Integration and Manifolds

Course No. 100 311

Fall 2007

Michael Stoll

Contents

1. Manifolds

Despite the title of the course, we will talk about manifolds first.

But before we go into that, let me remind you of two important theorems from Analysis II.

1.1. Implicit Function Theorem. *Let $W \subset \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ be open and $p = (p_1, p_2) \in W$. Let $f \in \mathcal{C}^q(W, \mathbb{R}^{n-k})$ (with $q \geq 1$) such that $f(p) = 0$; we write $f(z) = f(x, y)$ with $x \in \mathbb{R}^k$, $y \in \mathbb{R}^{n-k}$. If $D_y f_p$ is invertible, then there are open neighborhoods V_1 of p_1 in $\mathbb{R}^k$ and V_2 of p_2 in $\mathbb{R}^{n-k}$ such that $V_1 \times V_2 \subset W$, and there is $h \in \mathcal{C}^q(V_1, \mathbb{R}^{n-k})$ such that for all $x \in V_1$, $y \in V_2$, we have $f(x, y) = 0$ if and only if $y = h(x)$.*

This says that under an appropriate regularity condition on the derivative, we can locally solve an equation $f(x, y)$ by a smooth function $y = h(x)$.

1.2. Inverse Function Theorem. *Let $U \subset \mathbb{R}^n$ be open, $p \in U$, and $g \in \mathcal{C}^q(U, \mathbb{R}^n)$ (with $q \geq 1$). If Dg_p is invertible, then there are open neighborhoods $U' \subset U$ of p and V of $g(p)$ such that $g : U' \to V$ is bijective and $g^{-1} \in \mathcal{C}^q(V, \mathbb{R}^n)$.*

This says that if the derivative is invertible, then the function is locally invertible, and the inverse has the same "smoothness". The Inverse Function Theorem follows from the Implicit Function Theorem by considering the function $f(x, y) = x - g(y)$.

Back to manifolds.

Intuitively, k-dimensional manifolds are spaces that locally "look like $\mathbb{R}^k$". We distinguish several levels of "smoothness" that specify how closely like $\mathbb{R}^k$ our manifold is supposed to be. There are various ways to make this notion precise, and it turns out that they are all equivalent.

We first assume that the manifold-to-be M is embedded into real space $\mathbb{R}^n$.

1.3. Definition. Let $0 \leq k \leq n$ and $q \in \{1, 2, 3, \ldots, \infty\}$. A subset $M \subset \mathbb{R}^n$ is a *k-dimensional $\mathcal{C}^q$-submanifold of* $\mathbb{R}^n$ if the following equivalent conditions are satisfied.

(1) (M is locally a zero set)
For every $p \in M$ there is an open neighborhood W of p in $\mathbb{R}^n$ and a map $f \in \mathcal{C}^q(W, \mathbb{R}^{n-k})$ such that $W \cap M = f^{-1}(0)$ and such that Df_p has maximal rank ($= n - k$).

(2) (M locally looks like $\mathbb{R}^k \subset \mathbb{R}^n$)
For every $p \in M$ there is an open neighborhood W of p in $\mathbb{R}^n$ and open neighborhoods V_1 of 0 in $\mathbb{R}^k$ and V_2 of 0 in $\mathbb{R}^{n-k}$, and there is a $\mathcal{C}^q$-diffeomorphism $F : W \to V_1 \times V_2$ such that $F(W \cap M) = V_1 \times \{0\}$ and $F(p) = (0, 0)$.

(3) (M is locally diffeomorphic to $\mathbb{R}^k$)
For every $p \in M$ there is an open neighborhood U of p in M (i.e., $U = W \cap M$ with an open neighborhood W of p in $\mathbb{R}^n$), an open neighborhood V of 0 in $\mathbb{R}^k$, and a homeomorphism $\phi : U \to V$ such that $\phi^{-1} \in \mathcal{C}^q(V, \mathbb{R}^n)$, $\phi(p) = 0$ and $D(\phi^{-1})_0$ has maximal rank ($= k$).

(4) (M is locally a graph)
For every $p \in M$, there is a permutation matrix $P \in \mathrm{GL}_n(\mathbb{R})$, an open neighborhood W of p in $\mathbb{R}^n$ and open neighborhoods V_1 of 0 in $\mathbb{R}^k$ and V_2 of 0 in $\mathbb{R}^{n-k}$, and there is a map $g \in \mathcal{C}^q(V_1, V_2)$ with $g(0) = 0$ such that $W \cap M = \{p + P \cdot (x, g(x))^\top : x \in V_1\}$.

1.4. **Examples.** The unit circle in $\mathbb{R}^2$ is a 1-dimensional $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^2$, and the unit sphere in $\mathbb{R}^3$ is a 2-dimensional $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^3$. This can be seen in any of the four ways indicated in the definition above.

We still have to prove the equivalence of the four conditions.

Proof. We first show that (2) implies (1) and (3). So let $p \in M$, and V_1, V_2, W and F as in (2). For (1), we define $f = \pi_2 \circ F$, where $\pi_2 : V_1 \times V_2 \to V_2$ is the projection onto the second factor. Then $W \cap M = F^{-1}(V_1 \times \{0\}) = f^{-1}(0)$, and it remains to show that $\operatorname{rk} Df_p = n - k$. Since F is a diffeomorphism, we know that DF_p is invertible; also, $\operatorname{rk}(D\pi_2)_0 = n - k$, hence $Df_p = (D\pi_2)_0 \cdot DF_p$ has rank $n - k$ as well.

To show (3), we set $U = W \cap M$, $V = V_1$, and $\phi = \pi_1 \circ F \circ \iota$, where ι is the inclusion of U in W and $\pi_1 : V_1 \times V_2 \to V_1$ is the projection to the first factor. Then ϕ is continuous, and $\phi^{-1} = F^{-1} \circ \iota'$ (with $\iota' : V_1 \to V_1 \times \{0\} \subset V_1 \times V_2$) is a $\mathcal{C}^q$-function. We have $\phi(p) = \pi_1(F(p)) = 0$ and $D(\phi^{-1})_0 = (DF_p)^{-1} \cdot D\iota'_0$ has the same rank as $D\iota'_0$, which is k.

Let us show that (4) implies (2). Let $p \in M$, and assume P, W, V_1, V_2 and g given as in (4). Then we can define $F(z) = (x, y - g(x))$ where $(x, y)^\top = P^{-1} \cdot (z - p)$. It is clear that F is a $\mathcal{C}^q$-map. If we define for $x \in V_1$, $y \in \mathbb{R}^{n-k}$ the $\mathcal{C}^q$-map $G(x, y) = p + P \cdot (x, y + g(x))^\top$, then $G \circ F$ is the identity on W. Let V'_1, V'_2 be open neighborhoods of 0 in $\mathbb{R}^k, \mathbb{R}^{n-k}$ such that $W' = G(V'_1 \times V'_2) \subset W$. Then we obtain a $\mathcal{C}^q$-invertible map $F : W' \to V'_1 \times V'_2$ such that $F(p) = (0, 0)$ and $F(W' \cap M) = V'_1 \times \{0\}$.

The remaining implications "(1) $\Rightarrow$ (4)" and "(3) $\Rightarrow$ (4)" are less trivial, since there we need to reconstruct a piece of information. We will need to use some of the major theorems from the study of differentiable functions in several variables.

We show that (1) implies (4). Let $p \in M$ and take W and f as in (1). The matrix Df_p is a $(n - k) \times n$ matrix of rank $n - k$, therefore we can select $n - k$ of its columns that are linearly independent. Without loss of generality (and to ease notation), we can assume that these are the last $n - k$ columns and also that $p = 0$. If we write $z = (x, y)^\top$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^{n-k}$, then for $z \in W$, we have $z \in M \iff f(x, y) = 0$, and $D_y f_p$ is invertible. We can therefore apply the Implicit Function Theorem 1.1 and obtain neighborhoods V_1, V_2 of 0 in $\mathbb{R}^k, \mathbb{R}^{n-k}$ such that $V_1 \times V_2 \subset W$, and a map $g \in \mathcal{C}^q(V_1, \mathbb{R}^{n-k})$ such that for $z = (x, y) \in V_1 \times V_2$ we have $z \in M \iff f(x, y) = 0 \iff y = g(x)$.

Finally, we show that (3) implies (4). Let $p \in M$ and U, V, ϕ as in (3). We can assume that $p = 0$. Since $(D\phi^{-1})_0$ has rank r, we can pick r linearly independent rows of this matrix, and without loss of generality, these are the first r rows. If $\pi_1 : \mathbb{R}^n \to \mathbb{R}^k$ is the projection to the first k coordinates and $\pi_2 : \mathbb{R}^n \to \mathbb{R}^{n-k}$ is the projection to the last $n - k$ coordinates, then we have that $D(\pi_1 \circ \phi^{-1})_0$ is invertible. By the Inverse Function Theorem 1.2, there are open neighborhoods V_1, V' of 0 in $\mathbb{R}^k$ with $V' \subset V$ such that $\pi_1 \circ \phi^{-1} : V' \to V_1$ is bijective and the inverse h of this map is in $\mathcal{C}^q(V_1, \mathbb{R}^k)$. We let $g = \pi_2 \circ \phi^{-1} \circ h : V_1 \to \mathbb{R}^{n-k}$.

Also let $W \subset \mathbb{R}^n$ open such that $U = M \cap W$. If necessary, we replace V_1 and V' by smaller neighborhoods such that there is an open neighborhood V_2 of 0 in $\mathbb{R}^{n-k}$ with $g(V_1) \subset V_2$ and $V_1 \times V_2 \subset W$. Then for $z = (x, y) \in V_1 \times V_2$, we have $z \in M \iff z \in \phi^{-1}(V') \iff z \in (\phi^{-1} \circ h)(V_1)$. Further, $(\phi^{-1} \circ h)(x) = (x, g(x))$, hence $M \cap V_1 \times V_2$ is exactly the graph of g. □

1.5. **Examples.**

(1) Let $U \subset \mathbb{R}^n$ be open. Then U is an n-dimensional $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^n$. Indeed, we can use part (3) of Def. 1.3 above, with $U = V$ and $\phi = \mathrm{id}_U$.

(2) Let $p \in \mathbb{R}^n$. Then $\{p\}$ is a 0-dimensional $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^n$. We can see this, for example, from part (1) of Def. 1.3, by taking $W = \mathbb{R}^n$ and $f(x) = x - p$. More generally, any discrete set of points in $\mathbb{R}^n$ is a 0-dimensional submanifold. A subset S of $\mathbb{R}^n$ is *discrete* if every $p \in S$ has a neighborhood W in $\mathbb{R}^n$ such that $W \cap S = \{p\}$. For example, the set $S = \{\frac{1}{n} : n \in \{1, 2, \dots\}\} \subset \mathbb{R}$ is a 0-dimensional submanifold of $\mathbb{R}$. On the other hand, $S \cup \{0\}$ is not (Exercise!).

(3) Generalizing Example 1.4 above, define for $n \geq 1$ the *unit* $(n-1)$*-sphere*

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\} = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \dots + x_n^2 = 1\}.$$

Then $\mathbb{S}^{n-1}$ is an $(n-1)$-dimensional $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^n$. This is most easily seen using part (1) of Def. 1.3: we take $W = \mathbb{R}^n$ and $f(x) = \|x\|^2 - 1$; then $Df_p = 2p \neq 0$ for $p \in \mathbb{S}^{n-1}$.

At first sight, it may seem natural to consider submanifolds of $\mathbb{R}^n$, since in many cases, manifolds arise naturally in this context, for example as zero sets like the spheres. However, there are situations where spaces that one would like to consider as manifolds appear in a setting where there is no natural embedding into some $\mathbb{R}^n$. One instance of this is when we want to "quotient out" by a (nice) group action. For example, we may want to identify antipodal points on $\mathbb{S}^n$ (which means that we quotient by the action of the group $\{\pm 1\}$ acting by scalar multiplication). Since antipodal points never collapse, it is clear that the quotient (not-quite-yet-)manifold locally looks exactly like $\mathbb{S}^n$, and since we understand the quality of being a manifold as a local property, we would like to consider this quotient space as a manifold. However, there is no obvious way to realize this space as a submanifold of some $\mathbb{R}^N$.

Another situation where we run into similar problems is the construction of "parameter spaces". For example, consider the problem of describing the "space of all lines through the origin in $\mathbb{R}^n$". Given one line, it is fairly easy to describe (or parametrize) all nearby lines, and it is also fairly clear how to "glue together" these local descriptions. Again, we obtain something that locally looks like a piece of some $\mathbb{R}^k$ (here $k = n - 1$) and so morally should be a manifold. But again, there is no obvious or natural way to realize the whole of the space we construct as a submanifold of some $\mathbb{R}^N$.

(Note that the two sample problematic cases we have described are in fact equivalent, up to a shift of n — Exercise!)

The way out of these problems is to define manifolds in a more abstract and intrinsic way, without relying on the structure of an ambient space. If we look at the various parts of our Definition 1.3 above, we see that (1), (2), and (4) make use of the ambient $\mathbb{R}^n$ in essential way. On the other hand, (3) is less dependent on the ambient space; it is only used to ensure the correct smoothness (by requiring that ϕ^{-1} is of class $\mathcal{C}^q$ and is regular at 0). So we will base our new approach on this part of the definition. We have to remove the dependence on the ambient space for the smoothness part. The idea is to compare different "charts" of M with each other instead of a comparison of M sitting in $\mathbb{R}^n$ with something in $\mathbb{R}^k$.

We begin with some observations on submanifolds of $\mathbb{R}^n$.

1.6. **Lemma and Definition.** *Let M be a k-dimensional $\mathcal{C}^q$-submanifold of $\mathbb{R}^n$, $p \in M$. Let $\phi : U \to V$ be a map as in part (3) of Def. 1.3. Then there is an open neighborhood $p \in U' \subset U$ such that $(D\phi^{-1})_x$ has maximal rank k at all $x \in V' = \phi(U')$.*

A homeomorphism $\phi : U \to V$, where $p \in U \subset M$ and $0 \in V \subset \mathbb{R}^k$ are open neighborhoods, $\phi(p) = 0$, ϕ^{-1} is $\mathcal{C}^q$, and $D\phi^{-1}$ has rank k everywhere on V, is called a *$\mathcal{C}^q$-chart* of M centered at p.

Proof. By assumption, $(D\phi^{-1})_0$ has rank k, so there are k rows of this matrix that are linearly independent; in particular, the determinant of the corresponding $k \times k$ submatrix is nonzero. Since $(D\phi^{-1})_x$ varies continuously with $x \in V$, there is an open neighborhood $0 \in V' \subset V$ such that this determinant is nonzero for all $x \in V'$; in particular, $\mathrm{rk}(D\phi^{-1})_x = k$ for all $x \in V'$. The claim follows by setting $U' = \phi^{-1}(V')$. □

In this language, we can say that $M \subset \mathbb{R}^n$ is a k-dimensional $\mathcal{C}^q$-submanifold if and only if for every $p \in M$ there is a $\mathcal{C}^q$-chart $\phi : U \to \mathbb{R}^k$ centered at p.

1.7. **Lemma.** *Let M be a k-dimensional $\mathcal{C}^q$-submanifold of $\mathbb{R}^n$. Let $\phi : U \to V$ and $\phi' : U' \to V'$ be two $\mathcal{C}^q$-charts of M. Then the map*

$$\Phi = \phi' \circ \phi^{-1} : \phi(U \cap U') \longrightarrow \phi'(U \cap U')$$

is a $\mathcal{C}^q$-diffeomorphism.

Proof. Since ϕ and ϕ' are homeomorphisms, it is clear that Φ is also a homeomorphism (and well-defined). We have to show that Φ and Φ^{-1} are both of class $\mathcal{C}^q$. We would like to use the Implicit Function Theorem on $y = \Phi(x) \iff (\phi')^{-1}(y) - \phi^{-1}(x) = 0$, but this does not immediately work, as the values of $(\phi')^{-1}$ and of ϕ^{-1} are in $\mathbb{R}^n$ and not in $\mathbb{R}^k$. However, we can remedy this by (locally) picking k coordinates in a suitable way. So let $x_0 \in \phi(U \cap U')$, $y_0 \in \phi'(U \cap U')$ with $\Phi(x_0) = y_0$. We know that $\mathrm{rk}(D(\phi')^{-1})_{y_0} = k$, hence there is a projection $\pi : \mathbb{R}^n \to \mathbb{R}^k$ selecting k coordinates such that $(D(\pi \circ (\phi')^{-1}))_{y_0}$ is invertible. By the Implicit Function Theorem 1.1, there is a unique $\mathcal{C}^q$-function f such that $y = f(x)$ solves $\pi\big((\phi')^{-1}(y) - \phi^{-1}(x)\big) = 0$ in a neighborhood of (x_0, y_0). On the other hand, we know that $y = \Phi(x)$ is a solution as well. Hence $\Phi = f \in \mathcal{C}^q$ on this neighborhood, and since y_0 was arbitrary, we have $\Phi \in \mathcal{C}^q$. Interchanging the roles of ϕ and ϕ', we find that Φ^{-1} is also of class $\mathcal{C}^q$. □

This now allows us to reduce the definition of the "$\mathcal{C}^q$-structure" on a manifold to a comparison between charts.

