X, Y) = ((\phi_{\gamma_1, \epsilon} \circ \phi_{\gamma_2, \epsilon}^{-1})(X), D(\phi_{\gamma_1, \epsilon} \circ \phi_{\gamma_2, \epsilon}^{-1})(X)(Y)).$$

Thus the charts transform exactly the way they should for the tangent bundle.

In a quite similar fashion, we can show that if $\Sigma$ is a compact Riemann surface and $p > 2$,

$$TL^p_1(\Sigma, M) = L^p_1(\Sigma, TM).$$
It is important to observe that just as the imbedding \( i : M \to \mathbb{R}^N \) induces an imbedding \( \omega_i : L^2_1(S^1, M) \to L^2_1(S^1, \mathbb{R}^N) \), so the imbedding \( i : TM \to T\mathbb{R}^N = \mathbb{R}^{2N} \) induces an imbedding
\[
\omega_i : TL^2_1(S^1, TM) = L^2_1(S^1, TM) \longrightarrow L^2_1(S^1, T\mathbb{R}^N) = L^2_1(S^1, \mathbb{R}^{2N}),
\]
allowing us to realize \( TL^2_1(S^1, TM) \) as a subspace of a Banach space. Similarly, \( TL^1_p(\Sigma, M) \) can be regarded as a subspace of a Banach space.

Note that the Hilbert space inner product allows us to identify the model space \( L^2_1(S^1, \mathbb{R}^n) \) for \( \mathcal{M} = L^2_1(S^1, M) \) with its dual. Using this fact, it is not difficult to verify that
\[
T^*L^2_1(S^1, M) = L^2_1(S^1, T^*M), \quad \otimes^k T^*L^2_1(S^1, M) = L^2_1(S^1, \otimes^k T^*M),
\]
and
\[
\Lambda^k T^*L^2_1(S^1, M) = L^2_1(S^1, \Lambda^k T^*M).
\]

It is actually the “sections” of the tangent bundle and the exterior powers of the cotangent bundle that will be of most importance for us; these are called vector fields and differential forms, respectively.

**Definition.** A smooth vector field on \( \mathcal{M} \) is a smooth map
\[
\mathcal{X} : \mathcal{M} \to TM \quad \text{such that} \quad \pi \circ \mathcal{X} = \text{id}_\mathcal{M}.
\]

Note that if \( \mathcal{X} : \mathcal{M} \to TM \) is a smooth vector field and \( f : \mathcal{M} \to \mathbb{R} \) is a smooth function, we can define the vector field \( f\mathcal{X} : \mathcal{M} \to TM \) by \( (f\mathcal{X})(p) = f(p)\mathcal{X}(p) \). This makes the space of smooth vector fields into a module over the ring of smooth real-valued functions.

**Example.** If \( M \) is a finite-dimensional Riemannian manifold and \( X : M \to TM \) is a smooth vector field on \( M \), then
\[
\omega_X : L^2_1(S^1, M) \longrightarrow L^2_1(S^1, TM) = TL^2_1(S^1, M)
\]
and
\[
\omega_X : L^p_1(\Sigma, M) \longrightarrow L^p_1(\Sigma, TM) = TL^p_1(\Sigma, M)
\]
are smooth vector fields on \( L^2_1(S^1, M) \) and \( L^p_1(\Sigma, M) \) when \( p > 2 \).

**1.8. Differential forms**

For calculations on smooth manifolds, differential forms are often more convenient to use than general tensor fields.

**Definition.** A smooth covariant tensor field of rank \( k \) on \( \mathcal{M} \) is a smooth map
\[
\phi : \mathcal{M} \longrightarrow \otimes^k T^*\mathcal{M} \quad \text{such that} \quad \pi \circ \phi = \text{id}_\mathcal{M}.
\]
A *smooth differential form* of degree $k$ on $\mathcal{M}$ (or a smooth $k$-form) is a smooth map

$$\phi : \mathcal{M} \rightarrow \Lambda^k T^* \mathcal{M}$$

such that $\pi \circ \phi = \text{id}_\mathcal{M}$.

We let $\Omega^k(\mathcal{M})$ denote the space of smooth differential forms of degree $k$ on $\mathcal{M}$.

As in the case of vector fields, we can multiply covariant tensor fields or differential forms by functions.

An important example of differential one-form occurs when $f : \mathcal{M} \rightarrow \mathbb{R}$ is a smooth function. Then for the coordinate chart $(U_{\alpha}, \phi_{\alpha})$, we have

$$D(f \circ \phi_{\alpha}^{-1}) : U_{\alpha} \rightarrow L(E, \mathbb{R}) = E^*,$$

where $E$ is the model space of $\mathcal{M}$. The *differential* of $f$ is the smooth one-form $df$ such that

$$df(p) = [\alpha, p, D(f \circ \phi_{\alpha}^{-1})(\phi_{\alpha}(p))], \quad \text{for } p \in U_{\alpha}.$$

It is readily verified that the local representatives transform as they should under change of coordinates. Moreover, it is easily checked that the differentials of functions at a point $p$ generate the cotangent space. The Leibniz rule for differentiation implies that $d(fg) = gdf + fdg$.

**Definition.** A point $p \in \mathcal{M}$ is a *critical point* for the real-valued function $f : \mathcal{M} \rightarrow \mathbb{R}$ if $df(p) = 0$.

An important special case of this construction is the real-valued function

$$J : L^2_1(S^1, M) \rightarrow M, \quad J(\gamma) = J(\gamma) = \frac{1}{2} \int_{S^1} |\gamma'(t)|^2 dt,$$

when $M$ is a Riemannian manifold. In this case, regularity theory will show that a critical point in this case is actually a $C^\infty$ map, hence a smooth closed geodesic in $M$.

Using the notion of differential of a function, we can define directional derivatives:

**Definition.** The *directional derivative* of a function $f$ in the direction of a vector field $\mathcal{X}$, or the *Lie derivative* of $f$ in the direction of $\mathcal{X}$ is the function $\mathcal{X}(f)$ defined by

$$\mathcal{X}(f)(p) = df(p)(\mathcal{X}(p)),$$

the right-hand side being the dual pairing between cotangent and tangent spaces.
Lemma 1.8.1. Let $\mathcal{X}$ and $\mathcal{Y}$ be smooth vector fields on the Banach manifold $\mathcal{M}$. Then there is a unique vector field $[\mathcal{X}, \mathcal{Y}]$ on $\mathcal{M}$ which satisfies the equation

$$([\mathcal{X}, \mathcal{Y}](f))(p) = (\mathcal{X}\mathcal{Y}(f))(p) - (\mathcal{Y}\mathcal{X}(f))(p).$$

Sketch of proof. It suffices to show that the above formula is equivalent to an expression for $[\mathcal{X}, \mathcal{Y}]$ in terms of a local coordinate chart $(U_\alpha, \phi_\alpha)$. Suppose that $\bar{\mathcal{X}}, \bar{\mathcal{Y}} : U_\alpha \to E$ are defined by $\mathcal{X}(p) = [p, \alpha, \bar{\mathcal{X}}(p)]$, $\mathcal{Y}(p) = [p, \alpha, \bar{\mathcal{Y}}(p)]$. Using the chain rule, one can check that the vector field $[\mathcal{X}, \mathcal{Y}]$ is then defined by

$$[\mathcal{X}, \mathcal{Y}](p) = [p, \alpha, D\bar{\mathcal{Y}}(p)\bar{\mathcal{X}}(p) - D\bar{\mathcal{X}}(p)\bar{\mathcal{Y}}(p)].$$

QED

The vector field $[\mathcal{X}, \mathcal{Y}]$ is known as the Lie bracket of $\mathcal{X}$ and $\mathcal{Y}$; it is easily verified that it satisfies the identity:

$$[f\mathcal{X}, g\mathcal{Y}] = fg[\mathcal{X}, \mathcal{Y}] + f\mathcal{X}(g)\mathcal{Y} - g\mathcal{Y}(f)\mathcal{X}.$$ 

If $\mathcal{X}_1, \ldots, \mathcal{X}_k$ are smooth vector fields on $\mathcal{M}$ and $\phi$ is a smooth $k$-form on $\mathcal{M}$, then the smooth function

$$\phi(\mathcal{X}_1, \ldots, \mathcal{X}_k) : \mathcal{M} \to \mathbb{R}$$

is defined fiberwise via the continuous $(k + 1)$-linear map

$$L^k_a(E, \mathbb{R}) \times E \times \cdots \times E \to \mathbb{R},$$

where $E$ is the model space for $\mathcal{M}$.

Definition. If $\phi \in \Omega^k(\mathcal{M})$ and $\omega \in \Omega^l(\mathcal{M})$, the wedge product of $\phi$ and $\omega$ is the $(k + l)$-form $\phi \wedge \omega$ defined by

$$(\phi \wedge \omega)(\mathcal{X}_1, \ldots, \mathcal{X}_{k+l}) = \frac{k!!}{(k + l)!!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) \phi(\mathcal{X}_{\sigma(1)}, \ldots, \mathcal{X}_{\sigma(k)})\omega(\mathcal{X}_{\sigma(k+1)}, \ldots, \mathcal{X}_{\sigma(k+l)}).$$

Here $S_{k+l}$ is the symmetric group on $k + l$ letters and sgn$(\sigma)$ is the sign of the permutation $\sigma \in S_{k+l}$.