1.8. **Definition.** Let M be a topological space, which we assume to be *Hausdorff* and to have a *countable basis.* A *(k-dimensional) chart* on M is a homeomorphism $\phi : U \to V$, where $U \subset M$ and $V \subset \mathbb{R}^k$ are open sets. We say that ϕ is *centered at* $p \in M$ if $p \in U$ and $\phi(p) = 0$.

Two k-dimensional charts $\phi : U \to V$ and $\phi' : U' \to V'$ on M are $\mathcal{C}^q$*-compatible* if the *transition map*

$$\Phi = \phi' \circ \phi^{-1} : \phi(U \cap U') \longrightarrow \phi'(U \cap U')$$

is a $\mathcal{C}^q$-diffeomorphism.

A set $\mathcal{A}$ of charts of M is a $\mathcal{C}^q$*-atlas* of M if for every $p \in M$ there is a chart $\phi : U \to V$ in $\mathcal{A}$ such that $p \in U$, and every pair of charts in $\mathcal{A}$ is $\mathcal{C}^q$-compatible. In this case, the pair $(M, \mathcal{A})$ is an *(abstract) k-dimensional $\mathcal{C}^q$-manifold.* Usually, we will just write M, with the atlas being understood.

A $\mathcal{C}^q$*-structure* on M is a maximal $\mathcal{C}^q$-atlas on M. If $(M, \mathcal{A})$ is a k-dimensional $\mathcal{C}^q$-manifold, there is a unique corresponding $\mathcal{C}^q$-structure, which is given by the set of all k-dimensional charts on M that are $\mathcal{C}^q$-compatible with all charts in the atlas $\mathcal{A}$. We do not distinguish between manifold structures on M given by atlases inducing the same $\mathcal{C}^q$-structure.

1.9. **Exercise.** Let $\mathcal{A}$ be a $\mathcal{C}^q$-atlas on M and let ϕ, ϕ' be two charts on M. Show that if both ϕ and ϕ' are $\mathcal{C}^q$-compatible with all charts in $\mathcal{A}$, then ϕ and ϕ' are $\mathcal{C}^q$-compatible with each other.

This implies that the set of all charts on M that are $\mathcal{C}^q$-compatible with $\mathcal{A}$ forms a $\mathcal{C}^q$-atlas, which then must be maximal. It is the $\mathcal{C}^q$-structure on M induced by the given atlas $\mathcal{A}$.

1.10. **Remarks.** A topological space X is *Hausdorff* if for any two distinct points $x, y \in X$ there are disjoint neighborhoods of x and y in X. The reason for this requirement is that we want to avoid pathological examples like the “real line with a double origin”. We obtain it by identifying to copies of $\mathbb{R}$ everywhere except at the origins. We wouldn't want to consider this a manifold, even though it satisfies the other required properties.

A *basis* of a topological space X is a collection of open sets of X such that every open set is a union of sets from the collection. For example, a (countable) basis of $\mathbb{R}^n$ is given by all open balls of rational radius centered at points with rational coordinates. The requirement of a countable basis is of a technical nature; it is needed to allow certain constructions (in particular, partitions of unity) which are necessary to obtain a useful theory.

Any subspace (subset with the induced topology) of a Hausdorff space with countable basis has the same two properties, so they are automatically satisfied for submanifolds of $\mathbb{R}^n$.

1.11. **Remark.** One could start with just a set M; then the topology is uniquely defined by the charts (by requiring that they are homeomorphisms). However, we still have to require that with this topology, M is a Hausdorff space with countable basis.

1.12. **Remark.** One can replace the condition "$\mathcal{C}^q$-diffeomorphism" in the definition of compatibility of charts by "homeomorphism". What one then gets are *topological manifolds*. We will not look at them in more detail, since we want to do analysis, and for this we need a differentiable (i.e., $\mathcal{C}^q$ for some q) structure. So in the following, "manifold" always means "differentiable manifold", i.e., "$\mathcal{C}^q$-manifold" for some $q \geq 1$.

1.13. **Example.** If M is a k-dimensional $\mathcal{C}^q$-submanifold of $\mathbb{R}^n$, then M is an abstract k-dimensional $\mathcal{C}^q$-manifold in a natural way: we take as atlas $\mathcal{A}$ the set of all charts $\phi : U \to V$ on M such that ϕ^{-1} is of class $\mathcal{C}^q$ as a map from V into $\mathbb{R}^n$. These are essentially the maps from part (3) of Def. 1.3. Lemma 1.7 shows that $\mathcal{A}$ really is an atlas, hence $(M, \mathcal{A})$ is a manifold.

1.14. **Example.** Let us show that $M = \mathbb{S}^n/\{\pm 1\}$ is a manifold in a natural way. We consider $\mathbb{S}^n$ as a $\mathcal{C}^\infty$-submanifold of $\mathbb{R}^{n+1}$. We take as atlas on $\mathbb{S}^n$ the orthogonal projections to hyperplanes, restricted to open half-spheres. Let $\iota : x \mapsto -x$ be the antipodal map on $\mathbb{S}^n$. Then for any ϕ in our atlas, $\phi \circ \iota$ is compatible with all charts in the atlas again.

The topology on M is defined by saying that $U \subset M$ is open if and only if its preimage $\pi^{-1}(U)$ is open in $\mathbb{S}^n$, where $\pi : \mathbb{S}^n \to M$ is the canonical map. It is then also true that the open sets of M are exactly the images under π of open sets of $\mathbb{S}^n$, which implies that M has a countable basis of the topology. Also, M is Hausdorff: Let $p, q \in M$ be distinct, and let $x, y \in \mathbb{S}^n$ with $\pi(x) = p$, $\pi(y) = q$. Then $x, -x, y, -y$ are pairwise distinct, hence we can find neighborhoods U of x and V of y in $\mathbb{S}^n$ such that $U, -U, V, -V$ are disjoint in pairs. Then $\pi(U)$ and $\pi(V)$ are disjoint neighborhoods of p and q, respectively.

We now define charts on M. Let $\phi : U \to \mathbb{R}^n$ be a chart in our atlas of $\mathbb{S}^n$; we have $U \cap -U = \emptyset$. Then $\pi|_U : U \to \pi(U)$ is a homeomorphism, and we let $\phi' = \phi \circ (\pi|_U)^{-1} : \pi(U) \to \mathbb{R}^n$ be a chart on M. Since for every $x \in \mathbb{S}^n$, there is a chart on $\mathbb{S}^n$ centered at x, it is clear that our charts cover M. Also, the transition maps between charts locally are transition maps between charts on $\mathbb{S}^n$ or charts composed with ι and therefore are $\mathcal{C}^\infty$-compatible. This gives us a $\mathcal{C}^\infty$-atlas on M.

1.15. **Remark.** In the same way, one can prove the following. Let M be a $\mathcal{C}^q$-manifold, and let Γ be a *finite* group acting on M such that no $x \in M$ is fixed by a non-identity element of Γ and such that for every ϕ in the differentiable structure of M and every $\gamma \in \Gamma$, $\phi \circ \gamma$ is again in the differentiable structure. (In the language we will soon introduce, this means that Γ acts on M via $\mathcal{C}^q$-diffeomorphisms.) Then the quotient M/Γ can be given the structure of a $\mathcal{C}^q$-manifold in a natural way.

(For infinite groups Γ, one has to require that for all $x, y \in M$ not in the same orbit under Γ, there are neighborhoods $x \in U$ and $y \in V$ such that $U \cap \gamma V = \emptyset$ for all $\gamma \in \Gamma$; this is to make sure the quotient space is Hausdorff.)

1.16. **Example.** Let us look at the second motivational example for introducing abstract manifolds and see how to turn the set of all lines through the origin in $\mathbb{R}^n$ into a manifold. We will first define charts, which then will also determine the topology. So let ℓ be a line through $0 \in \mathbb{R}^n$, and let $0 \neq x \in \ell$. Let E be the hyperplane orthogonal to x. Then we have a chart ϕ_x mapping the set of all lines not contained in E to E (which we can identify with $\mathbb{R}^{n-1}$) by setting $\phi^{-1}(y) = \mathbb{R}(x+y)$, the line through the point $x+y$. If x and x' are two generators of the same line ℓ, then the transition map between the charts ϕ_x and $\phi_{x'}$ is just a scaling. In general, the transition map is easily seen to be of the form

$$z \longmapsto \frac{\|y\|^2}{(x+z)\cdot y}(x+z) - y\,,$$

which is a $\mathcal{C}^\infty$ map. This gives us a $\mathcal{C}^\infty$-atlas for our manifold. We still need to check the topological conditions. We can take as a basis of the topology the images of balls around the origin of rational radius under ϕ_x^{-1}, where x has rational coordinates, so we have a countable basis. Also, it is easy to see that for any pair of distinct lines, we can find a chart that contains both of them, and then find disjoint neighborhoods within that chart, so the space is also Hausdorff.

1.17. **Remark.** Whitney proved in 1936 that every sufficiently smooth k-dimensional manifold can be embedded into $\mathbb{R}^{2k+1}$ (he later improved that to $\mathbb{R}^{2k}$ for $k > 2$). Given this, it would be sufficient to only consider submanifolds of $\mathbb{R}^n$. On the other hand, such an embedding may not be natural or easy to construct (try to do it for our examples above!) and is to some extent arbitrary, whereas the abstract notion of manifold only captures the intrinsic features of the given object, independently of the surrounding space. It thus lets us focus on the essentials, and we are not encumbered with the unnecessary construction of an embedding into $\mathbb{R}^n$.

2. Differentiable Maps and Tangent Spaces

Now after we have defined what manifolds are, we want to use them. For this, we need to consider maps between them. These maps should respect the differentiable structure. The way to define the concept of a differentiable map between manifolds is to make use of the charts. (The general philosophy here is to use the charts to reduce everything to known concepts on $\mathbb{R}^k$.)

2.1. **Definition.** Let M and M' be two $\mathcal{C}^q$-manifolds, let $f : M \to M'$ be continuous and $1 \le r \le q$. Let $p \in M$. Then f is *differentiable of class* $\mathcal{C}^r$ *at* p, if there are charts ϕ of M centered at p and ϕ' of M' centered at $f(p)$ such that $\phi' \circ f \circ \phi^{-1}$ is a $\mathcal{C}^r$-map in a neighborhood of 0. If f is $\mathcal{C}^r$ at all $p \in M$, then f is a $\mathcal{C}^r$*-map* or *differentiable of class* $\mathcal{C}^r$.

If $M' = \mathbb{R}$, then we speak of $\mathcal{C}^r$*-functions* on M; we write $\mathcal{C}^r(M)$ for the vector space of all such functions.

It is clear that if the defining property holds for some pair of charts ϕ, ϕ' (centered at p, $f(p)$), it will hold for all such pairs.

Note that $\phi' \circ f \circ \phi^{-1}$ is a map between subsets of spaces $\mathbb{R}^k$ and $\mathbb{R}^{k'}$ (if k and k' are the dimensions of M and M', respectively), hence for this map it makes sense to talk about being of class $\mathcal{C}^r$.

2.2. **Definition.** Let M and M' be two $\mathcal{C}^q$-manifolds, and let $f : M \to M'$ be a homeomorphism. If both f and f^{-1} are $\mathcal{C}^r$-maps, then f is called a $\mathcal{C}^r$-*diffeomorphism*, and M and M' are said to be $\mathcal{C}^r$*-diffeomorphic*.

2.3. **Example.** The antipodal map $x \mapsto -x$ is a $\mathcal{C}^\infty$-diffeomorphism of the sphere $\mathbb{S}^n$ to itself.

Our next goal will be to define derivatives of differentiable maps. Recall that the derivative of a map $\mathbb{R}^n \supset U \to \mathbb{R}^m$ at a point is a linear map that gives the best linear approximation to the given map near the given point. This works because $\mathbb{R}^n$ and $\mathbb{R}^m$ are vector spaces, so we can put a linear structure on the neighborhoods of our point and its image. With manifolds, we have the problem that we do not have such a linear structure. Therefore we first need to construct one, which means that we have to linearize the manifolds, too, not only the map. This leads to the notion of *tangent space.*

The tangent space to a manifold at a point is easy to define when we have a submanifold of $\mathbb{R}^n$.

2.4. **Definition.** Let M be a k-dimensional submanifold of $\mathbb{R}^n$, and let $p \in M$. Let $\phi : U \to \mathbb{R}^k$ be a chart of M centered at p. Then the *tangent space* to M at p, T_pM, is the linear subspace $(D\phi^{-1})_0(\mathbb{R}^k)$ of $\mathbb{R}^n$.

It is clear that this does not depend on the chart: changing the chart, we precompose with the derivative of the transition map, which is an invertible linear map and so does not change the image space.

The idea is that near p, M is the image of the differentiable map ϕ^{-1}, hence the best linear approximation to M is by the image of the derivative of ϕ^{-1} at the point we are interested in.

Note that the tangent space T_pM is a linear subspace of $\mathbb{R}^n$. If we shift its origin to p, then $p+T_pM$ is tangent to M at p in the geometric sense: it is the tangent line to M at p when M is of dimension 1, the tangent plane, when M is of dimension 2, and so on.

2.5. **Exercise.** Write the tangent space T_pM in terms of the data given in parts (1), (2), and (4) of Def. 1.3.

2.6. **Example.** Consider the "north pole" $N = (0, \dots, 0, 1) \in \mathbb{S}^n$. To find the tangent space $T_N\mathbb{S}^n$, we take the chart

$$\phi : \{x \in \mathbb{S}^n : x_{n+1} > 0\} \longrightarrow B_1(0) \subset \mathbb{R}^n\,, \quad x \longmapsto (x_1, \dots, x_n)\,;$$

its inverse is $\phi^{-1}(y) = (y, \sqrt{1 - \|y\|^2})$, with derivative $(D\phi^{-1})_0 : z \mapsto (z, 0)$. Hence $T_N\mathbb{S}^n$ is the "equatorial hyperplane" $x_{n+1} = 0$.

(Using the previous exercise and writing $\mathbb{S}^n$ as the zero set of $\|x\|^2 - 1$, we easily find that $T_x\mathbb{S}^n$ is the hyperplane orthogonal to x, for all $x \in \mathbb{S}^n$.)

If we want to define tangent spaces for abstract manifolds, we cannot use the vector space structure of an ambient $\mathbb{R}^n$, as we did above. We have to resort to using charts again. Note that if M is a submanifold of $\mathbb{R}^n$ and $p \in M$, then any chart ϕ of M centered at p sets up an isomorphism of $\mathbb{R}^k$ with T_pM via $(D\phi^{-1})_0$ (where $k = \dim M$ is the dimension of M). So we can take the tangent space to be $\mathbb{R}^k$, but we have to be careful how the various copies of $\mathbb{R}^k$ given by the various charts are identified with each other.

2.7. **Definition.** Let M be an abstract k-dimensional $\mathcal{C}^q$-manifold, and let $p \in M$. A *tangent vector* to M at p is an equivalence class of pairs (ϕ, v), consisting of a chart $\phi : U \to V \subset \mathbb{R}^k$ of M centered at p and a vector $v \in \mathbb{R}^k$.

Two such pairs (ϕ, v) and (ϕ', v') are equivalent if $D\Phi_0 \cdot v = v'$, where $\Phi = \phi' \circ \phi^{-1}$ is the transition map.

The *tangent space* T_pM of M at p is the set of all tangent vectors to M at p.

2.8. **Exercise.** Verify that the relation between pairs (ϕ, v) defined above really is an equivalence relation.

We need to justify the name tangent *space*.

2.9. **Lemma.** *In the situation of the definition above, T_pM is a k-dimensional real vector space in a natural way.*

Proof. First note that if ϕ and ϕ' are two charts centered at p, then for every $v \in \mathbb{R}^k$, there is a unique $v' = D\Phi_0 \cdot v \in \mathbb{R}^k$ such that $(\phi, v) \sim (\phi', v')$. We fix one chart ϕ; then we get an identification of T_pM with $\mathbb{R}^k$, which we use to define the vector space structure on T_pM. Since the identifying maps $D\Phi_0$ are invertible linear maps, this structure does not depend on which chart ϕ we pick. $\square$

2.10. **Derivations on $\mathbb{R}^k$.** Our definition of tangent vectors is fairly concrete, but maybe not very satisfying, since it is not intrinsic. There are alternative definitions that are better in this respect.