We remind the reader that some authors prefer to define the wedge product using the factor

$$\frac{1}{(k + l)!!} \text{ instead of } \frac{k!!}{(k + l)!!}.$$
With either convention, the wedge product is bilinear and associative, just as in the case of finite-dimensional manifolds, but not commutative. If $\phi$ is a $k$-form and $\omega$ an $l$-form, then

$$\phi \wedge \omega = (-1)^{kl} \omega \wedge \phi.$$ 

We next note that differential forms are “functorial”. If $F : \mathcal{M} \to \mathcal{N}$ is a smooth map and $\phi$ is a smooth differential form of degree $k$ on $\mathcal{N}$, we can define a differential form $F^* \phi$ on $\mathcal{M}$ by

$$[F^* \phi](p) = F^*_p(\phi(F(p)));$$

it follows from (1.16) that $F^* \phi$ is smooth. Moreover, the wedge product is natural under smooth maps: if $F : \mathcal{M} \to \mathcal{N}$ is a smooth map and $\phi$ and $\omega$ are smooth differential forms on $\mathcal{N}$, then

$$F^*(\phi \wedge \omega) = F^* \phi \wedge F^* \omega.$$ 

**Definition.** If $\phi$ is a smooth $k$-form on $\mathcal{M}$ and $\mathcal{X}$ is a smooth vector field, the interior product $\iota_{\mathcal{X}} \phi$ is the smooth $(k - 1)$-form on $\mathcal{M}$ defined by the formula

$$\iota_{\mathcal{X}} \phi(\mathcal{X}_2, \ldots, \mathcal{X}_k) = \phi(\mathcal{X}, \mathcal{X}_2, \ldots, \mathcal{X}_k),$$

whenever $\mathcal{X}_1, \ldots, \mathcal{X}_k$ are smooth vector fields on $\mathcal{M}$. (It is readily checked that there is a unique such differential form.) It is easily checked that

$$\iota_{\mathcal{X}}(\phi \wedge \psi) = (\iota_{\mathcal{X}} \phi) \wedge \psi + (-1)^{\deg(\phi)} \phi \wedge \iota_{\mathcal{X}} \psi.$$ 

Finally, we claim that there is an exterior derivative $d$, a the collection of $\mathbb{R}$-linear maps from $k$-forms to $(k + 1)$-forms which satisfy the following axioms, familiar from finite-dimensional differential topology:

1. If $\omega$ is a $k$-form, the value $d\omega(p)$ depends only on $\omega$ and its derivatives at $p$.
2. If $f$ is a smooth real-valued function regarded as a differential 0-form, $d(f)$ is the differential of $f$ defined before.
3. $d \circ d = 0$.
4. If $\omega$ is a $k$-form and $\phi$ is an $l$-form, then

$$d(\omega \wedge \phi) = (d\omega) \wedge \phi + (-1)^k \omega \wedge (d\phi).$$

5. If $F : \mathcal{N} \to \mathcal{M}$ is a smooth map, $F^* \circ d = d \circ F^*$ on differential forms.
Just as in the finite-dimensional case, one can prove the following:

**Theorem 1.8.2.** There is a unique of linear maps of real vector spaces,

\[ d : \Omega^k(M) \rightarrow \Omega^{k+1}(M), \]

which satisfy the five axioms above. Moreover, these linear maps satisfy the explicit formula

\[
(1.18) \quad d\omega(X_0, \ldots, X_k) = \sum (-1)^i X_i \left( \omega(X_0, \ldots, \hat{X}_i, \ldots, X_k) \right) + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_k),
\]

where the hats denote elements which are left out.

We sketch the proof under the assumption that the corresponding theorem for finite-dimensional manifolds has been established. Using the fifth axiom, we can reduce the proof of uniqueness to the case where the manifold is an open subset of the model space. Since (1.18) is linear over functions, it suffices to establish the formula in the case where \( X_0, \ldots, X_k \) are constant in terms of the local chart, in which case all brackets \([X_i, X_j]\) vanish. Thus it suffices to prove uniqueness when the vector fields \( X_0, \ldots, X_k \) are tangent to a \((k+1)\)-dimensional affine subspace of the model space, and in this case (1.18) follows from the corresponding formula on this \((k+1)\)-dimensional subspace, a finite-dimensional submanifold.

To prove the local existence, one merely defines the exterior derivatives by the formula (1.18) and check that they satisfy all of the axioms. It is simplest to verify the first four axioms by showing that if any of these axioms were to fail, it would have to fail already on finite-dimensional manifolds.

Finally, one notes that if operators satisfying the five axioms are defined and unique for any open subset of the model space, they are defined and unique for any set \( U_\alpha \) in a smooth atlas \( \{U_\alpha : \alpha \in A\} \) for \( \mathcal{M} \). These operators must agree on overlaps \( U_\alpha \cap U_\beta \), for \( \alpha, \beta \in A \), and hence must fit together to form well-defined operators on \( \mathcal{M} \), and the resulting operators must be unique because their restrictions to each \( U_\alpha \) are unique.

Differential forms of degree \( k \) should be thought of as integrands for integration over compact oriented \( k \)-dimensional submanifolds. From this point of view, it is evident that they should be determined by their restrictions to finite-dimensional manifolds, where the above axioms for exterior derivative are familiar.
1.9. Riemannian and Finsler metrics

To construct critical points of functions such as the action or energy, we need to develop the “method of steepest descent” within the context of infinite-dimensional manifolds. And to find the direction of steepest descent, we are led to seek a notion of gradient, which in the finite-dimensional context depends upon a Riemannian metric. Thus it is important to consider how to extend the notion of Riemannian metric to infinite-dimensional manifolds. It is to be expected that a somewhat stronger theory is possible for Hilbert manifolds than for Banach manifolds. Indeed, we will see that there is no fully satisfactory notion of Riemannian metric or gradient on Banach manifolds. We will need to use the weaker notions of Finsler metrics and pseudogradients.

Suppose therefore that $\mathcal{M}$ is a Hilbert manifold modeled on the Hilbert space $E$. If $(U_\alpha, \phi_\alpha)$ is a smooth chart on $\mathcal{M}$, the Hilbert space inner product $\langle \cdot, \cdot \rangle$ pulls back via $\phi_\alpha$ to an inner product on $T_p\mathcal{M}$: we define

$$\langle [\alpha, p, v], [\alpha, p, w] \rangle_\alpha = (v, w).$$

**Definition.** A Riemannian metric on a Hilbert manifold $\mathcal{M}$ is a function which assigns to each $p \in \mathcal{M}$ an inner product

$$\langle \cdot, \cdot \rangle_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$$

such that:

1. For each $p \in \mathcal{M}$, there is some constant $c_p > 0$ such that

$$\frac{1}{c_p} \langle [\alpha, p, v], [\alpha, p, v] \rangle_\alpha < \langle [\alpha, p, v], [\alpha, p, v] \rangle_p < c_p \langle [\alpha, p, v], [\alpha, p, v] \rangle_\alpha,$$

   for all $[\alpha, p, v] \in T_p\mathcal{M}$ with $v \neq 0$. (Thus the topology induced by the Riemannian metric on $T_p\mathcal{M}$ agrees with the model space topology.)

2. $\langle \cdot, \cdot \rangle_p$ varies smoothly with $p$; in other words, $p \mapsto \langle \cdot, \cdot \rangle_p$ is a smooth covariant tensor field of rank two.

**Example 1.9.1.** Suppose that we have an isometric imbedding of the smooth compact Riemannian manifold $M$ into $\mathbb{R}^N$. Then the Hilbert manifold $L^2_1(S^1, M)$ can be given the Riemannian metric $\langle \cdot, \cdot \rangle$ defined as follows: if $X, Y \in T_\gamma L^2_1(S^1, M)$, we can regard $X$ and $Y$ as maps $X, Y : S^1 \to \mathbb{R}^N$ such that $X(t) \in T_{\gamma(t)}M$ for each $t \in S^1$. We set

$$\langle X, Y \rangle_\gamma = \int_{S^1} [X(t) \cdot Y'(t) + X'(t) \cdot Y(t)] dt.$$
This can be regarded as the pullback of the “flat” Riemannian metric on the Hilbert space $L^2_\text{f}(S^1, \mathbb{R}^N)$, and it is smooth because the pullback of a smooth covariant tensor field via a smooth map is smooth.

**Definition.** If $\mathcal{M}$ is a Hilbert manifold with Riemannian metric $p \mapsto \langle \cdot, \cdot \rangle_p$, the gradient of a $C^1$ function $f : \mathcal{M} \to \mathbb{R}$ is the vector field $\text{grad}(f)$ defined by

$$\langle \text{grad}(f)(p), v \rangle_p = df_p(v),$$

for all $v \in T_p \mathcal{M}$.

The idea behind the method of steepest descent is to follow flowlines for the vector field $-\text{grad}(f)$; in favorable cases, these flowlines will converge to a critical point for $f$.

**Example 1.9.2.** Once again suppose that the smooth compact Riemannian manifold $M$ is isometrically imbedded in $\mathbb{R}^N$. If $\Sigma$ is a compact Riemann surface with a constant curvature metric of total area one in its equivalence class, the Hilbert manifold $L^2_\text{f}(\Sigma, M)$ can be given a Riemannian metric $\langle \cdot, \cdot \rangle$ by the following formula: If $X, Y \in T_f L^2_\text{f}(\Sigma, M)$, we can regard $X$ and $Y$ as maps $X, Y : \Sigma \to \mathbb{R}^N$ such that $X(p) \in T_{f(p)} M$ for each $p \in \Sigma$. We set

$$\langle X, Y \rangle_f = \int_{\Sigma} [X \cdot Y + DX \cdot DY + D^2X \cdot D^2Y] dt,$$

where $D$ and $D^2$ represent the first and second covariant differentials computed with respect to the Riemannian metric on $\Sigma$. This is the pullback of a “flat” Riemannian metric on the Hilbert space $L^2_\text{f}(\Sigma, \mathbb{R}^N)$. If $\Sigma$ is $S^2$ or $T^2$, the group of isometries of the metric on $\Sigma$ acts as isometries of this Riemannian metric. Of course, we can do a similar construction for $L^2_k(\Sigma, M)$, when $k \geq 3$.