We will now assume that everything (manifolds, differentiable maps, ...) is $\mathcal{C}^\infty$. This is sufficient for most applications and saves us the trouble of keeping track of how differentiable our objects are.

First consider an open neighborhood of $0 \in \mathbb{R}^k$. To every $v \in \mathbb{R}^k$, considered as a tangent vector to $\mathbb{R}^k$ in 0, we can associate a linear form $\mathcal{C}^\infty(V) \to \mathbb{R}$, given by the directional derivative in the direction of v:

$$\partial_v : f \longmapsto \frac{d}{dt} f(tv)\Big|_{t=0} .$$

This map has the additional property $\partial_v(fg) = f(0)\partial_v(g) + g(0)\partial_v(f)$, coming from the Leibniz rule. Conversely, every linear form ∂ on $\mathcal{C}^\infty(V)$ that satisfies $\partial(fg) = f(0)\partial(g) + g(0)\partial(f)$ is of the form $\partial = \partial_v$ for some $v \in \mathbb{R}^k$. To see this, note that every $f \in \mathcal{C}^\infty(V)$ can be written as

$$f(x_1, \dots, x_k) = c + f_1(x)x_1 + \cdots + f_k(x)x_k$$

with functions $f_1, \dots, f_k \in \mathcal{C}^\infty(V)$ and $c \in \mathbb{R}$. Then

$$\partial(f) = f_1(0)\partial(x_1) + \cdots + f_k(0)\partial(x_k) = \frac{\partial f}{\partial x_1}(0)v_1 + \cdots + \frac{\partial f}{\partial x_k}(0)v_k = \partial_v(f) ,$$

where $v_j = \partial(x_j)$ and $v = (v_1, \dots, v_k)$. We can therefore identify $\mathbb{R}^k = T_0V$ with the space of all these *derivations* ∂ on $\mathcal{C}^\infty(V)$. With respect to this identification, it makes sense to denote the standard basis of the space by $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_k}$.

2.11. **Derivations on Manifolds.** Now suppose we have a k-dimensional manifold M and a point $p \in M$. Assume ϕ and ϕ' are two charts of M centered at p and that $f \in \mathcal{C}^\infty(M)$. Let $(\phi, v) \sim (\phi', v')$ represent a tangent vector to M at p. Then I claim that

$$\partial_v(f \circ \phi^{-1}) = \partial_{v'}(f \circ (\phi')^{-1}).$$

In particular, we obtain the same derivation

$$\partial : \mathcal{C}^\infty(M) \longrightarrow \mathbb{R}, \qquad f \longmapsto \partial_v(f \circ \phi^{-1}) = \partial_{v'}(f \circ (\phi')^{-1}).$$

To see the claimed equality, let $\Phi = \phi' \circ \phi^{-1}$ be the transition map. Also note that $\partial_v(h) = Dh_0 \cdot v$. We obtain

$$\begin{aligned} \partial_v(f \circ \phi^{-1}) &= D(f \circ \phi^{-1})_0 \cdot v = D(f \circ (\phi')^{-1} \circ \Phi)_0 \cdot v \\ &= D(f \circ (\phi')^{-1})_0 D\Phi_0 \cdot v = D(f \circ (\phi')^{-1})_0 \cdot v'. \end{aligned}$$

What we get out of this is:

The tangent space T_pM can be identified in a natural way with the space of all derivations on M at p.

Here, a *derivation on M at p* is a linear form $\partial : \mathcal{C}^\infty(M) \to \mathbb{R}$ such that

$$\partial(fg) = f(p)\partial(g) + g(p)\partial(f) \quad \text{for all } f, g \in \mathcal{C}^\infty(M).$$

Our intuition about directional derivatives is still valid on a manifold M. However, we cannot use straight lines $t \mapsto tv$, so we have to resort to more general curves.

2.12. **Definition.** Let M be a manifold. A *smooth curve* on M is a $\mathcal{C}^\infty$-map from an open interval in $\mathbb{R}$ to M.

Now let $p \in M$, where M is a manifold, and let $\gamma :]-\varepsilon, \varepsilon[\to M$ be a smooth curve with $\gamma(0) = p$. Then

$$\partial : f \longmapsto \frac{d}{dt} f\big(\gamma(t)\big)\Big|_{t=0}$$

is a derivation on M at p and hence corresponds to a tangent vector to M at p, which we can consider to be the *velocity* of the curve γ at $t = 0$. Conversely, every tangent vector can be obtained in this way — just transport a piece of the straight line $t \mapsto tv$ from the chart ϕ to M, where the tangent vector is represented by (ϕ, v).

In this way, we can also express the tangent space as the set of equivalence classes of all curves γ as above, with two curves being considered equivalent if they induce the same derivation on M at p.

2.13. **Exercise.** Write out what this equivalence means in terms of charts.

With the tangent space at hand, we can now define what the derivative of a differentiable map should be.

2.14. **Definition.** Let $f : M \to M'$ be a differentiable map between manifolds, and let $p \in M$ be a point. Then the *derivative* of f at p is the linear map

$$Df_p : T_pM \longrightarrow T_{f(p)}M' , \quad \partial \longmapsto \big(h \mapsto \partial(h \circ f)\big) .$$

Here, we identify the tangent spaces with the corresponding spaces of derivations.

In terms of our earlier definition of tangent vectors as equivalence classes of pairs (ϕ, v), we get the following. Let ϕ and ψ be charts of M and M', centered at p and $f(p)$, respectively. Then

$$Df_p\big((\phi, v)\big) = \big(\psi, D(\psi \circ f \circ \phi^{-1})_0 \cdot v\big) .$$

So what we get is just the derivative of our map when expressed in terms of the charts.

To see this, recall that (ϕ, v) corresponds to the derivation $g \mapsto D(g \circ \phi^{-1})_0 \cdot v$. The image under Df_p therefore is

$$h \longmapsto D(h \circ f \circ \phi^{-1})_0 \cdot v = D(h \circ \psi^{-1})_0 \cdot D(\psi \circ f \circ \phi^{-1})_0 \cdot v ,$$

which in turn is the derivation corresponding to $\big(\psi, D(\psi \circ f \circ \phi^{-1})_0 \cdot v\big)$.

2.15. **Lemma.** *Let M be a manifold, $p \in M$. Then for $f, g \in \mathcal{C}^\infty(M)$, we have the usual rules*

$$D(f \pm g)_p = Df_p \pm Dg_p \qquad \textit{and} \qquad D(fg)_p = f(p)Dg_p + g(p)Df_p .$$

Proof. Exercise. □

2.16. **Lemma.** *Let M be a manifold, $p \in M$, $v \in T_pM$. Then v corresponds to the derivation $f \mapsto Df_p \cdot v$ on M at p.*

Proof. Exercise. □

The following generalizes the Chain Rule to maps between manifolds.

2.17. **Proposition.** *Let M, M' and M'' be manifolds, $f : M \to M'$, $g : M' \to M''$ differentiable maps. Let $p \in M$. Then $D(g \circ f)_p = Dg_{f(p)} \circ Df_p$.*

Proof. Let $h \in \mathcal{C}^\infty(M'')$ and $\partial \in T_pM$ a derivation on M at p. Then on the one hand,

$$\big(D(g \circ f)_p(\partial)\big)(h) = \partial(h \circ g \circ f) ,$$

and on the other hand,

$$\big((Dg_{f(p)} \circ Df_p)(\partial)\big)(h) = \big(Dg_{f(p)}(Df_p(\partial))\big)(h) = \big(Df_p(\partial)\big)(h \circ g) = \partial(h \circ g \circ f) .$$

□

2.18. **Example.** Let M be a manifold and $\gamma : \,]-\varepsilon, \varepsilon[\, \to M$ a smooth curve on M with $\gamma(0) = p$. Then the velocity of γ at 0 is the same as $D\gamma_0$ applied to $1 \in \mathbb{R} = T_0\,]-\varepsilon, \varepsilon[$. Indeed, the velocity $v \in T_pM$ is the derivation

$$f \longmapsto \frac{d}{dt} f\big(\gamma(t)\big)|_{t=0} = D(f \circ \gamma)_0 \cdot 1 = Df_p \cdot (D\gamma_0 \cdot 1)$$

(by the Chain Rule), which by Lemma 2.16 corresponds to $D\gamma_0 \cdot 1 \in T_pM$.

On submanifolds, we can get differentiable maps and their derivatives quite easily.

2.19. **Proposition.** *Let $M \subset \mathbb{R}^n$, $M' \subset \mathbb{R}^{n'}$ be submanifolds, and let $f : U \to \mathbb{R}^{n'}$ be a differentiable map such that $M \subset U \subset \mathbb{R}^n$ and $f(M) \subset M'$. Then the restriction $\tilde{f} : M \to M'$ is a differentiable map between manifolds, and for $p \in M$, we have $D\tilde{f}_p = Df_p|_{T_pM} : T_pM \to T_{f(p)}M'$, where the tangent spaces are identified with linear subspaces of the ambient spaces.*

Proof. Let ϕ be a chart of M centered at p. Let $\Psi : U' \to V \subset \mathbb{R}^{n'}$ be a map as in part (2) of Def. 1.3, such that $f(p) \in U'$ and $\Psi(f(p)) = 0$. Then (denoting $k' = \dim M'$) $\psi = \Psi|_{U' \cap M'} \to V \cap (\mathbb{R}^{k'} \times \{0\})$ is a chart on M' centered at $f(p)$. Let $\pi : \mathbb{R}^{n'} \to \mathbb{R}^{k'}$ be the projection to the first k' coordinates. Then $\psi \circ \tilde{f} \circ \phi^{-1} = \pi \circ \Psi \circ f \circ \phi^{-1}$ is a composition of differentiable maps and therefore again a differentiable map (recall that 'differentiable' means '$\mathcal{C}^\infty$'). By definition, this means that $\tilde{f}$ is differentiable. Using the identifications of the various versions of the tangent spaces, we find that

$$D\tilde{f}_p = (D\psi^{-1})_0 \circ D(\psi \circ f \circ \phi^{-1})_0 \circ (D\phi^{-1})_0^{-1} \,,$$

where the latter is the inverse of the isomorphism $(D\phi^{-1})_0$ between $\mathbb{R}^{\dim M}$ and $T_pM \subset \mathbb{R}^n$. By the chain rule, this gives

$$D\tilde{f}_p = D(f \circ \phi^{-1})_0 \circ (D\phi^{-1})_0^{-1} = Df_p|_{T_pM} \,.$$

Since $(D\psi^{-1})_0$ maps into $T_{f(p)}M'$, we obtain a linear map as claimed. □

2.20. **Example.** Consider the "z-coordinate map" $z : \mathbb{S}^2 \to \mathbb{R}$. By Prop. 2.19, it is a differentiable map. For $x \in \mathbb{S}^2$, we have $T_x\mathbb{S}^2 = x^\perp$ and

$$Dz_x : x^\perp \ni v \longmapsto \langle v, e_3 \rangle \,.$$

This linear map is nonzero and therefore of rank 1 if and only if x is not a multiple of e_3. Otherwise (i.e., when x is the "north pole" or "south pole"), we have $Dz_x = 0$.

2.21. **Definition.** Let M be a manifold, $f \in \mathcal{C}^\infty(M)$. A point $p \in M$ is called a *critical point* of f if $Df_p = 0$. In this case, $f(p) \in \mathbb{R}$ is called a *critical value* of f.

So in our example above, the critical points are the north and south poles, and the critical values are ± 1. We will come back to these notions later.

3. Vector Bundles and the Tangent Bundle

We would like to take the set of all tangent vectors at all points p on a manifold M together and turn it into a manifold TM, the *tangent bundle* of M. This should come with a map $TM \to M$ that associates to a tangent vector the point it is based at, with fiber over p the tangent space T_pM. We want to include the linear structure of the tangent spaces as part of the structure of TM. This leads to the notion of *vector bundles* over M.

3.1. **Definition.** Let E and M be manifolds and $\pi : E \to M$ a surjective differentiable map. Then E is a *vector bundle* over E of rank n, if for every $p \in M$, there is a neighborhood U of p and a diffeomorphism $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^n$ such that $\mathrm{pr}_1 \circ \psi = \pi$; furthermore, we require that for each pair $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^n$ and $\psi' : \pi^{-1}(U') \to U' \times \mathbb{R}^n$ of such diffeomorphisms, the transition map $\psi' \circ \psi^{-1} : (U \cap U') \times \mathbb{R}^n \to (U \cap U') \times \mathbb{R}^n$ is of the form $(p, v) \mapsto (p, g(p) \cdot v)$, where $g : U \cap U' \to \mathrm{GL}(n, \mathbb{R})$ is a $\mathcal{C}^\infty$-map.

It is then clear that the vector space structure induced on fibers of π by these charts ψ is independent of the chart. We can therefore think of $E \to M$ as a smoothly varying family of n-dimensional vector spaces, one for each $p \in M$. We denote the fiber $\pi^{-1}(p)$ by E_p.

If $n = 1$, we call E a *line bundle* over M.

A *homomorphism* of vector bundles over M is a differentiable map $f : E_1 \to E_2$, where $\pi_1 : E_1 \to M$ and $\pi_2 : E_2 \to M$ are vector bundles over M, such that $\pi_2 \circ f = \pi_1$ and such that the induced map $f_p : (E_1)_p \to (E_2)_p$ is linear for every $p \in M$. It is then clear what an *isomorphism* of vector bundles should be.

We call E a *trivial* vector bundle, if we can take $U = M$ in the definition above. Then $E \cong M \times \mathbb{R}^n$ as vector bundles. Accordingly, we call a map like ψ above a *local trivialization* of $E \to M$.

3.2. **Example.** Let $M = \mathbb{S}^1$ be the circle, and let

$$E = \{(x, y) \in \mathbb{S}^1 \times \mathbb{R}^2 : y \in \mathbb{R}x\} \subset \mathbb{R}^4 .$$

Then E is a trivial line bundle over $\mathbb{S}^1$, since we can map E to $\mathbb{S}^1 \times \mathbb{R}$ by sending (x, y) to (x, λ), where $y = \lambda x$.

On the other hand, identifying $\mathbb{S}^1 \subset \mathbb{C}$ with the group of complex numbers of absolute value 1, we can define a line bundle

$$E' = \{(z, y) \in \mathbb{S}^1 \times \mathbb{C} : y \in \mathbb{R}\sqrt{z}\} \subset \mathbb{R}^4 ,$$

where $\sqrt{z}$ is one of the two square roots of z (they differ by a sign, so span the same real subspace of $\mathbb{C}$). This line bundle is non-trivial (Exercise).

In fact, E and E' are the only two line bundles on $\mathbb{S}^1$, up to isomorphism. Roughly speaking, this comes from the fact that when you walk around the circle once, you have the choice of identifying the fiber with the copy you were carrying with you either in an orientation-preserving or in an orientation-reversing way.

3.3. **The Tangent Bundle.** The motivation for introducing vector bundles was our wish to construct the tangent bundle. So we do that now. Let M be a manifold. Then locally, via a chart, we can identify the tangent bundle with $V \times \mathbb{R}^k$ (where $k = \dim M$ and V is some open subset of $\mathbb{R}^k$). We use the chart to go to $U \times \mathbb{R}^k$, where $U \subset M$ is open.

More precisely, let $TM = \{(p, v) : p \in M, v \in T_pM\}$ as a set, and let $\pi : TM \to M$ be the projection to the first component. Let $\phi : U \to V \subset \mathbb{R}^k$ be a chart on M. Then we define $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ by $\psi(p, v) = (p, w)$, where $(\phi - \phi(p), w)$ represents the tangent vector $v \in T_pM$ (note that $\phi - \phi(p)$ is a chart centered at p). This is clearly a bijection. We check the compatibility. So let ψ' be another map as above, constructed from another chart $\phi' : U' \to V'$. We find that

$$(\psi' \circ \psi^{-1})(p, w) = (p, D\Phi_{\phi(p)} \cdot w)$$

where $\Phi = \phi' \circ \phi^{-1}$ is the transition map between the charts of M, compare Def. 2.7. Since Φ is a $\mathcal{C}^\infty$ map, the same is true of $p \mapsto D\Phi_{\phi(p)} \in \mathrm{GL}(k, \mathbb{R})$.

The compatibility of the 'local trivializations' ψ then allows us to endow TM with a topology and differentiable structure. The topology is Hausdorff, since (p, v) and (q, w) can be separated by neighborhoods of the form $\pi^{-1}(U)$ and $\pi^{-1}(V)$ if $p \neq q$, and (p, v) and (p, v') can be separated by neighborhoods pulled back from sets of the form $U \times V$ and $U \times V'$ under a suitable map ψ. We can construct a countable basis of the topology by letting (U_i) be a countable basis of the topology on M such that each U_i is the domain of a chart and (V_j) a countable basis of the topology of $\mathbb{R}^k$; then $\big(\psi_i^{-1}(U_i \times V_j)\big)_{i,j}$ is a countable basis of the topology of TM, where ψ_i is the local trivialization of TM constructed from the chart ϕ_i defined on U_i.