In the case of Banach manifolds, the metrics best suited to our applications are not Riemannian, but Finsler. If $(U_\alpha, \phi_\alpha)$ is a smooth chart on a Banach manifold $\mathcal{M}$, the Banach space norm $\| \cdot \|$ pulls back via $\phi_\alpha$ to $T_p \mathcal{M}$: we can define

$$\| \cdot \|_\alpha : T_p \mathcal{M} \to \mathbb{R} \quad \text{by} \quad \| (\alpha, p, v) \|_\alpha = \| v \|.$$

**Definition.** A Finsler metric on a Banach manifold $\mathcal{M}$ is a function which assigns to each $p \in \mathcal{M}$ a norm

$$\| \cdot \|_p : T_p \mathcal{M} \longrightarrow \mathbb{R}$$

such that:

1. For each $p \in \mathcal{M}$, there is some constant $c_p > 0$ such that

$$\frac{1}{c_p} \| (\alpha, p, v) \|_\alpha < \| (\alpha, p, v) \|_p < c_p \| (\alpha, p, v) \|_\alpha,$$
for all \([\alpha, p, v] \in T_p \mathcal{M}\) with \(v \neq 0\). (Thus the Finsler norm on \(T_p \mathcal{M}\) is equivalent to the Banach space norm for the model space.)

(2) \(\| \cdot \|_p\) varies continuously with \(p\): Given \(p \in \mathcal{M}\) and \(\varepsilon > 0\), there exists an open neighborhood \(U\) of \(p\) such that \(q \in U\) implies

\[
(1 - \varepsilon)\| [\alpha, q, v] \|_q < \| [\alpha, p, v] \|_p < (1 + \varepsilon)\| [\alpha, q, v] \|_q,
\]

Simplification of notation: We will henceforth write \(v\) instead of \([\alpha, p, v]\) for a typical element of \(T_p \mathcal{M}\).

Note that any Riemannian metric on a Hilbert manifold determines a Finsler metric: simply define

\[
\| \cdot \|_p : T_p \mathcal{M} \to \mathbb{R} \quad \text{by} \quad \| v \|_p = \sqrt{\langle v, v \rangle_p}.
\]

Even in this special case, however, the norm \(\| \cdot \|_p\) is only continuous, not smooth as a function on \(T \mathcal{M}\).

The Riemannian metric on a Hilbert manifold gives a norm-preserving vector bundle isomorphism from \(T_p \mathcal{M}\) to \(T^*_p \mathcal{M}\). We do not have such an isomorphism in the case of a Finsler metric on a Banach manifold, but the norm \(\| \cdot \|_p\) on \(T_p \mathcal{M}\) induces a dual norm (which we also denote by \(\| \cdot \|_p\) for simplicity) on \(T^*_p \mathcal{M}\):

\[
\| \phi \|_p = \sup \{ |\phi(v)| : v \in T_p \mathcal{M} \text{ and } \| v \|_p = 1 \}.
\]

**Example 1.9.3.** Once again suppose that the smooth compact Riemannian manifold \(\mathcal{M}\) is isometrically imbedded in \(\mathbb{R}^N\). Given a compact Riemann surface \(\Sigma\) and a real number \(p > 2\), the Banach manifold \(L^p_1(\Sigma, \mathcal{M})\) can be given a Finsler metric as follows: If \(X \in T_f(L^p_1(\Sigma, \mathcal{M}))\), we can regard \(X\) as a map \(X : \Sigma \to \mathbb{R}^N\) such that \(X(p) \in T_f(p)\mathcal{M}\) for each \(p \in \Sigma\). We then let \(\| X \|_f\) be the \(L^p_1\)-norm of \(X\) as a mapping into Euclidean space. Alternatively, this Finsler metric \(f \mapsto \| \cdot \|_f\) can be regarded as the pullback of the “flat” Finsler metric on the Banach space \(L^p_1(\Sigma, \mathbb{R}^N)\).

If \(\mathcal{M}\) is a connected Banach manifold with reflexive model space \(E\) and Finsler metric \(\| \cdot \|\) and \(\gamma : [0, 1] \to \mathcal{M}\) is a \(C^1\) curve, we can define its length \(L(\gamma)\) by

\[
L(\gamma) = \int_0^1 \| \gamma'(t) \| dt,
\]

where integration along a path can be defined as follows. First we use (1.4) to define path integrals in the Banach space \(E\), then we transport this to paths in the domain of a single coordinate chart and show it is independent of such a chart, then we extend to paths in the Banach manifold itself. We
1. Infinite-dimensional Manifolds

can then define a distance function \( d : \mathcal{M} \times \mathcal{M} \to \mathbb{R} \) by

\[
(1.19) \quad d(p, q) = \inf \{ L(\gamma) : \gamma : [0, 1] \to \mathcal{M} \text{ is } C^1, \, \gamma(0) = p, \gamma(1) = q \}.
\]

**Proposition 1.9.4.** Given a Finsler metric on a regular Banach manifold, the distance function \( d \) defined above defines a metric in the sense of metric spaces, and the metric topology agrees with the manifold topology.

For the general proof, we refer to [Pal70] (see the appendix to §2).

We merely describe a proof for our Examples 1.9.1, 1.9.2 and 1.9.3. It is quite easily verified that the distance function \( d \) satisfies

\[
d(p, q) = d(q, p), \quad d(p, r) \leq d(p, q) + d(q, r), \quad \text{and} \quad d(p, p) = 0.
\]

That only leaves the property \( d(p, q) = 0 \Rightarrow p = q \).

**Proposition 1.9.5.** In each of our examples, \( L^2_1(S^1, M) \), \( L^2_k(\Sigma, M) \) for \( k \geq 2 \) and \( L^p(\Sigma, M) \) with \( p > 2 \), \( d \) is a metric and the metric topology agrees with the manifold topology. Moreover, \( L^2_1(S^1, M) \), \( L^2_k(\Sigma, M) \) for \( k \geq 2 \) and \( L^p(\Sigma, M) \) are complete as metric spaces.

**Proof.** Let us consider \( L^p_1(\Sigma, M) \). If \( f, g \in L^p_1(\Sigma, M) \) and \( d(f, g) = 0 \), then there exist arbitrarily short paths connecting \( f \) and \( g \) in \( L^p_1(\Sigma, M) \). But a path connecting \( f \) and \( g \) in \( L^p_1(\Sigma, M) \) also connects \( f \) and \( g \) in the ambient Banach space \( E = L^p_1(\Sigma, \mathbb{R}^N) \), so there are arbitrarily short curves connecting \( f \) and \( g \) in the ambient Banach space. However, by a straightforward modification of the finite-dimensional argument, it is easily verified that if \( \gamma : [0, 1] \to E \) is a \( C^1 \)-map into a Banach space, then

\[
\int_0^1 \| \gamma'(t) \| dt \geq \| \gamma(0) - \gamma(1) \|,
\]

so two distinct points in \( E \) cannot be joined by curves of arbitrarily small length.

Since the map \( \omega_i : L^p_1(\Sigma, M) \to L^p_1(\Sigma, \mathbb{R}^N) \) induced by the inclusion \( i : M \to \mathbb{R}^N \) is an imbedding, it is now easy to verify that the metric topology agrees with the manifold topology. Finally, since \( \omega_i : L^p_1(\Sigma, M) \to L^p_1(\Sigma, \mathbb{R}^N) \) is distance-decreasing, a Cauchy sequence \( \{ f_i \} \) in \( L^p_1(\Sigma, M) \) is also a Cauchy sequence in \( L^p_1(\Sigma, \mathbb{R}^N) \), and must therefore converge. Since \( L^p_1(\Sigma, M) \) is a closed subset of \( L^p_1(\Sigma, \mathbb{R}^N) \), we see that \( L^p_1(\Sigma, M) \) must be complete as a metric space.

The other cases are treated in the same way. QED
1.10. Vector fields and ODEs

It is well-known that the global qualitative theory of systems of ordinary differential equations is best formulated within the language of vector fields on finite-dimensional manifolds. This theory, including the fundamental existence and uniqueness theorem for systems of ordinary differential equations, can be extended to infinite-dimensional manifolds with almost no change in the proofs. A detailed exposition of this extension is presented in Chapter IV of [Lan95].

Definition. A \( C^1 \) curve \( \gamma : (a, b) \to M \) is called an integral curve for the vector field \( \mathcal{X} \) if

\[
\mathcal{X}(\gamma(t)) = \gamma'(t), \quad \text{for } t \in (a, b),
\]

where \( \gamma'(t) \) is the velocity vector to \( \gamma \) at \( t \).

Just as in the finite-dimensional case, a fundamental Existence and Uniqueness Theorem for Ordinary Differential Equations states that given a smooth vector field \( \mathcal{X} \) on \( M \), there is a unique integral curve for \( \mathcal{X} \) which passes through any point of \( M \):

Theorem 1.10.1. Suppose that \( \mathcal{X} \) is a \( C^2 \) vector field on \( M \) and \( p \in M \). Then there is an open neighborhood \( U \) of \( p \), an \( \epsilon > 0 \) and a \( C^2 \) map

\[
\phi : (-\epsilon, \epsilon) \times U \longrightarrow M
\]

such that if \( \phi_t(q) = \phi(t, q) \) for \( t \in (-\epsilon, \epsilon) \) and \( q \in U \), then

1. each curve \( t \mapsto \phi_t(q) \) is an integral curve for \( \mathcal{X} \),
2. any integral curve for \( \mathcal{X} \) which passes through \( U \) is of the form \( t \mapsto \phi_t(q) \) for some \( q \in U \),
3. \( \phi_0 \) is the inclusion \( U \subset M \), and
4. \( \phi_t \circ \phi_s = \phi_{t+s} \), whenever, both sides are defined.

We will call \( (-\epsilon, \epsilon) \times U \) a local flow box for \( \mathcal{X} \).

Idea of proof (following Chapter IV of Lang [Lan95]). The proof is based upon the Contraction Lemma, just like the proof of the Inverse Function Theorem.

We can replace the differential equation (1.20) by its local coordinate representation, and consider the initial value problem

\[
\gamma'(t) = f(\gamma(t)), \quad \gamma(0) = q,
\]
for \( \gamma : (a, b) \to V \) and \( f : V \to E \), where \( V \) is a suitable open subset of a Banach space \( E \). We can assume that
\[
\|f\| \leq K \quad \text{and} \quad \|Df\| \leq L
\]
on \( V \). Integrating both sides of (1.21) yields the equivalent integral equation
\[(1.22) \quad \gamma(t) = q + \int_0^t f(\gamma(u))du.\]

We can assume that the closed ball \( B_{\delta}(p) \) of radius \( \delta \) about \( p \) is contained in \( V \) and suppose that \( q \in B_{\delta}(p) \), the open ball of radius \( \delta \) about \( p \). Let \( I = [-\epsilon, \epsilon] \), where \( \epsilon > 0 \) will be chosen later, and let
\[
X = \left\{ \gamma : I \to U : \gamma \text{ is continuous}, \gamma(0) = q \text{ and } \gamma(I) \subset B_{2\delta}(p) \right\}.
\]

We can make \( X \) into a complete metric space by defining the distance function \( d \) by
\[
d(\gamma_1, \gamma_2) = \sup \{\|\gamma_1(t) - \gamma_2(t)\| : t \in I\}.
\]
If \( \gamma \in X \), we set
\[
T(\gamma)(t) = q + \int_0^t f(\gamma(u))du.
\]
We choose \( \epsilon \) so that \( \epsilon < \delta/K \), and hence
\[
\|T(\gamma)(t) - q\| \leq \epsilon K \leq \delta,
\]
so \( T(\gamma) \in X \). Finally, we note that
\[
d(T(\gamma_1), T(\gamma_2)) = \sup \{\|T(\gamma_1)(t) - T(\gamma_2)(t)\| : t \in I\}
\leq \epsilon \sup \{\|f(\gamma_1(t)) - f(\gamma_2(t))\| : t \in I\}
\leq \epsilon L \sup \{\|\gamma_1(t) - \gamma_2(t)\| : t \in I\} = \epsilon L d(\gamma_1, \gamma_2).
\]
Thus by choosing \( \epsilon \) so that \( \epsilon L < 1 \), we can ensure that \( T : X \to X \) will be a contraction. Then, by the Contraction Lemma, \( T \) has a unique fixed point \( \gamma_q \in X \), which must be a solution to the integral equation (1.22). This fixed point \( \gamma_q \) is \( C^1 \) and is the unique solution to the initial value problem (1.21).

Thus we can define
\[
\phi : (-\epsilon, \epsilon) \times B_{\delta}(p) \to V \quad \text{by} \quad \phi(t, q) = \gamma_q(t).
\]
It is relatively easy to check that properties (2), (3) and (4) of the theorem hold and that \( \phi \) is continuous. It takes a little more work to check that the map \( \phi \) is \( C^2 \), and for that we refer the reader to the excellent presentation in [Lan95]. QED
Once we have the Existence and Uniqueness Theorem, we can piece together the locally defined maps to form a map

$$\phi : V \rightarrow M,$$

where $V$ is an open neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$. which we call the local flow for $X$. We say that the maps $\{\phi_t\}$ defined by $\phi_t(q) = \phi(t, q)$ form the one-parameter group of local diffeomorphisms of $M$ which corresponds to the vector field $X$.

### 1.11. Condition C

We want to apply the existence and uniqueness theorem from the preceding section to find critical points of a $C^2$ real-valued map $f : M \rightarrow [0, \infty)$ via the method of steepest descent, where $M$ is an infinite-dimensional manifold. In order to get this method to work, we need to assume that the function $f$ assumes a “compactness condition” introduced by Palais and Smale [PS64]:

**Definition.** Suppose that $M$ is a Banach manifold with a complete Finsler metric. (For example, $M$ might be a Hilbert manifold with a complete Riemannian metric.) Then a $C^2$ function $f : M \rightarrow [0, \infty)$ is said to satisfy condition $C$ if whenever $\{p_i\}$ is a sequence in $M$ such that

1. $f(p_i)$ is bounded and
2. $\|df(p_i)\|$ is not bounded away from zero,

then $\{p_i\}$ possesses a subsequence which converges to a critical point for $f$.

Note that if $M$ is finite-dimensional, $f : M \rightarrow [0, \infty)$ satisfies Condition C if and only if it is proper. However, Condition C also makes sense and is easy to utilize when $M$ is an infinite-dimensional Hilbert manifold:

**Theorem 1.11.1.** Suppose that $M$ is a Hilbert manifold with a complete Riemannian metric $\langle \cdot, \cdot \rangle$. If $f : M \rightarrow [0, \infty)$ is a $C^2$ function which satisfies condition $C$, $X = \text{grad}(f)$ and $\{\phi_t\}$ is the local one-parameter group of diffeomorphisms corresponding to $-X$, then

1. for each $p \in M$, $\phi_t(p)$ is defined for all $t \geq 0$, and
2. there is a sequence $t_i \rightarrow \infty$ such that $\{\phi_{t_i}(p)\}$ converges to a critical point for $f$.

The proof will be based upon a collection of inequalities. It follows from the chain rule, that if $g : M \rightarrow \mathbb{R}$ is any smooth function,

$$dg(\gamma(t))(X(\gamma(t))) = dg(\gamma(t))(\gamma'(t)) = \frac{d}{dt}(g \circ \gamma)(t).$$
We use this fact, together with the fundamental theorem of calculus, to show that if $0 < t_1 < t_2$,

\[(1.24) \quad f(\phi_{t_1}(p)) - f(\phi_{t_2}(p)) = -\int_{t_1}^{t_2} \frac{d}{dt} f(\phi_t(p)) dt\]

\[= \int_{t_1}^{t_2} df(\phi_t(p)) (X(\phi_t(p))) dt = \int_{t_1}^{t_2} (\nabla f(\phi_t(p)), X(\phi_t(p))) dt\]

\[= \int_{t_1}^{t_2} \|\nabla f(\phi_t(p))\|^2 dt = \int_{t_1}^{t_2} \|df(\phi_t(p))\|^2 dt.\]

On the other hand, using the metric $d$ on the Riemannian manifold $\mathcal{M}$, we have

\[(1.25) \quad d(\phi_{t_1}(p), \phi_{t_2}(p)) \leq \int_{t_1}^{t_2} \left\| \frac{d}{dt}(\phi_t(p)) \right\| dt\]

\[\leq \int_{t_1}^{t_2} \|X(\phi_t(p))\| dt = \int_{t_1}^{t_2} \|df(\phi_t(p))\| dt.\]

Squaring this and using the Cauchy-Schwarz inequality, we obtain

\[(1.26) \quad d(\phi_{t_1}(p), \phi_{t_2}(p))^2 \leq \left[ \int_{t_1}^{t_2} \|df(\phi_t(p))\| dt \right]^2\]

\[\leq (t_2 - t_1) \int_{t_1}^{t_2} \|df(\phi_t(p))\|^2 dt = (t_2 - t_1)(f(\phi_{t_1}(p)) - f(\phi_{t_2}(p))).\]

Let $\bar{t} = \sup\{t \in \mathbb{R} : \phi_t(p) \text{ is defined}\}$, and if $\bar{t} < \infty$, we choose a sequence $\{t_i\}$ of real numbers strictly less than $\bar{t}$ such that $t_i \to \bar{t}$. It follows from (1.26) that $\{\phi_{t_i}(p)\}$ is a Cauchy sequence in $\mathcal{M}$. Since $(\mathcal{M}, d)$ is a complete metric space by Proposition 1.9.5, $\phi_{t_i}(p) \to q$, for some $q \in \mathcal{M}$. But by the Existence and Uniqueness Theorem there is a flow box $(-\epsilon, \epsilon) \times U$ containing $(0, q)$, and this implies that the curve $t \mapsto \phi_t(p)$ can be extended beyond $\bar{t}$, giving a contradiction. Thus $\bar{t} = \infty$ and we see that $\phi_t(p)$ is defined for all $t \geq 0$, and the first statement of the theorem is proven.

Finally, it follows from (1.24) that

\[\int_0^{\infty} \|\nabla f(\phi_t(p))\|^2 dt < \infty,\]

and hence there must exist a sequence $t_i \to \infty$ such that

\[\|df(\phi_{t_i}(p))\| = \|\nabla f(\phi_{t_i}(p))\| \to 0.\]

By Condition C, a subsequence of $\{\phi_{t_i}(p)\}$ converges to a critical point for $f$, finishing the proof of the theorem.

Of course, we would like a version of the above theorem for the case of a $C^2$ function $f : \mathcal{M} \to [0, \infty)$, where $\mathcal{M}$ is only a Banach manifold. However, it
is not possible to define Riemannian metrics on \( M \) in this case, so we need a replacement for the notion of gradient, such as the following, a definition essentially due to Palais [Pal66]:

**Definition.** Suppose that \( f : M \to \mathbb{R} \) is a \( C^2 \) function on a Banach manifold which has a Finsler metric and let \( U \) be an open subset of \( M \). A \( C^1 \) vector field \( X : U \to T M \) is called a pseudogradient for \( f \) over \( U \) if there is a fixed constant \( \varepsilon > 0 \) such that for each \( p \in U \),

1. \( df_p(X(p)) > \varepsilon \| dp \|^2 \),
2. \( \| X(p) \| < (1/\varepsilon) \| dp \| \).

In both inequalities, \( \| dp \| = \sup \{| dp(v)| : v \in T_p M \text{ and } \| v \| \leq 1 \} \), which is the dual norm on the cotangent space to \( M \) at \( p \).