3.4. **Example.** If $M \subset \mathbb{R}^n$ is a k-dimensional submanifold, then we can realize TM as a $2k$-dimensional submanifold of $\mathbb{R}^{2n}$: we have seen that we can identify T_pM with a k-dimensional linear subspace of $\mathbb{R}^n$, so

$$TM = \{(p, v) \in M \times \mathbb{R}^n : v \in T_pM\} \subset \mathbb{R}^{2n} ,$$

and one can check that the differentiable structure induced by this embedding in $\mathbb{R}^{2n}$ agrees with the abstract one we have defined.

3.5. **Lemma.** *Let $f : M \to M'$ be a differentiable map between manifolds. Then its derivative gives a differentiable map $Df : TM \to TM'$, $(p, v) \mapsto \big(f(p), Df_p(v)\big)$.*

Proof. This follows easily by looking at suitable charts. □

3.6. **Corollary.** *Every diffeomorphism $M \to M'$ extends to a diffeomorphism $TM \to TM'$ that induces linear isomorphisms on the fibers.*

Proof. Clear. □

The tangent bundle now allows us, for example, to define the notion of a vector field on M. Intuitively, a vector field gives us a tangent vector in every point of M, so it can be described by a map $M \to TM$. More precisely, it is a *section*.

3.7. **Definition.** Let $\pi : E \to M$ be a vector bundle. A *section* of E is a differentiable map $s : M \to E$ such that $\pi \circ s = \mathrm{id}_M$ (i.e., $s(p)$ lies in the fiber E_p above p, for all $p \in M$).

The section s *vanishes* at $p \in M$ if $s(p) = 0 \in E_p$ (recall that E_p is a vector space).

3.8. **Exercise.** A line bundle $E \to M$ is trivial if and only if it allows a non-vanishing section.

More generally, a rank n vector bundle $E \to M$ is trivial if and only if it allows n sections that are linearly independent at every point.

3.9. **Example.** The tangent bundle $T\mathbb{S}^1$ of the circle is a trivial line bundle. Indeed, we can realize it as

$$T\mathbb{S}^1 = \{(z, w) \in \mathbb{C}^2 : |z| = 1, \mathrm{Re}(\bar{z}w) = 0\} ,$$

and $z \mapsto (z, iz)$ is a non-vanishing section.

3.10. **Definition.** Let M be a manifold. A *vector field* on M is a section $M \to TM$.

3.11. **Example.** It can be shown that $T\mathbb{S}^2$ is a non-trivial vector bundle. (The proof is non-trivial, too!) Given this, we show that there is no non-vanishing vector field on the sphere $\mathbb{S}^2$: assume there was such a vector field $v : \mathbb{S}^2 \to T\mathbb{S}^2$ with $v(x) \neq 0$ for all $x \in \mathbb{S}^2$. Then $w(x) = x \times v(x)$ (vector product in $\mathbb{R}^3$) would give us a second vector field such that $v(x)$ and $w(x)$ were linearly independent for every $x \in \mathbb{S}^2$. But then we would have a trivialization $\mathbb{S}^2 \times \mathbb{R}^2 \to T\mathbb{S}^2$, $\big(x, (a, b)\big) \mapsto \big(x, av(x) + bw(x)\big)$, which, however, does not exist. Hence:

Every vector field on $\mathbb{S}^2$ must vanish somewhere.

("You cannot smoothly comb a hedgehog.")

3.12. **Constructions with Vector Bundles.** There are various ways in which to construct new vector bundles out of given ones. For example, if $\pi : E \to M$ is a vector bundle and $f : M' \to M$ is a differentiable map, then we can construct the *pull-back* of E to M', often denoted by f^*E. It is a vector bundle $\pi' : E' \to M'$ with a map $\tilde{f} : E' \to E$ such that

$$\begin{array}{ccc} E' & \xrightarrow{\tilde{f}} & E \\ {\scriptstyle \pi'}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ M' & \xrightarrow{f} & M \end{array}$$

commutes and gives isomorphisms $\tilde{f}_p : E'_p \to E_{f(p)}$ of the fibers. Furthermore, if E is trivial above $U \subset M$, then E' is trivial above $f^{-1}(U) \subset M'$, in a way compatible with the maps in the diagram above. (Exercise.)

There is a whole class of constructions of a different type coming from constructions of linear algebra. For example, let $\pi : E \to M$ and $\pi' : E' \to M$ be two vector bundles over M. Then we can form their *direct sum* $E \oplus E' \to M$ in the following way. Let $U \subset M$ be open such that both E and E' are trivial over U. (We can get a family of such open subsets that covers M by taking intersections of subsets over which E or E' is trivial, respectively.) Let the trivializations be $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^m$ and $\psi' : (\pi')^{-1}(U) \to U \times \mathbb{R}^n$. Then we take

$$E \oplus E' = \bigcup_{p \in M} \big(\{p\} \times (E_p \oplus E'_p)\big)$$

to be the total space and declare

$$\Psi : \bigcup_{p \in U} \big(\{p\} \times (E_p \oplus E'_p)\big) \longrightarrow U \times (\mathbb{R}^m \oplus \mathbb{R}^n)\,, \quad (p, v \oplus v') \longmapsto (p, w \oplus w')\,,$$

where $\psi(v) = (p, w)$ and $\psi'(v') = (p, w')$, to be a local trivialization of $E \oplus E'$. The compatibility condition is easily checked; this allows us to put a topology and differentiable structure on $E \oplus E'$.

In the same way, we can construct the *tensor product* $E \otimes E'$, the *dual bundle* E^* (whose fiber at p is the vector space dual to E_p), *tensor powers* $E^{\otimes k}$, *symmetric powers* $S^k E$ and *alternating* (or *exterior*) powers $\bigwedge^k E$. The latter will be important for us when discussing differential forms later.

3.13. **Example.** If $\pi : E \to M$ is a rank n vector bundle, then $\bigwedge^n E$ is a line bundle over M.

Even though $T\mathbb{S}^2$ is nontrivial, $\bigwedge^2 T\mathbb{S}^2$ is. Consider $T\mathbb{S}^2$ as a subvector bundle of $\mathbb{S}^2 \times \mathbb{R}^3$, via the realization of $T\mathbb{S}^2 \subset \mathbb{S}^2 \times \mathbb{R}^3$ coming from viewing $\mathbb{S}^2 \subset \mathbb{R}^3$ as a submanifold. Then the wedge product of two tangent vectors at $p \in \mathbb{S}^2$ can be identified with their vector product, which has its value in the bundle $E = \{(x, y) \in \mathbb{S}^2 \times \mathbb{R}^3 : y \in \mathbb{R}x\}$. So we see that $\bigwedge^2 T\mathbb{S}^2 \cong E$. But E admits a non-vanishing section $x \mapsto (x, x)$, so E is a trivial line bundle over $\mathbb{S}^2$.

4. Orientation and Orientability

You will remember from calculus that (for $f \in \mathcal{C}(\mathbb{R})$, say)

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx\,;$$

the value of the integral depends on the *orientation* of the interval.

Later, we will introduce two kinds of integrals on manifolds: one that is not sensitive to orientation (and can be used to compute volumes and the like), and one that is. In order to be able to deal with the second kind, we will need to orient our manifolds, and so we have to define what we mean by this.

4.1. **Orientations of $\mathbb{R}^n$.** The key fact is that $\mathrm{GL}(n, \mathbb{R})$, which gives us all the ways to go from one basis of $\mathbb{R}^n$ to another, has two connected components that are distinguished by the sign of the determinant (except when $n = 0$ — you can't orient points). We can therefore say that two bases of $\mathbb{R}^n$ have the *same* or *opposite orientation* if the determinant of the basis change matrix is positive or negative, respectively.

Since the determinant is defined intrinsically for automorphisms $V \to V$ of a finite-dimensional vector space, the notion generalizes to arbitrary finite-dimensional real vector spaces. We can then define an *orientation* of such a space V as a maximal set of bases of V that all have the same orientation. An automorphism $f : V \to V$ is then said to be *orientation-preserving* (or *orientation-reversing*) if it maps a basis to another basis of the same (or the opposite) orientation; this is the case if and only if $\det f$ is positive (or negative).

Note that $\mathbb{R}^n$ has a *canonical orientation*, which is given by the orientation that contains the canonical basis.

We can extend the notion of preserving or reversing orientation to differentiable maps $f : U \to \mathbb{R}^n$ where $U \subset \mathbb{R}^n$ is open: we say that f *preserves* (resp., *reverses*) *orientation* if $\det Df_x > 0$ (resp., $\det Df_x < 0$) for all $x \in U$.

We now extend this to manifolds.

4.2. **Definition.** Let M be a manifold. An *orientation* of M is a maximal atlas consisting of charts such that the transition maps Φ satisfy $\det D\Phi_x > 0$ for all x in the domain of definition of Φ. I.e., we require all transition maps to preserve orientation.

We say that a manifold is *orientable* if it has an orientation, and we say that M is *oriented* when we have chosen an orientation of M.

4.3. **Lemma.** *Let M be an oriented and connected manifold that is not just a point. Then M has exactly one other orientation than the given one.*

This other orientation is the *opposite orientation.*

Proof. Let $\mathcal{O}^+$ be the given orientation of M, and let $\mathcal{O}$ be another orientation. Let $\mathcal{O}^-$ be the orientation that consists of all charts in $\mathcal{O}^+$, post-composed with $(x_1, x_2, \ldots, x_k) \mapsto (-x_1, x_2, \ldots, x_k)$. Define subsets $A^+, A^- \subset M$ by saying that $p \in A^\pm$ if there is a neighborhood U of p such that $\mathcal{O}^\pm|_U$ and $\mathcal{O}|_U$ are the same orientation of U. Since all charts in $\mathcal{O}^-$ have orientation-reversing transition maps with all charts in $\mathcal{O}^+$, $A^+ \cap A^- = \emptyset$. On the other hand, let $\phi^+ : U \to V$ be a chart centered at $p \in M$, which is in $\mathcal{O}^+$, and let $\phi : U \to V'$ be a chart centered at p that is in $\mathcal{O}$. We can assume U (and hence V and V') to be connected. Let $\Phi : V \to V'$ be the transition map. Then $\det D\Phi : V \to \mathbb{R}^\times$ must either be always positive or always negative, which means that $\phi \in \mathcal{O}^+$ or $\phi \in \mathcal{O}^-$. This implies that $\mathcal{O}|_U = \mathcal{O}^+|_U$ or $\mathcal{O}|_U = \mathcal{O}^-|_U$, hence $U \subset A^+$ or $U \subset A^-$. This in turn implies that $A^+ \cup A^- = M$ and that A^+ and A^- are both open. Since M is connected, we must have either $A^+ = M$ (then $\mathcal{O} = \mathcal{O}^+$) or $A^- = M$ (then $\mathcal{O} = \mathcal{O}^-$). □

4.4. **Definition.** Let M be a manifold, $p \in M$. An *orientation of M near p* is an equivalence class of oriented manifolds $(U, \mathcal{O})$, where $p \in U \subset M$ is open, and $(U, \mathcal{O})$ and $(U', \mathcal{O}')$ are equivalent if $\mathcal{O}|_{U \cap U'} = \mathcal{O}'|_{U \cap U'}$.

Such local orientations always exist (compare the proof of the preceding lemma); this comes from the fact that open subsets of $\mathbb{R}^k$ inherit the canonical orientation of $\mathbb{R}^k$. This notion is important for the definition of the *orientable double cover* of M; this is a manifold M^* whose points correspond to pairs $(p, \mathcal{O})$, where $p \in M$ and $\mathcal{O}$ is an orientation of M near p. By the lemma above, the projection $M^* \to M$ (that forgets the local orientation) is a two-to-one map, and M^* carries a natural orientation. If M is connected, then M^* is connected if and only if M is *not* orientable (Exercise).

For example, the orientable double cover of the Möbius Strip is a cylinder; since the cylinder is connected, the Möbius Strip is not orientable. As another example, you may want to show that the orientable double cover of the real projective plane is $\mathbb{S}^2$, hence the real projective plane is not orientable, either. (Which is no surprise, since one can embed the Möbius Strip into it!)

Note that an orientation of M near p fixes an orientation of the tangent space T_pM. The manifold is orientable if we can orient all the tangent spaces in a consistent manner. We can call a vector bundle $E \to M$ *orientable* if we can consistently orient the fibers E_p.

4.5. **Lemma.** *A line bundle $L \to M$ is orientable if and only if it is trivial.*

Proof. If L is trivial, then $L \cong M \times \mathbb{R}$, and we can just transfer the canonical orientation of $\mathbb{R}$ to each of the fibers of L.

Conversely, assume that L is orientable; fix an orientation. Let (U_i) be an open covering of M such that $L|U_i$ is trivial. We will see later that there is a *partition of unity* subordinate to this covering; this is a family (ψ_i) of $\mathcal{C}^\infty$-functions on M such that $\psi_i(p) = 0$ if $p \notin U_i$, for any $p \in M$, only finitely many ψ_i are nonzero on a neighborhood of p, and $\sum_i \psi_i(p) = 1$ (the sum then makes sense). We have "oriented trivializations" $\phi_i : U_i \times \mathbb{R} \xrightarrow{\cong} L|_{U_i}$, meaning that the orientation of L corresponds to the canonical orientation of $\mathbb{R}$. We can then define a section $s : p \mapsto \sum_i \psi_i(p)\phi_i(p, 1)$ (where we set $\psi_i(p)\phi_i(p, 1) = 0$ for $p \notin U_i$). Since for a given $p \in M$, all $\phi_i(p, 1)$ for i such that $p \in U_i$ differ only by scaling with positive factors, we have that $s(p) \neq 0$. So we have a non-vanishing section of L, hence L is trivial. □

4.6. **Example.** Let $M \subset \mathbb{R}^n$ be an $(n-1)$-dimensional submanifold. Then we can define a line bundle $NM \to M$, the *normal bundle* of M as

$$NM = \{(p, v) \in \mathbb{R}^n \times \mathbb{R}^n : p \in M, v \perp T_pM\}$$

(this works for submanifolds of any dimension; in general it is a vector bundle of rank $n - \dim M$). Then M is orientable if and only if NM is a trivial line bundle.

To see this, note that the latter condition is equivalent to saying that there is a section $n : M \to NM$ such that $\|n(p)\| = 1$ for all $p \in M$ (there must be a non-vanishing section, which we can then normalize). Such a section n defines a smoothly varying *unit normal vector* to M. Now note that a nonzero vector $y \in (T_pM)^\perp$ defines an orientation on T_pM by taking as oriented bases $b_1, \ldots, b_{n-1}$ those that when extended by y give a canonically oriented basis of $\mathbb{R}^n$. A section n as above therefore provides us with a consistent orientation of the tangent spaces, and conversely.

4.7. **Proposition.** *A manifold M is orientable if and only if $\bigwedge^{\dim M} TM$ is a trivial line bundle over M.*

Proof. Let $k = \dim M$. An orientation of T_pM defines an orientation of $\bigwedge^k T_pM$ (given by the set of all $v_1 \wedge \cdots \wedge v_k$ where $v_1, \ldots, v_k$ is an oriented basis of T_pM), and conversely. So M is orientable if and only if $\bigwedge^k TM$ is an orientable line bundle if and only if $\bigwedge^k TM$ is a trivial line bundle. □

4.8. **Example.** We can now see in two different ways that the sphere $\mathbb{S}^2$ is orientable. On the one hand, there is the obvious "outer unit normal field" given by $x \mapsto (x, x)$, so $\mathbb{S}^2$ is orientable by Example 4.6. On the other hand, we have seen in Example 3.13 that $\bigwedge^2 T\mathbb{S}^2$ is trivial, so $\mathbb{S}^2$ is orientable by Prop. 4.7.

In fact, both ways are really the same, as

$$NM \cong \operatorname{Hom}(\textstyle\bigwedge^{n-1} TM, M \times \bigwedge^n \mathbb{R}^n) \cong \bigwedge^{n-1} TM^*$$

in the situation of Example 4.6. (If U and U' are complementary subspaces of a vector space V, with $\dim U = k$, $\dim U' = n - k$ and $\dim V = n$, then the wedge product gives a bilinear map

$$\textstyle\bigwedge^k U \times \bigwedge^{n-k} U' \to \bigwedge^n V$$

which induces an isomorphism $\bigwedge^{n-k} U' \cong \operatorname{Hom}\big(\bigwedge^k U, \bigwedge^n V\big)$.)

5. Partitions of Unity

We have already seen a *partition of unity* in the proof of Lemma 4.5. Such a partition of unity is very useful if we want to split a global object into local ones in a smooth way, or conversely, if we want to glue together local objects to a global one in a smooth way, like in the proof mentioned above. In this section, we will prove that partitions of unity exist on manifolds. We follow [W].