In the case of a Hilbert manifold, \( X = \text{grad}(f) \) satisfies both conditions in the definition with \( \varepsilon = 1 \); in other words, a gradient on a Hilbert manifold is also a pseudogradient. The definition was set up so that the following theorem would be true:

**Theorem 1.11.2.** Suppose that \( M \) is a Banach manifold with a complete Finsler metric \( \| \cdot \| \). Suppose that \( f : M \to [0, \infty) \) is a \( C^2 \) function which satisfies condition \( C \). Let

\[ K = \{ p \in M : df(p) = 0 \} \]

and let \( U = M - K \). If \( X \) is a pseudogradient for \( f \) on \( U \), and \{\( \phi_t \)\} is the local one-parameter group of diffeomorphisms corresponding to \(-X\), then

1. for each \( p \in U \), \( \phi_t(p) \) is defined for all \( t \geq 0 \), and
2. there is a sequence \( t_i \to \infty \) such that \{\( \phi_{t_i}(p) \)\} converges to a critical point for \( f \).

The proof is almost identical to the proof of Theorem 1.11.1. Assuming that \( p \) is not a critical point for \( f \) and that \( 0 < t_1 < t_2 \), we use the first condition in the definition of pseudogradient to replace (1.24) by the inequality

(1.27) \[ f(\phi_{t_1}(p)) - f(\phi_{t_2}(p)) = -\int_{t_1}^{t_2} \frac{d}{dt} f(\phi_t(p)) dt \]

\[ = \int_{t_1}^{t_2} df(\phi_t(p))(X(\phi_t(p))) dt \geq \varepsilon \int_{t_1}^{t_2} \| df(\phi_t(p)) \|^2 dt. \]

We use the second condition in the definition of pseudogradient to replace (1.25) by

(1.28) \[ d(\phi_{t_1}(p), \phi_{t_2}(p)) \leq \int_{t_1}^{t_2} \| X(\phi_t(p)) \| dt \leq \frac{1}{\varepsilon} \int_{t_1}^{t_2} \| df(\phi_t(p)) \| dt. \]
We square as before and use the Cauchy-Schwarz inequality to obtain

\begin{align*}
(1.29) \quad d(\phi_t^1(p), \phi_t^2(p))^2 &\leq \frac{1}{\varepsilon^2} \left[ \int_{t_1}^{t_2} \| df(\phi_t(p)) \| dt \right]^2 \\
&\leq \frac{t_2 - t_1}{\varepsilon^2} \int_{t_1}^{t_2} \| df(\phi_t(p)) \|^2 dt \leq \frac{t_2 - t_1}{\varepsilon^3} \left( f(\phi_t^1(p)) - f(\phi_t^2(p)) \right).
\end{align*}

We can now use (1.29) instead of (1.26) to show that \( \phi_t(p) \) is defined for all \( t \geq 0 \). Finally, it follows from (1.27) that

\[ \int_0^\infty \| df(\phi_t(p)) \|^2 dt < \infty, \]

which enables us to find a sequence \( t_i \to \infty \) such that \( \| df(\phi_{t_i}(p)) \| \to 0 \), and by condition C, a subsequence of \( \{ \phi_{t_i}(p) \} \) must converge to a critical point for \( f \).

Given a \( C^2 \) function \( f : \mathcal{M} \to [0, \infty) \) on a Banach manifold, how can we construct a corresponding pseudogradient on \( U = \mathcal{M} - K \), where \( K \) is the set of critical points for \( f \)? The standard technique consists of constructing pseudogradients over each set of an open cover of \( U \), and then piecing these together using a partition of unity.

Let \( \mathcal{M} \) be a metrizable infinite-dimensional smooth manifold modeled on a Banach space. According to a well-known theorem of Stone, \( \mathcal{M} \) must be paracompact. That means that every open cover of \( \mathcal{M} \) must have an open locally finite refinement.

**Definition.** Let \( \mathcal{U} = \{ U_\alpha : \alpha \in A \} \) be an open cover of \( \mathcal{M} \). A **partition of unity** subordinate to \( \mathcal{U} \) is a collection \( \{ \psi_\alpha : \alpha \in A \} \) of continuous real-valued functions on \( \mathcal{M} \) such that

1. \( \psi_\alpha : \mathcal{M} \to [0, 1] \),
2. the support of \( \psi_\alpha \) is a closed subset of \( U_\alpha \),
3. if \( p \in \mathcal{M} \), there is an open neighborhood \( V \) of \( p \) which intersects the supports of only finitely many \( \psi_\alpha \), and
4. \( \sum \psi_\alpha = 1 \).

It is known (and proven in topology texts) that any open cover of a paracompact Hausdorff space possesses a continuous partition of unity subordinate to a given open cover \( \mathcal{U} = \{ U_\alpha : \alpha \in A \} \).

Moreover, as proven in Lang [Lan95], \( C^\infty \) Hilbert manifolds possess \( C^\infty \) partitions of unity. However, for Banach manifolds, we encounter an obstacle: Banach manifolds need not possess \( C^\infty \) partitions of unity. As pointed out in the survey of Eells [Eel66], to construct \( C^k \) partitions of
1.11. Condition C

unity one needs to be able to construct nontrivial $C^k$ real-valued functions on the model Banach space $E$ with bounded support.

Fortunately, in the case where $\Sigma$ is a Riemann surface and $p > 2$ the Banach manifold $L^p_1(\Sigma, M)$ does possess partitions of unity of class $C^2$. To see why, we notice that the function

$$f : L^p_1(\Sigma, E) \to \mathbb{R}$$

defined by

$$f(\phi) = \|\phi\|_{L^p_1}^p$$

is $C^2$, by an argument similar to that given in Example 1.2.6. Let $g : \mathbb{R} \to [0, 1]$ be a smooth function such that

1. $g(s) = 1$ when $|s| \leq 1$, and
2. $g(s) = 0$ when $|s| \geq 2$.

Then the map

$$f_\varepsilon : L^p_1(\Sigma, \mathbb{R}^n) \to [0, 1]$$

defined by

$$f_\varepsilon(\phi) = g\left(\frac{2f(\phi)}{\varepsilon}\right)$$

is a $C^2$ function which equals one on a small neighborhood of the origin and has support contained in the set $\{\phi \in L^p_1(\Sigma, \mathbb{R}^n) : \|\phi\|^p \leq \varepsilon\}$. Using local coordinates, we can transport this function to $L^p_1(\Sigma, M)$ thereby obtaining a $C^2$ function $f : L^p_1(\Sigma, M) \to \mathbb{R}$ which is one in a neighborhood of a given point $p$ and vanishes outside a given open neighborhood of $p$.

From this start we can follow the familiar argument (given in Lang [Lan95], Chapter II, §3) for constructing $C^2$ partitions of unity subordinate to any open cover on the metrizable smooth manifold $L^p_1(\Sigma, M)$.

**Lemma 1.11.3.** If $f : L^p_1(\Sigma, M) \to \mathbb{R}$ is a $C^2$ function, where $p > 2$, then $f$ possesses a $C^2$ pseudogradient $\mathcal{X}$ which is tangent to every $L^p_k(\Sigma, M) \subseteq L^2_k(\Sigma, M)$, for $k \in \mathbb{N}$, $k \geq 2$.

**Proof.** Suppose that $0 < \varepsilon < 1$. If $p$ is not a critical point for $f$, we can choose a unit vector $u \in T_pM$ such that $|df_p(u)| > \sqrt{\varepsilon}\|df_p\|$; then

$$v = \sqrt{\varepsilon}\|df_p\|u$$

satisfies $\|v\| < \|df_p\|$, $df_p(v) > \varepsilon\|df_p\|^2$,

the two conditions in the definition of pseudogradient at the point $p$. We can extend $v$ to a smooth vector field which is a pseudogradient on some neighborhood of $p$; for example, we could choose it to be constant in terms of some smooth chart. Thus we can construct a pseudogradient on an open neighborhood about any point $p$ which is not in the set $K$ of critical points of $F$. If $M$ admits $C^2$ partitions of unity, one can piece together a $C^2$ pseudogradient on $M - K$. 
In the above construction, we can choose \( v \) to lie in the dense subspace \( L_k^2(\Sigma, M) \) of \( L_1^p(\Sigma, M) \) and the \( C^2 \) partition of unity on \( L_k^p(\Sigma, M) \) can be chosen so that it restricts to a \( C^2 \) partition of unity on \( L_k^p(\Sigma, M) \subseteq L_k^2(\Sigma, M) \) for every \( k \geq 2 \). When this is done, the pseudogradient will be tangent to every \( L_k^p(\Sigma, M) \) for \( k \geq 2 \), finishing the proof of the lemma. QED

### 1.12. Birkhoff’s minimax principle

Condition C can often be used to construct critical points for a \( C^2 \) function \( f : \mathcal{M} \to [0, \infty) \) corresponding to various topological constraints of \( \mathcal{M} \). Indeed, Palais [Pal70] describes how it provides a rigorous derivation of George Birkhoff’s “minimax principle”.

Suppose that \( \mathcal{M} \) is a Banach manifold with a complete Finsler metric and that \( f : \mathcal{M} \to [0, \infty) \) is a \( C^2 \) function satisfying Condition C. We let

\[
\mathcal{M}^a = \{ p \in \mathcal{M} : f(p) \leq a \}
\]

so that \( \mathcal{M}^a \subseteq \mathcal{M}^b \) when \( a \leq b \).

We say that \( c \) is a critical value for \( f \) if there is a critical point \( p \in \mathcal{M} \) for \( f \) with \( f(p) = c \). Our goal in this section is to use pseudogradients to show that if there are no critical values between \( a \) and \( b \) where \( a < b \), then \( \mathcal{M}^a \) and \( \mathcal{M}^b \) are homotopy equivalent. This fact sometimes allows us to prove the existence of critical points for \( f \).