5.1. **Definition.** Let $(U_i)_{i\in I}$ and $(V_j)_{j\in J}$ be open covers of a topological space X. We say that $(V_j)_{j\in J}$ is a *refinement* of $(U_i)_{i\in I}$ if for every $j \in J$ there is some $i \in I$ such that $V_j \subset U_i$.

Let $(A_i)_{i\in I}$ be a family of (not necessarily open) subsets of X. We say that $(A_i)_{i\in I}$ is *locally finite* if every point $p \in X$ has a neighborhood U such that $U \cap A_i \neq \emptyset$ only for finitely many $i \in I$.

5.2. **Definition.** Let M be a manifold. A *partition of unity* on M is a family of nonnegative $\mathcal{C}^\infty$-functions $(\psi_j)_{j\in J}$ on M such that $(\operatorname{supp}\psi_j)_{j\in J}$ is locally finite and $\sum_{j\in J}\psi_j(p) = 1$ for all $p \in M$. (The sum makes sense, since there are only finitely many non-zero terms.)

The partition of unity is *subordinate* to an open cover $(U_i)_{i\in I}$ of M if for every $j \in J$ there is some $i \in I$ such that $\operatorname{supp}\psi_j \subset U_i$. It is *subordinate with the same index set* if $I = J$ and $\operatorname{supp}\psi_i \subset U_i$ for all $i \in I = J$.

5.3. **Lemma.** *Let M be a manifold. Then there is a sequence $(W_n)_{n\ge1}$ of open subsets of M such that*

$$\overline{W_n} \text{ is compact}, \qquad \overline{W_n} \subset W_{n+1}, \qquad \text{and} \qquad M = \bigcup_{n=1}^{\infty} W_n\,.$$

Proof. In fact, we only need that M is Hausdorff, locally compact (i.e., every point as a compact neighborhood), and has a countable basis of the topology. Let $(B_m)_{m\ge1}$ be a countable basis of the topology of M; we can assume that $\overline{B_m}$ is compact for all m. (Just take any countable basis and remove all sets that do not satisfy the condition. The resulting collection will still be a basis, since M is locally compact and Hausdorff: Let $U \subset M$ be open and $p \in U$; let V be an open neighborhood of p with compact closure. Then there is a basis set B with $p \in B \subset U \cap V$, and the closure of B is a closed subset of $\overline{V}$, which is compact, hence $\overline{B}$ is again compact.) Now let $W_1 = B_1$ and recursively define an increasing sequence $(m_n)_{n\ge1}$ such that $W_n = \bigcup_{m=1}^{m_n} B_m$ by taking m_{n+1} to be the smallest integer $> m_n$ with $\overline{W_n} \subset \bigcup_{m=1}^{m_{n+1}} B_m$. This defines a sequence (W_n) with the desired properties. Note that $\overline{W_n} \subset \bigcup_{m=1}^{m_n} \overline{B_m}$ is a closed subset of a finite union of compact sets, hence compact. □

5.4. **Lemma.** *There exists a $\mathcal{C}^\infty$-function h_k on $\mathbb{R}^k$ with $0 \le h_k \le 1$, $h_k(x) = 1$ for $\|x\| \le 1$, and $h_k(x) = 0$ for $\|x\| \ge 2$.*

Proof. We start with the function

$$f : \mathbb{R} \longrightarrow \mathbb{R}\,, \qquad t \longmapsto \begin{cases} e^{-1/t} & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases}$$

It is a standard fact (or an exercise) that f is $\mathcal{C}^\infty$; $f(t)$ is positive for $t > 0$. Then

$$g(t) = \frac{f(t)}{f(t) + f(1-t)}$$

is again $\mathcal{C}^\infty$; it is zero for $t \le 0$ and $g(t) = 1$ for $t \ge 1$. We can then define

$$h_k(x) = g(4 - \|x\|^2)\,;$$

this satisfies the requirements. □

5.5. **Theorem (Existence of Partitions of Unity).** *Let M be a manifold, $(U_i)_{i \in I}$ an open cover of M.*

(1) *There exists a countable partition of unity $(\psi_j)_{j \in J}$ of M subordinate to $(U_i)_{i \in I}$ and such that $\operatorname{supp} \psi_j$ is compact for all $j \in J$.*
(2) *There exists a partition of unity $(\psi'_i)_{i \in I}$ subordinate to $(U_i)_{i \in I}$ with the same index set and with at most countably many of the ψ'_i not identically zero.*

Note that in the second case, we cannot require that ψ_i has compact support. Consider for example $M = \mathbb{R}$ with the one-element open cover given by $\mathbb{R}$ itself.

Proof. Let $(W_n)_{n \geq 1}$ be a sequence of open sets as in Lemma 5.3; let $W_0 = \emptyset$. For $p \in M$, let n_p be the largest integer such that $p \notin \overline{W_{n_p}}$. Pick i_p such that $p \in U_{i_p}$, and let ϕ_p be a chart centered at p whose domain is contained in $U_{i_p} \cap (W_{n_p+2} \setminus \overline{W_{n_p}})$ and whose codomain contains the closed ball of radius 2 in $\mathbb{R}^k$, where $k = \dim M$. Let h_k be the function from Lemma 5.4, and define $\alpha_p(x) = h_k(\phi_p(x))$ for x in the domain of ϕ_p and $\alpha_p(x) = 0$ else. Then $\alpha_p \in \mathcal{C}^\infty(M)$, α_p has compact support contained in $U_{i_p} \cap (W_{n_p+2} \setminus \overline{W_{n_p}})$ and takes the value 1 on some open neighborhood V_p of p.

For each $n \geq 1$, choose a finite set of points $p \in M$ such that the V_p cover $\overline{W_n} \setminus W_{n-1}$, and let J be the union of these finite sets. Then $(\alpha_j)_{j \in J}$ is a countable family of functions, whose supports form a locally finite family of subsets, so $\alpha = \sum_{j \in J} \alpha_j$ is a well-defined $\mathcal{C}^\infty$-function on M, and $\alpha \geq 1$ everywhere. We define $\psi_j = \alpha_j / \alpha$; this provides a partition of unity with compact supports and subordinate to $(U_i)_{i \in I}$.

For the second statement, fix for every $j \in J$ an $i_j \in I$ such that $\operatorname{supp} \psi_j \subset U_{i_j}$, and define

$$\psi'_i = \sum_{j : i_j = i} \psi_j \,.$$

Then $\operatorname{supp} \psi'_i \subset U_i$ (note that the union of a locally finite family of closed sets is closed), $\sum_{i \in I} \psi'_i = \sum_{j \in J} \psi_j = 1$, and ψ'_i is not the zero function only for the countably many $i \in I$ of the form i_j. □

The following is sometimes useful.

5.6. **Corollary.** *Let M be a manifold, $A \subset U \subset M$ with A closed and U open. Then there is $h \in \mathcal{C}^\infty(M)$ with $h|_A = 1$, $h|_{M \setminus U} = 0$, and $0 \leq h \leq 1$.*

Proof. There is a partition of unity $(\psi_U, \psi_{M \setminus A})$ subordinate to the open cover $(U, M \setminus A)$ of M with the same index set. We can then take $h = \psi_U$. □

6. Volume Integrals on Submanifolds

We now want to discuss how to define volumes of submanifolds of $\mathbb{R}^n$, or more generally, integrals of functions on submanifolds with respect to volume. The reason for looking at this case first is that the ambient space $\mathbb{R}^n$ gives us a natural notion of volume for k-dimensional parallelotopes; we can then use this to approximate the volume of the manifold.

6.1. **Reminder: "Flat" Volume Integrals.** Recall the "unoriented (Riemann) integral" of a sufficiently nice function $f : \mathbb{R}^k \to \mathbb{R}$ (e.g., f continuous) with compact support,

$$\int_{\mathbb{R}^k} f(x)\,|d^k x| = \lim_{N\to\infty} 2^{-kN} \sum_{\mathcal{C}\in\mathcal{D}_N(\mathbb{R}^k)} f(x_{\mathcal{C}})\,,$$

where $\mathcal{D}_N(\mathbb{R}^k)$ is the set of "lattice cubes" of side-length 2^{-N} in $\mathbb{R}^k$ and $x_{\mathcal{C}} \in \mathcal{C}$ is some choice of representative point. The limit exists and does not depend on the choice of the $x_{\mathcal{C}}$ if f is Riemann integrable (which is what we mean by 'nice').

The idea was to cut $\mathbb{R}^k$ (or the support of f) into little pieces (which are chosen to be cubes here), estimate the integral on each little piece by its volume times a representative value of f, and sum the numbers thus obtained.

6.2. **Submanifolds.** Now consider a k-dimensional submanifold $M \subset \mathbb{R}^n$. For simplicity, assume that M is described by a single chart $\phi : M \to V \subset \mathbb{R}^k$. Let $f : M \to \mathbb{R}$ be continuous (say) with compact support. Then $f \circ \phi^{-1} : V \to \mathbb{R}$ is continuous with compact support, and we can extend it (by zero) to a continuous function on all of $\mathbb{R}^k$. In order to approximate $\int_M f(x)\,|d^k x|$, we can use the chart. We cut $\mathbb{R}^k$ into pieces $\mathcal{C}$ as before and try to approximate the integral of f over the image in M of each little cube. Since locally ϕ^{-1} is very close to a linear map, the little cube $\mathcal{C}$ will map to something very well approximated by a parallelotope; this parallelotope is spanned by 2^{-N} times the partial derivatives of ϕ^{-1}. So we would like to take

$$\int_M f(x)\,|d^k x| = \lim_{N\to\infty} 2^{-kN} \sum_{\mathcal{C}\in\mathcal{D}_N(\mathbb{R}^k)} f\big(\phi^{-1}(x_{\mathcal{C}})\big)\,\mathrm{vol}_k\big(P(D(\phi^{-1})_{x_{\mathcal{C}}})\big)\,.$$

Here, $P(D(\phi^{-1})_x)$ denotes the parallelotope spanned by the columns of the matrix $D(\phi^{-1})_x$, and $x_{\mathcal{C}} \in \mathcal{C}$ is, as before, some choice of representative point.

6.3. **Volumes of Parallelotopes.** In order to turn this into something useful, we need to know how to compute the k-dimensional volume of a k-dimensional parallelotope in $\mathbb{R}^n$. For this, we recall that the volume of the parallelotope spanned by the columns of a $k \times k$ matrix A is given by $|\det A|$. This is not immediately useful, since the matrix $D(\phi^{-1})_x$ will not be a square matrix in general. However, note that we can also write the volume above as $\sqrt{\det(A^\top A)}$ (since $\det(A^\top A) = (\det A)^2$). Now if $a_1, \ldots, a_k$ are the columns of A, the (i,j)-entry of $A^\top A$ is the inner product $a_i \cdot a_j$. This means that the entries only depend on the lengths of the a_i and the angles between them. (Recall that $a \cdot a = \|a\|^2$ and $a \cdot b = \|a\|\,\|b\| \cos\alpha$, where $0 \le \alpha \le \pi$ is the angle between a and b.)

This is now something we can carry over to k-parallelotopes in $\mathbb{R}^n$: The k-dimensional volume of the parallelotope spanned in $\mathbb{R}^n$ by the columns of the $n \times k$ matrix T is

$$\mathrm{vol}_k P(T) = \sqrt{\det(T^\top T)}\,.$$

6.4. **Definition.** In the situation considered in 6.2 above, we define

$$\int_M f(x)\,|d^k x| = \int_V f\big(\phi^{-1}(x)\big)\,\sqrt{\det\big(D(\phi^{-1})_x^\top D(\phi^{-1})_x\big)}\,|d^k x|\,.$$

More generally, we can define such an integral whenever f has compact support contained in the domain of a chart of M.

We need to check that this definition does not depend on the chart. So let f have compact support contained in the domains of two charts ϕ and ϕ', and let $\Phi = \phi' \circ \phi^{-1}$ be the transition map. We can assume that $\phi : U \to V$ and $\phi' : U \to V'$ have the same domain; then $\Phi : V \to V'$, and the change-of-variables formula gives (note that $\phi^{-1} = (\phi')^{-1} \circ \Phi$):

$$\begin{aligned}
&\int_V f\big(\phi^{-1}(x)\big) \sqrt{\det\big(D(\phi^{-1})_x^\top D(\phi^{-1})_x\big)}\, |d^k x| \\
&\quad = \int_V f\big((\phi')^{-1}(\Phi(x))\big) \sqrt{\det\big((D((\phi')^{-1})_{\Phi(x)} D\Phi_x)^\top D((\phi')^{-1})_{\Phi(x)} D\Phi_x\big)}\, |d^k x| \\
&\quad = \int_V f\big((\phi')^{-1}(\Phi(x))\big) \sqrt{\det\big(D((\phi')^{-1})_{\Phi(x)}^\top D((\phi')^{-1})_{\Phi(x)}}\, |\det D\Phi_x|\, |d^k x| \\
&\quad = \int_{V'} f\big((\phi')^{-1}(y)\big) \sqrt{\det\big(D((\phi')^{-1})_y^\top D((\phi')^{-1})_y}\, |d^k y|
\end{aligned}$$

So we get the same result from both charts.

We now generalize to functions whose support is not necessarily contained in the domain of a chart. The tool we use for this is a partition of unity.

6.5. **Definition.** Let $M \subset \mathbb{R}^n$ be a k-dimensional submanifold, and let $f : M \to \mathbb{R}$ be a function. Let $(\psi_j)_j$ be a countable partition of unity on M such that each ψ_j has compact support contained in the domain of a chart of M. Then $\psi_j f$ is a function with compact support contained in the domain of a chart $\phi_j : U_j \to V_j$. We say that f is *Riemann integrable* if for each j, the function

$$V_j \longrightarrow \mathbb{R}, \quad x \longmapsto \psi_j\big(\phi_j^{-1}(x)\big) f\big(\phi_j^{-1}(x)\big) \sqrt{\det\big(D(\phi^{-1})_x^\top D(\phi^{-1})_x\big)}$$

is Riemann integrable, and $\sum_j \int_M |\psi_j(x) f(x)|\, |d^k x|$ converges.

In this case,

$$\int_M f(x)\, |d^k x| = \sum_j \int_M \psi_j(x) f(x)\, |d^k x| \in \mathbb{R}$$

exists and is defined to be the *(Riemann) integral* of f over M.

If f is continuous with compact support, then f is Riemann integrable over M. In particular, if f is continuous and M is compact, then f is Riemann integrable on M. If the constant function 1 is Riemann integrable on M, we call its integral $\int_M |d^k x|$ the *(k-dimensional) volume* of M.

We have to check that this definition does not depend on the partition of unity we choose. First, let $(\Psi_{j,\ell})$ be a refinement of (ψ_j); this is a partition of unity such that for each j, only finitely many $\Psi_{j,\ell}$ are not identically zero, and such that $\psi_j = \sum_\ell \Psi_{j,\ell}$. Since we can take a common chart for ψ_j and all the $\Psi_{j,\ell}$, it is clear that we get the same result with both partitions of unity.

Now consider two arbitrary partitions of unity (ψ_j) and (ψ'_ℓ) satisfying the assumptions of the definition above. Then $\Psi_{j,\ell} = \psi_j \psi'_\ell$ gives rise to a common refinement of both (reversing the indices j and ℓ to view it as a refinement of (ψ'_ℓ)), so by the argument given in the preceding paragraph both partitions give the same result.

In practice, when you want to do computations, it is not very convenient to work with partitions of unity. If we have a chart (or finitely many charts) that covers "most" of M (that have disjoint domains covering "most" of M), where "most" means everything except a subset of k-dimensional volume zero, then we can use this chart (these charts) to compute volumes and integrals.

6.6. **Example: Length of a Curve.** As a first example, take a curve (one-dimensional submanifold) in $\mathbb{R}^n$, parametrized by a smooth map $\gamma : I \to \mathbb{R}^n$, where I is an open interval. If $a, b \in I$, $a < b$, then the length of the part of the curve between $\gamma(a)$ and $\gamma(b)$ is given by

$$\int_a^b \|\gamma'(t)\| \, dt \,.$$

For example, if we consider the unit circle parametrized by $\gamma(t) = (\cos t, \sin t)$, then $\|\gamma'(t)\| = 1$, and we find that the length of the circle is 2π.

If we look at a logarithmic spiral $\gamma(t) = (e^{ct}\cos t, e^{ct}\sin t)$ (with $c > 0$), then we have

$$\|\gamma'(t)\| = \sqrt{\big(e^{ct}(c\cos t - \sin t)\big)^2 + \big(e^{ct}(c\sin t + \cos t)\big)^2} = e^{ct}\sqrt{c^2+1}\,.$$

We obtain for the length between $\gamma(a)$ and $\gamma(b)$ the result

$$\ell_{a,b} = \sqrt{c^2+1}\int_a^b e^{ct}\,dt = \frac{\sqrt{c^2+1}}{c}(e^{bc} - e^{ac})\,.$$

Note that this has a limit for $a \to -\infty$ and b fixed.

6.7. **Example: Area of Sphere.** Now consider the sphere $\mathbb{S}^2$. If we remove one 'meridian', we can parametrize it by spherical coordinates:

$$F : \,]0, 2\pi[\times]-\tfrac{\pi}{2}, \tfrac{\pi}{2}[\longrightarrow \mathbb{R}^3\,, \quad (t, u) \longmapsto (\cos t\cos u, \sin t\cos u, \sin u)\,.$$

The derivative of F is

$$DF_{(t,u)} = \begin{pmatrix} -\sin t\cos u & -\cos t\sin u \\ \cos t\cos u & -\sin t\sin u \\ 0 & \cos u \end{pmatrix},$$

so

$$DF_{(t,u)}^\top DF_{(t,u)} = \begin{pmatrix} \cos^2 u & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\sqrt{\det\big(DF_{t,u}^\top DF_{(t,u)}\big)} = \cos u\,.$$

(Note that $\cos u > 0$.) Therefore, the area of the sphere comes out as

$$\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} \cos u\,du\,dt = \int_0^{2\pi} 2\,dt = 4\pi\,.$$

6.8. **Example: Higher-Dimensional Spheres.** We can parametrize $\mathbb{S}^n$ in a similar way by

$$F : \left]0, 2\pi\right[\times \left]\tfrac{-\pi}{2}, \tfrac{\pi}{2}\right[^{n-1} \longrightarrow \mathbb{R}^n ,$$

$$(u_1, \cdots u_n) \longmapsto \begin{pmatrix} \cos u_1 \cos u_2 \cos u_3 \cdots \cos u_n \\ \sin u_1 \cos u_2 \cos u_3 \cdots \cos u_n \\ \sin u_2 \cos u_3 \cdots \cos u_n \\ \vdots \\ \sin u_{n-1} \cos u_n \\ \sin u_n \end{pmatrix}$$

We then find that

$$\sqrt{\det\bigl(DF^\top DF\bigr)} = \cos u_2 \cos^2 u_3 \cdots \cos^{n-1} u_n .$$

Therefore, the volume of $\mathbb{S}^n$ is

$$2\pi \int_{-\pi/2}^{\pi/2} \cos u_2 \, du_2 \int_{-\pi/2}^{\pi/2} \cos^2 u_3 \, du_3 \cdots \int_{-\pi/2}^{\pi/2} \cos^{n-1} u_n \, du_n = 2a_0 a_1 \dots a_{n-1} ,$$

where $a_k = \int_{-\pi/2}^{\pi/2} \cos^k t \, dt$. We have the recurrence

$$a_0 = \pi , \quad a_1 = 2 , \quad a_{k+2} = \frac{k+1}{k+2} a_k$$

(obtained via integration by parts). From this, we obtain by induction that

$$a_{k+1} a_k = \frac{2\pi}{k+1} .$$

Hence, by induction again,

$$\operatorname{vol} \mathbb{S}^{2n+1} = 2\pi (a_1 a_2) \cdots (a_{2n-1} a_{2n}) = \frac{2\pi^{n+1}}{n!}$$

and

$$\operatorname{vol} \mathbb{S}^{2n} = 2(a_0 a_1) \cdots (a_{2n-2} a_{2n-1}) = \frac{2^{n+1} \pi^n}{(2n-1)(2n-3) \cdots 3 \cdot 1} .$$

With the convention that $(-\frac{1}{2})! = \sqrt{\pi}$ and the usual recurrence for factorials, these can be combined into the single formula

$$\operatorname{vol} \mathbb{S}^n = \frac{2\pi^{(n+1)/2}}{\bigl(\frac{n-1}{2}\bigr)!} .$$

6.9. **Example (Möbius Strip).** As another example, consider the Möbius Strip M embedded in $\mathbb{R}^3$ as the image of the map

$$F : \mathbb{R} \times \left]-1, 1\right[\longrightarrow \mathbb{R}^3 , \quad (t, x) \longmapsto \bigl((1 + x \cos t) \cos 2t, (1 + x \cos t) \sin 2t, x \sin t\bigr) .$$

The inverse of F, after restricting it to $\left]0, \pi\right[\times \left]-1, 1\right[$ is a chart of M that covers all of M except the image of $\{0\} \times \left]-1, 1\right[$ under F (which is a line segment in $\mathbb{R}^3$). Using another chart (restrict F to $\left]-\varepsilon, \varepsilon\right[\times \left]-1, 1\right[$), we see that its characteristic function is integrable, with integral zero. So the volume (= area) of M is not changed when we remove this line segment. We can then compute the volume as

$$\begin{aligned}\mathrm{vol}_2(M) &= \int\limits_{]0,\pi[\times]-1,1[} \sqrt{\det(DF_z^\top DF_z)}\,|d^2z| \\ &= \int\limits_0^\pi \int\limits_{-1}^1 \sqrt{x^2+4(1+x\cos t)^2}\,dx\,dt \\ &\approx 13.254 > 4\pi\,.\end{aligned}$$

(The integral apparently cannot be evaluated in closed form.)

6.10. **Example (Graphs).** Consider an open subset $U \subset \mathbb{R}^k$ and a smooth function $f : U \to \mathbb{R}$. Then the graph of f is $\Gamma(f) = F(U)$, where

$$F : U \longrightarrow \mathbb{R}^{k+1}\,, x \longmapsto \big(x, f(x)\big)\,;$$

it is a k-dimensional submanifold of $\mathbb{R}^{k+1}$. The derivative of F is a $(k+1)\times k$ matrix whose upper k rows form an identity matrix, and whose last row is the gradient ∇f of f. Therefore, $DF^\top \cdot DF = (\nabla f)^\top \nabla f + I_k$, and $\det(DF^\top DF) = 1 + \|\nabla f\|^2$ (this is a linear algebra exercise). So finally,

$$\mathrm{vol}_k\,\Gamma(f) = \int\limits_U \sqrt{1+\|\nabla f_x\|^2}\,|d^kx|\,.$$

7. The Formalism of Differential Forms on $\mathbb{R}^n$

Our main goal in the following is to prove the general version of *Stokes' Theorem:*

$$\int\limits_M d\omega = \int\limits_{\partial M} \omega\,.$$

In order to do this, we first have to explain the various objects in this formula.

The main thing to be defined is the notion of a differential k-form, which is the type of thing ω is; it is the kind of object that naturally can be integrated over a k-dimensional domain.

7.1. **Definition.** A *k-form* on a real vector space V is an alternating multilinear map $V^k \to \mathbb{R}$.

In more down-to-earth terms, $\omega : V^k \to \mathbb{R}$ is linear in each of its arguments separately (this is the meaning of 'multilinear') and $\omega(v_1, \dots, v_k) = 0$ if $v_1, \dots, v_k \in V$ are linearly dependent (this is equivalent to 'alternating', given that ω is multilinear; 'alternating' by definition means that ω vanishes if two of its arguments are equal).

If $V = \mathbb{R}^n$, we can write down some k-forms. Namely, let $i_1, \dots, i_k \in \{1, 2, \dots, n\}$. Then we define the k-form $dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$ by

$$\big(dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}\big)\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1}\end{pmatrix}, \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2}\end{pmatrix}, \dots, \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk}\end{pmatrix}\right) = \det\begin{pmatrix} a_{i_11} & \cdots & a_{i_1k} \\ a_{i_21} & \cdots & a_{i_2k} \\ \vdots & & \vdots \\ a_{i_k1} & \cdots & a_{i_kk}\end{pmatrix}$$

That this *is* a k-form follows from the fact that the determinant is multilinear and alternating. Considering the determinant as a function of the rows, we also see that $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ is zero when two of the indices are equal and changes sign when two of them are swapped. Therefore, we can restrict to $1 \le i_1 < i_2 < \cdots < i_k \le n$. It can be shown that the set of k-forms obtained in this way is a basis of the space of all k-forms on $\mathbb{R}^n$, which therefore has dimension

$$\dim\left(\bigwedge^k \mathbb{R}^n\right)^* = \binom{n}{k}.$$

For example, up to scaling, there is only one n-form $dx_1 \wedge \cdots \wedge dx_n$, and it is given by the determinant of the matrix obtained from the entries of the n vectors.

7.2. **Definition.** Let ω be a k-form and η an l-form on $\mathbb{R}^n$. Then we define their *wedge product* as

$$\begin{aligned}(\omega \wedge \eta)(v_1, \ldots, v_{k+l}) &= \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} \varepsilon(\sigma)\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}) \\ &= \sum_{\sigma \in S_{k,l}} \varepsilon(\sigma)\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}) .\end{aligned}$$

Here $\varepsilon(\sigma)$ is the sign of the permutation σ, and $S_{k,l}$ denotes the set of (k,l)-*shuffles*; these are permutations $\sigma \in S_{k+l}$ such that $\sigma(1) < \cdots < \sigma(k)$ and $\sigma(k+1) < \cdots < \sigma(k+l)$. (Exercise: prove the second equality above.)

$\omega \wedge \eta$ is then a $(k+l)$-form on $\mathbb{R}^n$.

7.3. **Remark.** *The wedge product is associative and commutative up to sign:*

$$(\omega_1 \wedge \omega_2) \wedge \omega_3 = \omega_1 \wedge (\omega_2 \wedge \omega_3) =: \omega_1 \wedge \omega_2 \wedge \omega_3 \quad \textit{and} \quad \eta \wedge \omega = (-1)^{kl} \omega \wedge \eta$$

when ω is a k-form and η is an l-form.

Proof. Exercise. □

Note also that this notation is compatible with the notation $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ used earlier: this k-form really is the wedge product of $dx_{i_1}, \ldots, dx_{i_k}$.

What we will be interested in in the following are not just k-forms on $\mathbb{R}^n$, but k-forms that depend smoothly on a point $x \in U$, where $U \subset \mathbb{R}^n$ is an open set. (And later, we will generalize this notion to manifolds.)

7.4. **Definition.** Let $U \subset \mathbb{R}^n$ be open. A *differential k-form on U* is a smooth map $\omega : U \ni x \mapsto \omega_x \in (\bigwedge^k \mathbb{R}^n)^*$ that associates to every $x \in U$ a k-form ω_x on $\mathbb{R}^n$. In down-to-earth terms, this means that

$$\omega_x = \sum_{1 \le i_1 < \cdots < i_k \le n} f_{i_1,\ldots,i_k}(x)\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

with $f_{i_1,\ldots,i_k} \in \mathcal{C}^\infty(U)$.

We will usually just speak of "k-forms" when we mean "differential k-forms" and hope that this will not lead to confusion.

We define the wedge product of two differential forms on U point-wise:

$$(\omega \wedge \eta)_x = \omega_x \wedge \eta_x .$$

7.5. **Example.** On $\mathbb{R}^3$, there are

- 0-forms; these are just smooth functions;
- 1-forms $f\,dx + g\,dy + h\,dz$; they correspond to vector fields $\mathbf{F} = (f, g, h)$;
- 2-forms $s\,dy \wedge dz - t\,dx \wedge dz + u\,dx \wedge dy$; we can let them correspond to vector fields $\mathbf{U} = (s, t, u)$ again;
- 3-forms $r\,dx \wedge dy \wedge dz$, given by the function r.

However, one has to be careful with these identifications, since the transformation behavior is quite different — a 1-form really gives a *cotangent* vector at every point and therefore transforms differently under diffeomorphisms than a vector field, which gives a *tangent* vector at every point, and 2-forms transform still differently from that. Also, 3-forms transform differently than functions. See below where we introduce the pull-back of a differential form. On $\mathbb{R}^3$, we can make these identifications, because we have a fixed euclidean structure (i.e., we have a canonical inner product).

We need some more ingredients in order to understand the formula giving Stokes' Theorem. One of them is the exterior derivative of k-forms.

7.6. **Definition.** Let f be a 0-form on $U \subset \mathbb{R}^n$ (i.e., a smooth function). Then we define the 1-form df by

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dx_i\,.$$

If $\omega = \sum_{i_1<\cdots<i_k} f_{i_1,\dots,i_k}\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ is a k-form on U, then we define the $(k+1)$-form $d\omega$ by

$$d\omega = \sum_{i_1<\cdots<i_k} df_{i_1,\dots,i_k} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

(where $df_{i_1,\dots,i_k}$ is defined as above.)

7.7. **Remark.** df encodes the directional derivatives of f:

$$df_x(v) = \lim_{h\to 0} \frac{f(x+hv) - f(x)}{h}\,.$$

One can interpret $d\omega$ in a similar way, compare [HH].

7.8. **Examples.** Let us go back to $\mathbb{R}^3$ and see what the exterior derivative means for the various kinds of forms.

- If f is a function (0-form), then the vector field corresponding to df is just the *gradient* ∇f of f.

- If $\mathbf{F}$ is the vector field corresponding to the 1-form $\omega = f\,dx + g\,dy + h\,gz$, then we obtain

$$
\begin{aligned}
d\omega &= df \wedge dx + dg \wedge dy + dh \wedge dz \\
&= \frac{\partial f}{\partial x}\, dx \wedge dx + \frac{\partial f}{\partial y}\, dy \wedge dx + \frac{\partial f}{\partial z}\, dz \wedge dx \\
&\quad + \frac{\partial g}{\partial x}\, dx \wedge dy + \frac{\partial g}{\partial y}\, dy \wedge dy + \frac{\partial g}{\partial z}\, dz \wedge dy \\
&\quad + \frac{\partial h}{\partial x}\, dx \wedge dz + \frac{\partial h}{\partial y}\, dy \wedge dz + \frac{\partial h}{\partial z}\, dz \wedge dz \\
&= \Big(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\Big)\, dy \wedge dz - \Big(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}\Big)\, dx \wedge dz + \Big(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\Big)\, dx \wedge dy
\end{aligned}
$$

  (note that $dx \wedge dx = 0$, $dy \wedge dx = -\,dx \wedge dy$ etc.). The vector field corresponding to that is the *curl* $\nabla \times \mathbf{F}$ of $\mathbf{F}$.
- If $\mathbf{U}$ is the vector field corresponding to the 2-form

$$\eta = s\,dy \wedge dz - t\,dx \wedge dz + u\,dx \wedge dy\,,$$

  then $d\eta$ is given by

$$d\eta = ds \wedge dy \wedge dz - dt \wedge dx \wedge dz + du \wedge dx \wedge dy = \Big(\frac{\partial s}{\partial x} + \frac{\partial t}{\partial y} + \frac{\partial u}{\partial z}\Big)\, dx \wedge dy \wedge dz$$

  and corresponds to the *divergence* $\nabla \cdot \mathbf{U}$ of $\mathbf{U}$.

So we see that all these different operations from vector analysis really are just different incarnations of the same uniform principle. (This correspondence also explains the perhaps at first sight non-obvious way we did the identification of 2-forms and vector fields.)

7.9. **Remark.** *We have a modified Leibniz rule for the exterior derivative of a wedge product, and taking the exterior derivative twice gives zero:*

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta \quad \text{and} \quad dd\omega = 0\,,$$

where ω is a k-form.

Proof. Exercise. □

Also note that our notation dx_i is compatible with our general definition, if we think of x_i as the ith coordinate function on $\mathbb{R}^n$.

What is the purpose of a k-form? It wants to be integrated over some k-dimensional domain. Let us first do it for n-forms.

7.10. **Definition.** Let $\omega = f(x)\, dx_1 \wedge \cdots \wedge dx_n$ be an n-form on an open subset $U \subset \mathbb{R}^n$. Then we define

$$\int_U \omega = \int_U f(x)\, |d^n x|\,,$$

where the integral on the right is the multiple integral known from Analysis II.

Note that our integral is an *oriented* integral: if we change orientation by switching two of the coordinates, then ω (i.e., f) changes sign, and so does the integral.

Before we can define integrals of k-forms, we need to look at how k-forms should transform under diffeomorphisms, or more generally, smooth maps.

7.11. **Definition.** Let $\phi : U \to V$ be a smooth map, where $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ are open, and let

$$\omega = \sum_{i_1<\cdots<i_k} f_{i_1,\ldots,i_k}\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

be a k-form on V. Then we define the *pull-back* of ω to U as

$$\phi^*\omega = \sum_{i_1<\cdots<i_k} (f_{i_1,\ldots,i_k} \circ \phi)\, d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} .$$

Here $\phi = (\phi_1, \ldots, \phi_n)$, and $d\phi_i$ is the exterior derivative as defined above.