**Definition.** For \( k \in \mathbb{N} \), a \( C^k \) ambient isotopy of \( \mathcal{M} \) is a \( C^k \) map

\[
\Psi : [0, 1] \times \mathcal{M} \to \mathcal{M}
\]

such that each map \( \psi_t : \mathcal{M} \to \mathcal{M} \), defined by \( \psi_t(p) = \Psi(t, p) \) for \( t \in [0, 1] \), is a diffeomorphism, with \( \psi_0 = \text{id} \). We will sometimes denote the ambient isotopy by \( \{\psi_t : t \in [0, 1]\} \).

**Theorem 1.12.1 (Deformation Theorem).** Suppose that \( \mathcal{M} \) is a Banach manifold with a complete Finsler metric and \( C^2 \) partitions of unity. If \( f : \mathcal{M} \to [0, \infty) \) is a \( C^2 \) function satisfying condition C and there are no critical points \( p \) for \( f \) such that \( a \leq f(p) \leq b \) where \( a < b \), then

1. there is a \( C^2 \) ambient isotopy \( \Psi = \{\psi_t : t \in [0, 1]\} \) of \( \mathcal{M} \) such that \( \psi_1(\mathcal{M}^b) \subseteq \mathcal{M}^a \), and
2. \( \mathcal{M}^a \) is a strong deformation retract of \( \mathcal{M}^b \).

**Proof.** Since \( f \) satisfies Condition C, there is a constant \( k > 0 \) such that \( \|df\| \geq k \) on \( \{ p \in \mathcal{M} : a \leq f(p) \leq b \} \). Indeed, if not, there would exist a sequence \( \{p_i\} \) in \( \{ p \in \mathcal{M} : a \leq f(p) \leq b \} \) such that \( \|df(p_i)\| \to 0 \). By condition C, a subsequence of \( \{p_i\} \) would converge to a critical point for \( f \) in \( \{ p \in \mathcal{M} : a \leq f(p) \leq b \} \), contradicting the hypothesis of the theorem.
similar argument shows that there is an \( \varepsilon > 0 \) such that there are no critical points \( p \) for \( f \) such that \( a - \varepsilon < f(p) < b + \varepsilon \). Using a \( C^2 \) partition of unity, we construct a \( C^2 \) function \( \eta : \mathcal{M} \to [0, 1] \) such that

1. \( \eta \equiv 1 \) on \( \{ p \in \mathcal{M} : a \leq f(p) \leq b \} \),
2. \( \eta \equiv 0 \) outside \( \{ p \in \mathcal{M} : a - \varepsilon < f(p) < b + \varepsilon \} \).

Let \( \mathcal{Y} \) be a \( C^2 \) pseudogradient for \( f \) on \( U = \mathcal{M} - K \), where \( K \) is the critical locus of \( F \) and set \( \mathcal{X} = \eta \mathcal{Y} \).

Theorem 1.11.2 gives us a one-parameter group \( \{ \phi_t : t \in [0, \infty) \} \) of local diffeomorphisms determined by \( -\mathcal{X} \) with \( \phi_t(p) \) defined for all \( p \in \mathcal{M} \) when \( t \geq 0 \) with \( f(\phi_t(p)) \) decreasing. We claim that if \( p \in \mathcal{M}^b \), then \( \phi_t(p) \in \mathcal{M}^a \) for \( t > (b - a)/(\varepsilon k) \). Indeed, if \( p \in \mathcal{M}^b - \mathcal{M}^a \) and \( \phi_t(p) \notin \mathcal{M}^a \), it follows from (1.27) that

\[
b - a \geq f(p) - f(\phi_t(p)) \geq \varepsilon \int_0^t \| df(\phi_\tau(p)) \| d\tau \geq \varepsilon k t
\]
or, equivalently, \( t \leq (b - a)/(\varepsilon k) \). We now set

\[
\psi_t = \phi_{ct}, \quad \text{where} \quad c = \frac{b - a}{\varepsilon k},
\]
and define

\[
\Psi : [0, 1] \times \mathcal{M} \to \mathcal{M} \quad \text{by} \quad \Psi(t, p) = \psi_t(p).
\]

Then \( \Psi \) is an ambient isotopy such that \( \psi_1(\mathcal{M}^b) \subset \mathcal{M}^a \).

This proves the first assertion of the Theorem. For the second, we let \( \tau(p) \) be the first time \( t \) such that \( \phi_t(p) \in \mathcal{M}^a \) and define \( \Phi : [0, 1] \times \mathcal{M}^b \to \mathcal{M}^b \) by

\[
\Phi(t, p) = \begin{cases} 
\phi_t(\tau(p))(p), & \text{for } p \in \mathcal{M}^b - \mathcal{M}^a, \\
p, & \text{for } p \in \mathcal{M}^a.
\end{cases}
\]

One then checks that \( \Phi \) is a strong deformation retraction from \( \mathcal{M}^b \) to \( \mathcal{M}^a \), finishing the proof of the theorem. QED

**Definition.** Let \( \mathcal{F} \) be a family of subsets of \( \mathcal{M} \). We say that \( \mathcal{F} \) is **ambient isotopy invariant** if

\[
A \in \mathcal{F} \quad \Rightarrow \quad \psi_1(A) \in \mathcal{F}
\]
whenever \( \{ \psi_t : t \in [0, 1]\} \) is a \( C^1 \) ambient isotopy of \( \mathcal{M} \). If \( \mathcal{F} \) is an ambient isotopy invariant family of sets we set

\[
\text{Minimax}(f, \mathcal{F}) = \inf \{ \sup \{ f(p) : p \in A \} : A \in \mathcal{F} \}.
\]
Theorem 1.12.2 (Birkhoff Minimax Principle). Suppose that $\mathcal{M}$ is a smooth Banach manifold with a complete Finsler metric and $C^2$ partitions of unity. Suppose, moreover, that $f : \mathcal{M} \to [0, \infty)$ is a $C^2$ function satisfying condition C and $\mathcal{F}$ is a nonempty family of subsets of $\mathcal{M}$ which is ambient isotopy invariant. Then $\text{Minimax}(f, \mathcal{F})$ is a critical value for $f$.

**Proof.** Let $c = \text{Minimax}(f, \mathcal{F})$. If $c$ is not a critical value for $f$, then Condition C implies that there exists an $\varepsilon > 0$ such that there are no critical points $p \in \mathcal{M}$ with $f(p) \in (c - \varepsilon, c + \varepsilon)$. But by definition of $\text{Minimax}(f, \mathcal{F})$, there exists an $A \in \mathcal{F}$ such that $A \subseteq \mathcal{M}^{c+\varepsilon}$. Theorem 1.12.1 then gives a smooth isotopy \{\psi_t : t \in [0, 1]\} such that

$$\psi_1(\mathcal{M}^{c+\varepsilon}) \subseteq \mathcal{M}^{c-\varepsilon}, \text{ so } \psi_1(A) \subseteq \mathcal{M}^{c-\varepsilon}.$$  

Since $\mathcal{F}$ is ambient isotopy invariant, $\psi_1(A) \in \mathcal{F}$ showing that $\text{Minimax}(f, \mathcal{F}) \leq c - \varepsilon$, a contradiction. QED

For the simplest application of the minimax principle, we take $\mathcal{M}_0$ to be a component of $\mathcal{M}$, and set

$$\mathcal{F} = \{A \subseteq \mathcal{M} : A \subseteq \mathcal{M}_0\}.$$  

Then $\mathcal{F}$ is ambient isotopy invariant and the previous theorem implies the following:

**Corollary 1.12.3.** Suppose that $\mathcal{M}$ is a smooth Banach manifold with a complete Finsler metric, and that $f : \mathcal{M} \to [0, \infty)$ is a $C^2$ function satisfying condition C. Then $f$ assumes its minimum value on each component of $\mathcal{M}$.

We can also use invariants from algebraic topology to construct critical points. For example, suppose that $\mathcal{M}$ is simply connected and $[\alpha]$ is a nonzero element in $\pi_k(\mathcal{M})$, the $k$-th homotopy group of $\mathcal{M}$. In this case, elements of $\pi_k(\mathcal{M})$ can be identified with free homotopy classes of maps $f : S^k \to \mathcal{M}$ and we can set

$$\mathcal{F}_{[\alpha]} = \{h(S^k) \text{ such that } h : S^k \to \mathcal{M} \text{ is a continuous map representing } [\alpha]\},$$  

a space which is nonempty and ambient isotopy invariant. Hence there is a minimax critical point corresponding to the homotopy class $[\alpha]$.

Alternatively, suppose that $x$ is a nonzero element in $H_k(\mathcal{M}; G)$, the singular homology group of $\mathcal{M}$ of degree $k$ with coefficients in $G$, and let

$$\mathcal{F}_x = \{\text{supp}(z) \text{ such that } z \text{ is a singular cycle representing } x\},$$  

where supp($z$) denotes the support of $z$. Once again $\mathcal{F}_x$ is ambient isotopy invariant, and one obtains a minimax critical point corresponding to each homology class $x$. Of course, there is a dual version for cohomology. In fact, there is a de Rham theory version which makes use of differential forms: If
\( M \) has \( C^2 \) partitions of unity and \( a \) is a differential \( k \)-form on \( M \) such that \( da = 0 \), we can set

\[ \mathcal{F}_a = \{ h(A) \mid A \text{ a compact oriented manifold of dimension } k \text{ and } h : A \to M \text{ is a } C^1 \text{-map such that } \int_A h^*a \neq 0 \} . \]

It follows from the Homotopy Lemma to be proven in the next section that \( \mathcal{F}_a \) is ambient isotopy invariant, so once again if \( \mathcal{F}_a \) is nonempty we obtain a minimax critical point corresponding to the closed differential form \( a \).

**Remark 1.12.4.** Minimax critical points for different topological constraints can be the same. For example, a smooth function \( f : S^n \to \mathbb{R} \) might have only two critical points even though the homotopy group \( \pi_k(S^n) \) is nonvanishing for arbitrarily high \( k \) when \( n \geq 2 \) (as explained in §4.1 of [Hat02]). Thus minimax critical points corresponding to nonzero elements of \( \pi_k(S^n) \) for many different values of \( k \) must agree.