7.12. **Remark.** *Pull-back is compatible with exterior derivative and wedge product:*

$$\phi^*(d\omega) = d\phi^*\omega \quad \text{and} \quad \phi^*(\omega \wedge \eta) = \phi^*\omega \wedge \phi^*\eta .$$

If $\psi : W \to U$ is another smooth map, then $(\phi \circ \psi)^*\omega = \psi^*(\phi^*\omega)$.

Proof. Exercise. □

7.13. **Example.** If $\omega = f\, dx_1 \wedge \cdots \wedge dx_n$ is an n-form on $V \subset \mathbb{R}^n$ and $\phi : U \to V$ is a smooth map, where $U \subset \mathbb{R}^n$, then

$$\phi^*\omega = (f \circ \phi)\, d\phi_1 \wedge \cdots \wedge d\phi_n = \det(D\phi)\, (f \circ \phi)\, dx_1 \wedge \cdots \wedge dx_n .$$

This implies the following.

7.14. **Lemma.** *Let $\phi : U \to V$ be an orientation-preserving diffeomorphism, $U, V \subset \mathbb{R}^n$ open, and let ω be an n-form on V. Then*

$$\int_U \phi^*\omega = \int_{\phi(U)} \omega = \int_V \omega .$$

Proof. Let $\omega = f\, dx_1 \wedge \cdots \wedge dx_n$. We have

$$\begin{aligned}\int_U \phi^*\omega &= \int_U f\big(\phi(x)\big)\, \det(D\phi)\, |d^n x| \\ &= \int_U f\big(\phi(x)\big)\, |\det(D\phi)|\, |d^n x| = \int_V f(x)\, |d^n x| = \int_V \omega .\end{aligned}$$

Here we used that $\det(D\phi) > 0$ (ϕ preserves orientation) and the transformation formula for integrals. □

If ϕ reverses orientation, then we will get a change of sign.

Now we can define the integral of a k-form over a parametrized k-dimensional subset of $U \subset \mathbb{R}^n$.

7.15. **Definition.** Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^k$ open, let $\phi : V \to U$ be a smooth map, and let ω be a k-form on U. Then we define

$$\int_\phi \omega = \int_V \phi^*\omega .$$

7.16. **Remark.** *This integral is invariant under orientation-preserving re-parametrization: if $\psi : V' \to V$ is an orientation-preserving diffeomorphism, then*

$$\int_{\phi\circ\psi} \omega = \int_{\phi} \omega \,.$$

Proof.

$$\int_{\phi\circ\psi} \omega = \int_{V'} (\phi \circ \psi)^*\omega = \int_{V'} \psi^*(\phi^*\omega) = \int_{V} \phi^*\omega = \int_{\phi} \omega \,.$$

Note the use of Lemma 7.14 above. □

7.17. **Example.** We can now, for example, integrate a 1-form along a parametrized curve. Let $f\,dx + g\,dy + h\,dz$ be a 1-form on some open set $U \subset \mathbb{R}^3$, and let $\gamma : I \to U$ be smooth, where I is an interval. Then

$$\int_{\gamma} (f\,dx + g\,dy + h\,dz) = \int_{I} \big(f(\gamma(t))\gamma_1'(t) + g(\gamma(t))\gamma_2'(t) + h(\gamma(t))\gamma_3'(t)\big)\,dt \,.$$

In terms of the vector field $\mathbf{F}$ corresponding to the 1-form, this reads

$$\int_{\gamma} (f\,dx + g\,dy + h\,dz) = \int_{I} \mathbf{F}(\gamma(t)) \cdot \gamma'(t)\,dt \,.$$

We can now already prove a special case of Stokes' Theorem.

7.18. **Lemma (Baby Stokes).** *Let $H = \mathbb{R}_{<0} \times \mathbb{R}^{n-1}$ be the lower half-space in $\mathbb{R}^n$, let ω be an $(n-1)$-form on $\mathbb{R}^n$ with compact support (all the coefficient functions occurring in ω have compact support), and let $\phi : \mathbb{R}^{n-1} \to \mathbb{R}^n$, $x \mapsto (0, x)$. Then*

$$\int_{H} d\omega = \int_{\phi} \omega \,.$$

We use the lower (instead of the upper) half-space in order to get the correct orientation of the boundary $\{0\} \times \mathbb{R}^{n-1}$ when we use the obvious parametrization.

Proof. We can write

$$\omega = \sum_{i=1}^{n} (-1)^{i-1} f_i\, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n \,.$$

Then

$$d\omega = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}\, dx_1 \wedge \cdots \wedge dx_n \,.$$

By Fubini's Theorem and the Fundamental Theorem of Calculus, we have

$$\int_{H} \frac{\partial f_1}{\partial x_1}\, dx_1 \wedge \cdots \wedge dx_n = \int_{\mathbb{R}^{n-1}} \Big(\int_{0}^{\infty} \frac{\partial f_1}{\partial x_1}\, dx_1\Big) |dx_2 \ldots dx_n| = \int_{\mathbb{R}^{n-1}} f_1(0, x) |d^{n-1}x| = \int_{\phi} \omega \,.$$

In a similar way, we find for $i \ge 2$ that

$$\int_{H} \frac{\partial f_i}{\partial x_i}\, dx_1 \wedge \cdots \wedge dx_n = \int_{H_i} \Big(\int_{-\infty}^{\infty} \frac{\partial f_i}{\partial x_i}\, dx_i\Big) |dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n| = 0 \,.$$

Here $H_i = \{0\} \times \mathbb{R}^{n-2}$ is the projection of H that forgets the ith coordinate. Combining these relations gives the result. □

8. Differential Forms on Manifolds and Stokes' Theorem

In order to be able to formulate Stokes' Theorem, we have to look at open subsets of manifolds and their boundaries.

8.1. **Definition.** Let M be an n-dimensional manifold, $U \subset M$ open. A point $p \in \partial U$ is a *smooth boundary point* of U if there is a chart $\phi : U' \to V$ of M centered at p such that $\phi(U \cap U') = H \cap V$, where $H = \mathbb{R}_{<0} \times \mathbb{R}^{n-1}$ is the lower half-space.

Note that this implies that $\phi(\partial U \cap U') = \big(\{0\} \times \mathbb{R}^{n-1}\big) \cap V$ and that all points in $\partial U \cap U'$ are smooth boundary points.

We say that U has *smooth boundary* if all $p \in \partial U$ are smooth boundary points of U. This is equivalent to the requirement that ∂U can be covered by charts of M with the property required above.

8.2. **Remark.** *When U has smooth boundary, then ∂U is an $(n-1)$-dimensional manifold in a natural way. If M is oriented, then ∂U inherits an orientation.*

Proof. Let U have smooth boundary. Then ∂U is covered by charts ϕ as above, and we obtain charts of ∂U by restriction:

$$\phi' : \partial U \cap U' \xrightarrow{\phi} \big(\{0\} \times \mathbb{R}^{n-1}\big) \cap V \xrightarrow{\cong} V' \subset \mathbb{R}^{n-1} \,,$$

where $V' = \{v \in \mathbb{R}^{n-1} : (0, v) \in V\}$. The transition maps between these charts are restrictions of transition maps between charts of M and therefore diffeomorphisms.

If M is oriented, we can choose the charts to be compatible with the orientation (except when $\dim M = 1$; then the boundary point gets negative orientation when the chart is orientation-reversing). Since the transition maps preserve the lower half-space, their derivatives at points on the boundary hyperplane do as well, and we find that

$$D\Phi = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ * & & & \\ \vdots & & D\Phi' & \\ * & & & \end{pmatrix}$$

with $\lambda > 0$, hence $\det(D\Phi) > 0$ implies $\det(D\Phi') > 0$. □

For an oriented curve, this means that the endpoint gets positive orientation, and the starting point gets negative orientation. For a subset of $\mathbb{R}^2$, we orient the boundary curve counter-clockwise.

Now we extend the notion of differential forms to manifolds.

8.3. **Definition.** Let M be a manifold. A *differential k-form* ω on M is a section of the vector bundle $\left(\bigwedge^k TM\right)^*$ (whose fiber at $x \in M$ is the space of k-forms on the vector space T_xM). We denote by $\Omega^k(M)$ the $\mathbb{R}$-vector space of differential k-forms on M. For example, $\Omega^0(M) = \mathcal{C}^\infty(M)$ is just the space of smooth functions on M.

In more down-to-earth terms, for each chart $\phi : U \to V$ of M, ω is represented by a differential k-form ω_ϕ on V, such that for any two charts ϕ, ϕ' with transition map $\Phi = \phi' \circ \phi^{-1}$, we have $\omega_\phi = \Phi^* \omega_{\phi'}$ (on the relevant overlap).

The notions and properties of wedge products, exterior derivatives and pull-backs carry over to differential forms on manifolds. (Do it on charts and observe that everything is compatible.)

8.4. **Example.** Let us consider 1-forms on the circle $\mathbb{S}^1$. We use the two charts $\phi_1 = \left(f|_{]-\pi,\pi[}\right)^{-1}$ and $\phi_2 = \left(f|_{]0,2\pi[}\right)^{-1}$, where $f : \mathbb{R} \to \mathbb{S}^1$, $t \mapsto (\cos t, \sin t)$ is the usual parametrization. Note that the transition map $\Phi = \phi_2 \circ \phi_1^{-1}$ goes from $]-\pi, 0[\cup]0, \pi[$ to $]0, \pi[\cup]\pi, 2\pi[$ and is translation by 2π on the left interval and the identity on the right interval. If we write

$$\omega_{\phi_1} = h_1(t)\, dt \quad (-\pi < t < \pi) \qquad \text{and} \qquad \omega_{\phi_2} = h_2(t)\, dt \quad (0 < t < 2\pi)\,,$$

then we need to have that $h_1(t) = h_2(t + 2\pi)$ for $-\pi < t < 0$ and $h_1(t) = h_2(t)$ for $0 < t < \pi$. This can be interpreted by saying that h_1 and h_2 are both restrictions of a 2π-periodic function h on $\mathbb{R}$. In fact, we have $f^*\omega = h(t)\, dt$.

Finally, we need to define integrals of differential forms.

8.5. **Definition.** Let M be an oriented n-dimensional manifold, $\omega \in \Omega^n(M)$. Cover M by oriented charts $\phi_j : U_j \to V_j$, and let (ψ_j) be a subordinate partition of unity. We say that ω is *integrable* over M if

$$\sum_j \int\limits_{V_j} \left|(\psi_j \omega)_{\phi_j}\right| < \infty\,.$$

Here, we define

$$\int\limits_V |f\, dx_1 \wedge \dots \wedge dx_n| = \int\limits_V |f(x)|\, |d^n x|\,.$$

If ω is integrable over M, we set

$$\int\limits_M \omega = \sum_j \int\limits_{V_j} (\psi_j \omega)_{\phi_j}\,.$$

As before, it can be checked that this does not depend on the choice of charts or the partition of unity.

8.6. **Definition.** Let M be a manifold, let M' be an oriented k-dimensional manifold, let $\phi : M' \to M$ be a smooth map, and let $\omega \in \Omega^k(M)$. Then we set

$$\int\limits_\phi \omega = \int\limits_{M'} \phi^* \omega$$

(if $\phi^*\omega$ is integrable over M'). Note that as before, this integral is invariant under orientation-preserving re-parametrization.

If $U \subset M$ is open with smooth boundary, M an oriented n-dimensional manifold, $\omega \in \Omega^{n-1}(M)$, then we write

$$\int_{\partial U} \omega \qquad \text{for} \qquad \int_{\phi} \omega\,,$$

where $\phi : \partial U \to M$ is the inclusion, and ∂U is oriented as in Remark 8.2.

Now we can state and prove Stokes' Theorem.

8.7. **Theorem (Stokes).** *Let M be an oriented n-dimensional manifold, $U \subset M$ open with smooth boundary, $\omega \in \Omega^{n-1}(M)$. Assume that $\bar{U}$ is compact or that ω has compact support. Then*

$$\int_{U} d\omega = \int_{\partial U} \omega\,.$$

Proof. We cover U by charts $\phi_j : U_j \to V_j$ such that charts meeting the boundary ∂U are of the form required in Definition 8.1. Choose a subordinate partition of unity (ψ_j) with compact supports (we can re-number the charts to have the same index set, letting ϕ_j denote a chart whose domains contains $\operatorname{supp} \psi_j$; we may have repetitions of charts, but we don't care). There will only be finitely many ψ_j such that its support meets $\bar{U}$ and the support of ω (recall that the family of supports of the ψ_j is locally finite). We then have

$$\int_{U} d\omega = \int_{U} d\Big(\sum_j \psi_j \omega\Big) = \sum_j \int_{U} d(\psi_j \omega) = \sum_j \int_{V_j} \big(d(\psi_j \omega)_{\phi_j}\big) = \sum_j \int_{V_j} d\big((\psi_j \omega)_{\phi_j}\big)\,.$$

Note that the sum is finite, so there is no problem in interchanging it with the integration.

Now if $U_j \cap \partial U = \emptyset$, then (by an argument similar to that used in the proof of the 'Baby Stokes' result Lemma 7.18)

$$\int_{V_j} d\big((\psi_j \omega)_{\phi_j}\big) = 0\,.$$

In the other case, we find by Lemma 7.18 that

$$\int_{V_j} d\big((\psi_j \omega)_{\phi_j}\big) = \int_{V'_j} (\psi_j \omega)_{\phi'_j}\,,$$

where $\phi'_j : U_j \cap \partial U \to V'_j$ is the chart of ∂U obtained by restricting ϕ_j and projecting to the last $n-1$ coordinates (compare Remark 8.2). Let $\sum'_j$ denote the sum restricted to those j such that $U_j \cap \partial U \neq \emptyset$. Then we find

$$\int_{U} d\omega = {\sum_j}' \int_{V'_j} (\psi_j \omega)_{\phi'_j} = {\sum_j}' \int_{\partial U} \psi_j \omega = \int_{\partial U} {\sum_j}' \psi_j \omega = \int_{\partial U} \omega\,.$$

(Note that on ∂U, $\sum'_j \psi_j = 1$.) ∎

9. Interpretation of Integrals in $\mathbb{R}^n$

Let us see what Stokes' Theorem tells us about integration in $\mathbb{R}^3$ (or in $\mathbb{R}^n$). First we need to find out what integrals of k-forms on $\mathbb{R}^3$ correspond to.

9.1. **0-Forms.** A 0-form is just a function $f \in \mathcal{C}^\infty(\mathbb{R}^n) = \Omega^0(\mathbb{R}^n)$. We can integrate it over a 0-dimensional subset, which is just a finite collection of oriented points. On a single point $p \in \mathbb{R}^n$, we have

$$\int_{\pm p} f = \pm f(p)\,,$$

where the sign denotes the orientation of the point (and is *not* applied to the coordinates of p!).

9.2. **1-Forms.** A 1-form on $\mathbb{R}^n$ has the shape

$$\omega = f_1(x)\,dx_1 + f_2(x)\,dx_2 + \cdots + f_n(x)\,dx_n\,.$$

We can identify ω with the vector field $\mathbf{F} = (f_1, \ldots, f_n)^\top$.

If $\gamma : [a, b] \to \mathbb{R}^n$ is a curve, then

$$\int_\gamma \omega = \int_a^b \Big(f_1\big(\gamma(t)\big)\gamma_1'(t) + \cdots + f_n\big(\gamma(t)\big)\gamma_n'(t)\Big)\,dt = \int_a^b \mathbf{F}\big(\gamma(t)\big)\cdot\gamma'(t)\,dt = \int_\gamma \mathbf{F}\cdot\mathbf{v}\,|dx|\,;$$

the latter denoting the unoriented integral; $\mathbf{v}$ denotes the unit tangent vector in direction of the orientation of the curve. For this, we assume that $\gamma'(t)$ does not vanish; then $\gamma'(t) = \mathbf{v}(\gamma(t))\|\gamma'(t)\|$, and $\|\gamma'(t)\|$ is the factor $\sqrt{\det(D\gamma^\top\,D\gamma)}$ in the definition of the volume integral (which can be carried over to immersed manifolds, i.e., subsets parametrized by open subsets of $\mathbb{R}^k$ such that the derivative of the parametrization (like γ here) has maximal rank k everywhere).

Note that $\mathbf{F}\cdot\mathbf{v}$ gives the *tangential component* of the vector field along the curve.

9.3. **$(n-1)$-Forms.** An $(n-1)$-form on $\mathbb{R}^n$ looks like this:

$$\begin{aligned}\eta &= u_1(x)\,dx_2 \wedge \cdots \wedge dx_n - u_2(x)\,dx_1 \wedge dx_3 \wedge \cdots \wedge dx_n + \ldots \\ &\qquad + (-1)^{n-1}u_n(x)\,dx_1 \wedge \cdots \wedge dx_{n-1} \\ &= \sum_{i=1}^n (-1)^{i-1}u_i(x)\,dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n\,.\end{aligned}$$

We can identify η with the vector field $\mathbf{U} = (u_1, \ldots, u_n)^\top$.