### 1.13. de Rham cohomology

It is relatively straightforward to extend the standard treatments of de Rham cohomology to Banach manifolds which have \( C^2 \) partitions of unity. Indeed, if we let

\[ \Omega^k(M) = \{ C^1 \text{ differential forms } \omega \text{ on } M \text{ of degree } k : d\omega \in C^1 \} , \]

then whenever \( \psi \) is a \( C^2 \) smooth real-valued function,

\[ \omega \in \Omega^k(M) \Rightarrow \psi\omega \in \Omega^k(M) . \]

We can then make the \( \Omega^k(M) \)'s into a cochain complex in which the differential is the exterior derivative

\[ d : \Omega^k(M) \to \Omega^{k+1}(M) . \]

We say that an element \( \omega \in \Omega^k(M) \) is closed if \( d\omega = 0 \) and exact if \( \omega = d\theta \) for some \( \theta \in \Omega^{k-1}(M) \). Since \( d \circ d = 0 \), every exact \( k \)-form is closed. The quotient space

\[ H^k_{dR}(M; \mathbb{R}) = \frac{\text{closed elements of } \Omega^k(M)}{\text{exact elements of } \Omega^k(M)} \]

is called the de Rham cohomology of \( M \). If \( \omega \in \Omega^k(M) \) is closed, we let \([\omega] \) denote its cohomology class in \( H^k_{dR}(M; \mathbb{R}) \). In the terminology of algebraic topology, the de Rham cohomology is the cohomology of the cochain complex

\[ \cdots \to \Omega^{k-1}(M) \to \Omega^k(M) \to \Omega^{k+1}(M) \to \cdots . \]
Note that de Rham cohomology has a \textit{cup product} defined by
\[
[\omega] \cup [\phi] = [\omega \wedge \phi],
\]
which makes the direct sum
\[
H^*_dR(M; \mathbb{R}) = \sum_{k=0}^{\infty} H^k_dR(M; \mathbb{R})
\]
into a graded commutative algebra over \(\mathbb{R}\). Moreover, the cup product behaves well under smooth maps: If \(F : M \to N\) is a smooth map, the linear map \(F^*\) on differential forms induces a linear map
\[
F^* : H^k_dR(N; \mathbb{R}) \longrightarrow H^k_dR(M; \mathbb{R})
\]
such that \(F^*([\omega] \cup [\phi]) = F^*([\omega]) \cup F^*([\phi]).\)
Moreover, the identity map on \(M\) induces the identity on de Rham cohomology and if \(F : M \to N\) and \(G : N \to P\) are smooth maps, then \((G \circ F)^* = F^* \circ G^*\), so
\[
\mathcal{M} \mapsto H^k(M; \mathbb{R}), \quad (F : M \to N) \mapsto (F^* : H^k_dR(N; \mathbb{R}) \to H^k_dR(M; \mathbb{R}))
\]
is a contravariant functor from the category of smooth manifolds and smooth maps to the category of real vector spaces and linear maps.

\textbf{Lemma 1.13.1 (Poincaré Lemma).} If \(U\) is a convex open subset of a Banach space \(E\) or, more generally, any contractible open subset of \(E\), then the de Rham cohomology of \(U\) is trivial:
\[
H^k_dR(U; \mathbb{R}) \cong \begin{cases} 
\mathbb{R} & \text{if } k = 0, \\
0 & \text{if } k \neq 0.
\end{cases}
\]

One can modify the proof that is used in the finite-dimensional case. We only sketch the key ideas—the reader can refer to Lang [Lan95] for details. Since the inclusion from a point into the convex set \(U\) is a homotopy equivalence, the Poincaré Lemma is an immediate consequence of the next lemma.

\textbf{Lemma 1.13.2 (Homotopy Lemma).} Smoothly homotopic maps \(F, G : \mathcal{M} \to \mathcal{N}\) induce the same map on cohomology,
\[
F^* = G^* : H^k_dR(N; \mathbb{R}) \longrightarrow H^k_dR(M; \mathbb{R}).
\]
On the other hand, via functoriality, this follows from the special case of the Homotopy Lemma for the inclusion maps
\[
i_0, i_1 : \mathcal{M} \longrightarrow [0, 1] \times \mathcal{M}, \quad i_0(p) = (0, p), \quad i_1(p) = (1, p).
\]
Indeed, if $H : [0, 1] \times \mathcal{M} \rightarrow \mathcal{N}$ is a smooth homotopy from $F$ to $G$, then by definition of homotopy, $F = H \circ i_0$ and $G = H \circ i_1$, so

$$i_0^* = i_1^* \Rightarrow F^* = i_0^* \circ H^* = i_1^* \circ H^* = G^*.$$  

This special case, however, can be established by integrating over the fiber of the projection on the second factor $[0, 1] \times \mathcal{M} \rightarrow \mathcal{M}$. More precisely, let $t$ be the standard coordinate on $[0, 1]$, $T$ the vector field tangent to the fiber of $[0, 1] \times \mathcal{M}$ such that $dt(T) = 1$. We then define integration over the fiber

$$\pi_* : \Omega^k([0, 1] \times \mathcal{M}) \rightarrow \Omega^{k-1}(\mathcal{M}) \quad \text{by} \quad \pi_*(\omega)(p) = \int_0^1 (\iota_T \omega)(t, p) dt.$$  

Here the interior product $(\iota_T \omega)(t, p)$ is an element of $\Lambda^k T^*_t([0, 1] \times \mathcal{M})$ and the integration is possible because the exterior power at $(t, p)$ is canonically isomorphic to $\Lambda^k T^*_t([0, 1] \times \mathcal{M})$. The key to proving that $i_0^* = i_1^*$ in cohomology is the “cochain homotopy” formula

$$i_1^* \omega - i_0^* \omega = d(\pi_*(\omega)) + \pi_*(d\omega).$$  

This formula can be verified by using naturality to reduce the proof to the finite-dimensional case, and then calculating in local coordinates just as in the familiar finite-dimensional treatment found in [BT82]. Note that

$$d\omega = 0 \Rightarrow i_1^* \omega - i_0^* \omega = d(\pi_*(\omega)) \Rightarrow [i_1^* \omega] = [i_0^* \omega],$$  

and hence on the cohomology level $i_0^* = i_1^*$. This finishes our sketch of the proof of the Homotopy Lemma and the Poincaré Lemma.

**Remark 1.13.3.** If $\mathcal{N}$ is a submanifold of $\mathcal{M}$ with inclusion map $i : \mathcal{N} \rightarrow \mathcal{M}$, we let

$$\Omega^k(\mathcal{M}, \mathcal{N}) = \ker(i^* : \Omega^k(\mathcal{M}) \rightarrow \Omega^k(\mathcal{N})),$$  

and note that the exterior derivative makes this into the $k$-th cochain group of a cochain complex $\Omega^*(\mathcal{M}, \mathcal{N})$. The cohomology of this complex is called the *relative de Rham cohomology* of the pair $(\mathcal{M}, \mathcal{N})$ and is denoted by $H^k_{dR}(\mathcal{M}, \mathcal{N}; \mathbb{R})$.

The short exact sequence of cochain complexes

$$0 \rightarrow \Omega^*(\mathcal{M}, \mathcal{N}) \rightarrow \Omega^*(\mathcal{M}) \rightarrow \Omega^*(\mathcal{N}) \rightarrow 0$$  

yields a long exact sequence via the “snake lemma” (Theorem 2.16 in [Hat02]) from algebraic topology:

$$\cdots \rightarrow H^k_{dR}(\mathcal{M}, \mathcal{N}; \mathbb{R}) \rightarrow H^k_{dR}(\mathcal{M}; \mathbb{R}) \rightarrow H^k_{dR}(\mathcal{N}; \mathbb{R})$$  

$$\rightarrow H^{k+1}_{dR}(\mathcal{M}, \mathcal{N}; \mathbb{R}) \rightarrow H^{k+1}_{dR}(\mathcal{M}; \mathbb{R}) \rightarrow \cdots.$$  

This is very useful for calculating de Rham cohomology.
For us, one of the primary uses of differential forms will be to calculate cohomology of infinite-dimensional manifolds. It is important to realize that the de Rham cohomology is the same as the singular or Čech cohomology with real coefficients that is studied in algebraic topology:

**Theorem 1.13.4 (de Rham Theorem).** Suppose that either $M = L^p_k(\Sigma, M)$ where $p_k > \dim(\Sigma)$ or $M$ is finite-dimensional, and that $M$ admits $C^2$ partitions of unity. Then the cohomology of the cochain complex $\Omega^*(M)$ is isomorphic to the Čech cohomology of $M$.

The proof (due to André Weil) is via the zig-zag construction described in the excellent text on de Rham theory by Bott and Tu [BT82], which we follow closely.

In the case where $M$ is finite-dimensional, we let $U = \{U_\alpha : \alpha \in A\}$ be a locally finite open cover of $M$ by sets which are geodesically convex with respect to a Riemannian metric on $M$. If $M = L^p_k(\Sigma, M)$ where $p_k > \dim(\Sigma)$, we construct an open cover $U = \{U_\alpha : \alpha \in A\}$ in which each open set $U_\alpha$ is of the form

$$U_\alpha = \{g \in L^p_k(\Sigma, M) : \|g - f\|_{C^0} < \delta\}.$$  

The Sobolev Lemma guarantees the existence of such an open cover. We choose $\delta$ so small that if $p, q \in M$ and $d(p, q) < 2\delta$, then there is a unique minimizing geodesic

$$\gamma_{p,q} : [0, 1] \to M \text{ such that } \gamma_{p,q}(0) = p, \quad \gamma_{p,q}(1) = q,$$

and, moreover, this geodesic depends smoothly on $p$ and $q$. (Then any two points in a $\delta$-ball about a given point can be connected by a unique such geodesic.) If $g_1$ and $g_2$ are two elements of $U_\alpha$, we can then define a path

$$\Gamma_{g_1,g_2} : [0, 1] \to L^p_k(\Sigma, M) \text{ by } \Gamma_{g_1,g_2}(t)(p) = \gamma_{g_1(p),g_2(p)}(t).$$

It is easily checked in either case that the intersection of any collection of elements from $U$ is contractible. Readers familiar with cohomology theory will remember that the Čech cohomology of such an open covering is the same as the Čech cohomology of $M$ (and we do not need to take direct limits). Such an open cover is often called a “good” cover or “Leray” cover.

Given a good cover, we construct a double complex $K^{*,*}$ in which the $(p,q)$-element is

$$K^{p,q} = \check{C}^p(U, \Omega^q),$$

which is defined to be the space of functions $\omega$ which assign to each distinct ordered $(p+1)$-tuple $(\alpha_0, \ldots, \alpha_p)$ of indices such that $U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \neq \emptyset$ an element

$$\omega_{\alpha_0 \cdots \alpha_p} \in \Omega^q(U_{\alpha_0} \cap \cdots \cap U_{\alpha_p})$$
in such a way that if the order of elements in a sequence is permuted, \( \omega_{\alpha_0, \ldots, \alpha_p} \) changes by the change of the permutation; thus
\[
\omega_{\alpha_0 \alpha_1} = -\omega_{\alpha_1 \alpha_0}, \quad \omega_{\alpha \alpha} = 0,
\]
and so forth.

We have two differentials on the double complex, the exterior derivative
\[
d : \check{C}^p(U, \Omega^q) \to \check{C}^p(U, \Omega^{q+1})
\]
defined by \( (d\omega)_{\alpha_0 \ldots \alpha_p} = d\omega_{\alpha_0 \ldots \alpha_p} \),
and the Čech differential
\[
\delta : \check{C}^p(U, \Omega^q) \to \check{C}^{p+1}(U, \Omega^q)
\]
defined by
\[
(\delta \omega)_{\alpha_0 \ldots \alpha_p} = \sum_{i=0}^{p+1} (-1)^i \omega_{\alpha_0 \ldots \hat{\alpha}_i \ldots \alpha_{p+1}},
\]
the forms on the right being restricted to the intersection.

The first differential is exact except when \( q = 0 \) by the Poincaré Lemma, while in the case \( q = 0 \) we find that
\[
[\text{Kernel of } d : \check{C}^p(U, \Omega^0) \to \check{C}^p(U, \Omega^1)] = \check{C}^p(U; \mathbb{R}),
\]
the space of Čech cocycles for the covering \( U \) on \( M \). The Čech cohomology of the cover \( U \) is by definition the cohomology of the cochain complex
\[
(1.32) \quad \cdots \to \check{C}^{p-1}(U; \mathbb{R}) \to \check{C}^p(U; \mathbb{R}) \to \check{C}^{p+1}(U; \mathbb{R}) \to \cdots
\]
and is denoted by \( \check{H}^p(M; \mathbb{R}) \).

The second differential is exact except when \( p = 0 \) and it is at this point that the \( C^2 \) partition of unity \( \{ \psi_\alpha : \alpha \in A \} \) subordinate to \( U \) is used. Indeed, given a \( \delta \)-cocycle \( \omega \in \check{C}^p(U, \Omega^q) \), we set
\[
\tau_{\alpha_0 \ldots \alpha_{p-1}} = \sum_\alpha \psi_\alpha \omega_{\alpha_0 \ldots \alpha_{p-1}} \in \check{C}^{p-1}(U, \Omega^q),
\]
noting that since \( \psi_\alpha \) is \( C^2 \) we stay in the class of \( C^1 \) forms with \( C^1 \) exterior derivatives. Then
\[
(\delta \tau)_{\alpha_0 \ldots \alpha_p} = \sum_{\alpha, \alpha} (-1)^i \psi_\alpha \omega_{\alpha_0 \ldots \hat{\alpha}_i \ldots \alpha_p}
\]
and it follows from the fact that \( \delta \omega = 0 \) that
\[
\sum_{i=1}^{p} (-1)^i \omega_{\alpha_0 \ldots \hat{\alpha}_i \ldots \alpha_p} = \omega_{\alpha_0 \ldots \alpha_p}.
\]
Hence
\[(\delta \tau)_{\alpha_0 \cdots \alpha_p} = \left( \sum_{\alpha} \psi_\alpha \right) \omega_{\alpha_0 \cdots \alpha_p} = \omega_{\alpha_0 \cdots \alpha_p},\]
establishing exactness. When \(p = 0\), we find that
\[
\text{Kernel of } \delta : \check{C}^0(\mathcal{U}, \Omega^q) \to \check{C}^1(\mathcal{U}, \Omega^q) = \Omega^q(\mathcal{M}),
\]
the space of smooth \(q\)-forms on \(\mathcal{M}\).

We can summarize the previous discussion by stating that the rows and columns in the following commutative diagram are exact:

\[
\begin{array}{ccccccccc}
0 & \to & \Omega^2(\mathcal{M}) & \to & \check{C}^0(\mathcal{U}, \Omega^2) & \to & \check{C}^1(\mathcal{U}, \Omega^2) & \to & \check{C}^2(\mathcal{U}, \Omega^2) & \to & \cdots \\
& \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & \Omega^1(\mathcal{M}) & \to & \check{C}^0(\mathcal{U}, \Omega^1) & \to & \check{C}^1(\mathcal{U}, \Omega^1) & \to & \check{C}^2(\mathcal{U}, \Omega^1) & \to & \cdots \\
& \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
0 & \to & \Omega^0(\mathcal{M}) & \to & \check{C}^0(\mathcal{U}, \Omega^0) & \to & \check{C}^1(\mathcal{U}, \Omega^0) & \to & \check{C}^2(\mathcal{U}, \Omega^0) & \to & \cdots \\
& \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
\check{C}^0(\mathcal{U}, \mathbb{R}) & \to & \check{C}^1(\mathcal{U}, \mathbb{R}) & \to & \check{C}^2(\mathcal{U}, \mathbb{R}) & & & & & \uparrow & \uparrow & \uparrow \\
0 & 0 & 0 & & & & & & & & \\
\end{array}
\]

The remainder of the proof uses this diagram. Given a de Rham cohomology class \([\omega] \in H^p_{dR}(\mathcal{M}; \mathbb{R})\) with \(p\)-form representative \(\omega\) we construct a corresponding cohomology class \(s([\omega])\) in the Čech cohomology \(\check{H}^p(\mathcal{M}; \mathbb{R})\) as follows: The differential form defines an element \(\omega^{0p} \in \check{C}^0(\mathcal{U}, \Omega^p)\) by simply restricting \(\omega\) to the sets in the cover. It is readily checked that \(\omega^{0p}\) is closed with respect to the total differential \(D = \delta + (-1)^p d\) on the double complex \(K^{*,*} = \check{C}^*(\mathcal{U}, \Omega^*).\) Using the Poincaré Lemma, we construct an element \(\omega^{0,p-1} \in \check{C}^0(\mathcal{U}, \Omega^{p-1})\) such that \(d\omega^{0,p-1} = \omega^{0p}\). Let \(\omega^{1,p-1} = \delta \omega^{0,p-1}\) and observe that \(d\omega^{1,p-1} = 0\) and \(\omega^{1,p-1}\) is cohomologous to \(\omega^{0p}\) with respect to \(D\). Using the Poincaré Lemma again, we construct an element \(\omega^{1,p-2} \in \check{C}^1(\mathcal{U}, \Omega^{p-2})\) such that \(d\omega^{1,p-2} = \omega^{1,p-1}\). Let \(\omega^{2,p-2} = \delta \omega^{1,p-2}\) and note that \(\omega^{2,p-2}\) is cohomologous to \(\omega^{0p}\) with respect to \(D\). Continue in this fashion until we reach a \(D\)-cocycle \(\omega^{p0} \in \check{C}^p(\mathcal{U}, \Omega^0)\) which is cohomologous to \(\omega^{0p}\). Since \(d\omega^{p0} = 0\), each function \(\omega^{p0}_{\alpha_0 \cdots \alpha_p}\) is constant, and thus \(\omega^{p0}\) determines a Čech cocycle \(s(\omega)\) whose cohomology class is \(s([\omega])\).

By the usual diagram chasing, the cohomology class obtained is independent of choices made. Moreover, reversing the zig-zag construction described in the preceding paragraph yields an inverse to \(s\). This finishes our sketch of
the proof of de Rham’s theorem; for more details, one can consult [BT82], Chapter 2.

**Remark 1.13.5.** The proof shows that the cohomologies of the two cochain complexes (1.30) and (1.32) are isomorphic. It follows that the cohomology of the cochain complex (1.32) is independent of the choice of good cover. On the other hand, in the case where \( M \) has \( C^\infty \) partitions of unity the argument can be repeated with \( C^\infty \) differential forms to show that the de Rham cohomology is the same whether calculated with \( C^\infty \) forms or \( C^1 \) forms with \( C^1 \) exterior derivatives.

**Remark 1.13.6.** This remark assumes some familiarity with singular cohomology, as treated in Chapter 3 of [Hat02]. One could replace the double complex that occurs in the proof by

\[
K_{s}^{p,q} = C_{s}^{p}(U),
\]

which is defined to be the space of functions \( \omega \) which assign to each distinct ordered \((p + 1)\)-tuple \((\alpha_0, \ldots, \alpha_p)\) of indices such that \( U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \neq 0 \) an element \( s_{\alpha_0 \cdots \alpha_p} \) in the space of singular cochains within \( U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \) with coefficients in \( \mathbb{Z} \). The above proof can then be modified to give an isomorphism from singular cohomology to the Čech cohomology of \( M \).