If $\phi : W \to \mathbb{R}^n$ (with $W \subset \mathbb{R}^{n-1}$ open) parametrizes a hypersurface in $\mathbb{R}^n$, then we have

$$\begin{aligned}\int_\phi \eta &= \int_W \phi^*\eta \\ &= \int_W \Big(\sum_{i=1}^n (-1)^{i-1}u_i\big(\phi(x)\big)\,d\phi_1 \wedge \ldots d\phi_{i-1} \wedge d\phi_{i+1} \wedge \cdots \wedge d\phi_n\Big) \\ &\overset{(*)}{=} \int_W \det\Big(\mathbf{U}\big(\phi(x)\big), \frac{\partial\phi}{\partial x_1}, \ldots, \frac{\partial\phi}{\partial x_{n-1}}\Big)\,|d^{n-1}x| \\ &= \int_\phi \mathbf{U}\cdot\mathbf{n}\,|d^{n-1}x|\,.\end{aligned}$$

Here, $\mathbf{n}$ is the unit normal vector to the tangent space of the hypersurface such that $\mathbf{n}, \frac{\partial\phi}{\partial x_1}, \dots, \frac{\partial\phi}{\partial x_{n-1}}$ is a positively oriented basis. It can be checked that

$$\det\Bigl(\mathbf{U}\bigl(\phi(x)\bigr), \frac{\partial\phi}{\partial x_1}, \dots, \frac{\partial\phi}{\partial x_{n-1}}\Bigr) = \mathbf{U}\bigl(\phi(x)\bigr) \cdot \mathbf{n}\bigl(\phi(x)\bigr) \sqrt{\det(D\phi_x^\top\, D\phi_x)}\,.$$

("Volume of n-dimensional parallelotope is volume of base times height.")

So here the integral gives the *normal component* of the vector field. If we interpret the vector field as a flow (of some fluid, for example), then the integral gives us the total flow through the hypersurface in the direction given by $\mathbf{n}$ (which depends on the orientation of the hypersurface). Note that if the hypersurface is the boundary of some open subset $U \subset \mathbb{R}^n$, then $\mathbf{n}$ is the *outer* unit normal, pointing away from U.

It remains to check equality $(*)$ above. Expanding the $d\phi_j$'s, we find

$$\begin{aligned} &d\phi_1 \wedge \dots \wedge d\phi_{i-1} \wedge d\phi_{i+1} \wedge \dots \wedge d\phi_n \\ &\quad = \sum_{\sigma\in S_{n-1}} \varepsilon(\sigma) \frac{\partial\phi_{\tau_i(1)}}{\partial x_{\sigma(1)}} \cdots \frac{\partial\phi_{\tau_i(n-1)}}{\partial x_{\sigma(n-1)}}\, dx_1 \wedge \dots \wedge dx_{n-1} \\ &\quad = \det(D\phi)_i\, dx_1 \wedge \dots \wedge dx_{n-1}\,, \end{aligned}$$

where $\tau_i(k) = k$ if $k < i$, $\tau_i(k) = k+1$ if $k \geq i$, and $(D\phi)_i$ denotes the matrix $D\phi$ with the ith row removed. We then have

$$\begin{aligned} &\sum_{i=1}^{n} (-1)^{i-1} u_i d\phi_1 \wedge \dots \wedge d\phi_{i-1} \wedge d\phi_{i+1} \wedge \dots \wedge d\phi_n \\ &\qquad = \sum_{i=1}^{n} (-1)^{i-1} u_i \det(D\phi)_i\, dx_1 \wedge \dots \wedge dx_{n-1} \\ &\qquad = \det(\mathbf{U}, D\phi)\, dx_1 \wedge \dots \wedge dx_{n-1} \end{aligned}$$

as claimed.

9.4. n-**Forms.** Finally, an n-form on $\mathbb{R}^n$ is given by

$$r(x)\, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n\,.$$

for an open subset $U \subset \mathbb{R}^n$, we then simply have

$$\int_U r(x)\, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n = \int_U r(x)\, |d^n x|\,.$$

Now we can interpret Stokes' Theorem in these cases.

9.5. **Stokes for Curves and** 0**-Forms.** Let $\gamma : [a,b] \to \mathbb{R}^n$ be a curve, $f \in \mathcal{C}^\infty(\mathbb{R}^n)$. Then

$$\int_\gamma df = f\bigl(\gamma(b)\bigr) - f\bigl(\gamma(a)\bigr)\,,$$

and

$$\int_\gamma df = \int_\gamma \nabla f \cdot \mathbf{v}\, |dx| = \int_\gamma D_{\mathbf{v}} f\, |dx|$$

is the integral of the directional derivative of f in direction of the unit tangent vector of the curve. This generalizes the Fundamental Theorem of Calculus to line integrals.

9.6. **Stokes for Open Subsets and $(n-1)$-Forms.** Let $U \subset \mathbb{R}^n$ be an open subset with smooth boundary ∂U and assume (say) that $\bar{U}$ is compact. Let $\eta \in \Omega^{n-1}(\mathbb{R}^n)$, corresponding to the vector field $\mathbf{U}$ as above. Then

$$d\eta = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i}\, dx_1 \wedge \cdots \wedge dx_n = \nabla \cdot \mathbf{U}\, dx_1 \wedge \cdots \wedge dx_n\,;$$

$\nabla \cdot \mathbf{U}$ is the *divergence* of $\mathbf{U}$. We then obtain what is known as the *Divergence Theorem* (also called *Gauss' Theorem*):

$$\int\limits_U \nabla \cdot \mathbf{U}\, |d^n x| = \int\limits_{\partial U} \mathbf{U} \cdot \mathbf{n}\, |d^{n-1} x|\,,$$

where $\mathbf{n}$ is the outer unit normal vector. This says that the total flow out of the set U is the same as the total divergence of the vector field inside U; this justifies the interpretation of the divergence as the amount of flow that is 'generated' at a point.

9.7. **Green's Theorem.** This is the special case $n = 2$ of the preceding incarnation (or also the planar case of the following). If $S \subset \mathbb{R}^2$ is open and bounded with sufficiently nice boundary curve ∂S (oriented counter-clockwise), and $f, g \in \mathcal{C}^\infty(\mathbb{R}^2)$, then

$$\int\limits_S \Big(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\Big)\, |dx\, dy| = \int\limits_{\partial S} (f\, dx + g\, dy)\,.$$

9.8. **Example.** Let $U \subset \mathbb{R}^2$ be the upper semi-disk of radius R. To find

$$\int\limits_{\partial U} (x^2\, dx + 2xy\, dy)\,,$$

we can parametrize the two parts of the boundary and then have to integrate some polynomial in $\sin t$ and $\cos t$. Alternatively, we can use Green's Theorem and get

$$\int\limits_{\partial U} (x^2\, dx + 2xy\, dy) = \int\limits_U 2y\, |dx\, dy| = \int\limits_{-R}^{R} \int\limits_0^{\sqrt{R^2-x^2}} 2y\, dy\, dx = \int\limits_{-R}^{R} (R^2 - x^2)\, dx = \frac{4}{3}\, R^3\,.$$

9.9. **Stokes for Surfaces and 1-Forms on $\mathbb{R}^3$.** Let $\omega \in \Omega^1(\mathbb{R}^3)$, and let $\phi : W \to \mathbb{R}^3$ be a parametrized surface (with $W \subset \mathbb{R}^2$). Let $\mathbf{F}$ be the vector field corresponding to ω. If $\mathbf{U}$ is the vector field corresponding to $d\omega$, then we have $\mathbf{U} = \nabla \times \mathbf{F}$, i.e., $\mathbf{U}$ is the *curl* of $\mathbf{F}$. From Stokes' Theorem we obtain what is in fact the original result of Stokes:

$$\int\limits_S (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, |d^2 x| = \int\limits_{\partial S} \mathbf{F} \cdot \mathbf{v}\, |dx|\,.$$

Here, S is the oriented surface parametrized by ϕ, and ∂S is its boundary curve, oriented such that we move around S counter-clockwise when looking from the side in which the normal vector $\mathbf{n}$ points. This leads to the interpretation of the curl as giving the amount of rotation of the flow around a given axis (here, $\mathbf{n}$).

9.10. **Example.** Let $C \subset \mathbb{R}^3$ be the curve that is the intersection of the cylinder $x^2 + y^2 = 1$ and the graph $z = f(x, y)$ of a smooth function (for example, $z = \sin(10^{100}xy) + 2$), oriented clockwise when seen from 'above' (i.e., from large positive z values). We want to find

$$\int_C (y^3\,dx + x\,dy + z\,dz)\,.$$

We can parametrize C by $\gamma(t) = (\cos t, -\sin t, f(\cos t, -\sin t))$, but this is likely to result in a rather ugly integral (but see below). Instead, we can use Stokes' Theorem: let $a \in \mathbb{R}$ such that $f(x, y) > a$ for all (x, y) on the unit circle, and let S be the surface that is the cylinder intersected with $a < z < f(x, y)$. Let C' be the circle $x^2 + y^2 = 1$, $z = a$, oriented counter-clockwise when seen from above. Then $\partial S = C + C'$ (in suggestive notation), hence

$$\int_C (y^3\,dx+x\,dy+z\,dz) = \int_S 0\,|d^2x| - \int_{C'} (y^3\,dx+x\,dy+z\,dz) = -\int_{C'} (y^3\,dx+x\,dy+z\,dz)$$

(note that the curl of the vector field in question is $(0, 0, 1 - 3y^2)^\top$ and the unit normal vector on S is $(x, y, 0)^\top$, so the curl has vanishing normal component — there is no flow through the surface of the cylinder). For the integral over C', we can parametrize as usual and obtain

$$-\int_{C'} (y^3\,dx+x\,dy+z\,dz) = \int_0^{2\pi} (\sin^4 t - \cos^2 t)\,dt = \int_0^{2\pi} \frac{-1 - 8\cos 2t + \cos 4t}{8}\,dt = -\frac{\pi}{4}\,.$$

Here is a different approach. Note that

$$y^3\,dx + x\,dy + z\,dz = d\big(xy + \tfrac{1}{2}z^2\big) + (y^3 - y)\,dx\,.$$

By Stokes for 0-forms, this implies that

$$\int_C (y^3\,dx + x\,dy + z\,dz) = \int_C (y^3 - y)\,dx\,.$$

The latter integral can be computed fairly easily from the parametrization (since it does not involve z or dz); we get

$$\int_0^{2\pi} (\sin^4 t - \sin^2 t)\,dt = -\frac{\pi}{4}$$

as before.

10. Closed and Exact Forms

10.1. **Definition.** Let M be a manifold, $\omega \in \Omega^k(M)$. We say that ω is *closed*, if $d\omega = 0$. We say that ω is *exact*, if $\omega = d\eta$ for some $\eta \in \Omega^{k-1}(M)$. (When $k = 0$, then only $\omega = 0$ is exact.)

It is easy to see that every exact form is closed. What about the converse?

It is certainly not the case that a closed 0-form (which is a locally constant function) is necessarily exact (the zero function). So the question is only interesting for $k \geq 1$.

10.2. **Example.** Consider $\omega = f\,dx \in \Omega^1(\mathbb{R})$. Then $d\omega = 0$ (since there are no non-trivial 2-forms on $\mathbb{R}$). I claim that ω is exact. Indeed, we can define

$$F(x) = \int\limits_0^x f(t)\,dt = \int\limits_{[0,x]} \omega\,,$$

and then $\omega = F'(x)\,dx = dF$.

10.3. **Example.** Let $\phi : \mathbb{R} \to \mathbb{S}^1$, $t \mapsto (\cos t, \sin t)$, be the standard map. If $\omega \in \Omega^1(\mathbb{S}^1)$, then $\phi^*\omega = f\,dt$ with a 2π-periodic function f, and conversely. If $\omega = dF$, then $f(t) = (F \circ \phi)'(t)$; the function $F \circ \phi \in \mathcal{C}^\infty(\mathbb{R})$ is 2π-periodic. This implies that

$$\int\limits_0^{2\pi} f(t)\,dt = \int\limits_0^{2\pi} F\bigl(\phi(t)\bigr)\,dt = F(2\pi) - F(0) = 0\,.$$

So, for example, if $\phi^*\omega = dt$, then ω cannot be exact, since $\int_0^{2\pi} dt = 2\pi \neq 0$. This argument can be extended to show that the linear map

$$\Omega^1(\mathbb{S}^1) \longrightarrow \mathbb{R}\,, \qquad \omega \longmapsto \int\limits_{\mathbb{S}^1} \omega$$

is surjective and has kernel the subspace of exact forms $d\Omega^0(\mathbb{S}^1)$.

We will now generalize the first example above and prove the following.

10.4. **Theorem.** *If $\omega \in \Omega^k(\mathbb{R}^n)$ is closed and $k \geq 1$, then ω is exact.*

Proof. We want to do induction on the number of variables. So we define $\Omega^k_l(\mathbb{R}^n) \subset \Omega^k(\mathbb{R}^n)$ to be the subspace of all k-forms that only involve $dx_1, \dots, dx_l$, and for

$$\omega = \sum_{1 \le i_1 < \cdots < i_k \le l} f_{i_1,\dots,i_k}\, dx_{i_1} \wedge \cdots \wedge dx_{i_k} \in \Omega^k_l(\mathbb{R}^n)\,,$$

we set

$$d^{(l)}\omega = \sum_{1 \le i_1 < \cdots < i_k \le l} d^{(l)} f_{i_1,\dots,i_k} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}\,,$$

where

$$d^{(l)} f = \sum_{i=1}^{l} \frac{\partial f}{\partial x_i}\, dx_i\,.$$

The statement we now prove by induction on l is the following.

If $\omega \in \Omega^k_l(\mathbb{R}^n)$ with $k \geq 1$ and $d^{(l)}\omega = 0$, then there is $\eta \in \Omega^{k-1}_l(\mathbb{R}^n)$ such that $\omega = d^{(l)}\eta$.

For $l = n$, we have $\Omega^k_n(\mathbb{R}^n) = \Omega^k(\mathbb{R}^n)$ and $d^{(n)} = d$, so we obtain the statement of the theorem.

We start with $l = 1$. If $\omega \neq 0$, then we must have $\omega = f\,dx_1$, and $d^{(1)}\omega = 0$ is automatic. We define $\eta(x) = \int_0^{x_1} f(t, x_2, \dots, x_n)\,dt$; then $\omega = d^{(1)}\eta$, compare Example 10.2 above.

Now assume the statement is proved for l, and we want to prove it for $l+1$. So let $\omega \in \Omega^k_{l+1}(\mathbb{R}^n)$ with $k \geq 1$ and $d^{(l+1)}\omega = 0$. Then we can write $\omega = \omega_1 + \omega_2 \wedge dx_{l+1}$ with $\omega_1 \in \Omega^k_l(\mathbb{R}^n)$ and $\omega_2 \in \Omega^{k-1}_l(\mathbb{R}^n)$. We have

$$0 = d^{(l+1)}\omega = d^{(l)}\omega_1 + \omega' \wedge dx_{l+1}$$

for some $\omega' \in \Omega_l^k(\mathbb{R}^n)$. Since $d^{(l)}\omega_1$ does not involve terms with dx_{l+1}, this implies that $d^{(l)}\omega_1 = 0$ as well (looking at coefficients of the standard basis). By induction, there is $\eta_1 \in \Omega_l^{k-1}(\mathbb{R}^n)$ such that $d^{(l)}\eta_1 = \omega_1$. Write

$$d^{(l+1)}\eta_1 = d^{(l)}\eta_1 + \eta_2 \wedge dx_{l+1} = \omega_1 + \eta_2 \wedge dx_{l+1}$$

and set

$$\tilde{\omega} = \omega - d^{(l+1)}\eta_1 = (\omega_2 - \eta_2) \wedge dx_{l+1}\,.$$

We have $d^{(l+1)}\tilde{\omega} = d^{(l+1)}\omega - d^{(l+1)}d^{(l+1)}\eta_1 = 0$, so

$$d^{(l)}(\omega_2 - \eta_2) \wedge dx_{l+1} = d^{(l+1)}(\omega_2 - \eta_2) \wedge dx_{l+1} = d^{(l+1)}\tilde{\omega} = 0\,,$$

which implies that $d^{(l)}(\omega_2 - \eta_2) = 0$. If $k \geq 2$, then by induction again, there is $\zeta \in \Omega_l^{k-2}(\mathbb{R}^n)$ such that $d^{(l)}\zeta = \omega_2 - \eta_2$. Then

$$\begin{aligned} d^{(l+1)}(\eta_1 + \zeta \wedge dx_{l+1}) &= d^{(l+1)}\eta_1 + d^{(l)}\zeta \wedge dx_{l+1} \\ &= \omega_1 + \eta_2 \wedge dx_{l+1} + (\omega_2 - \eta_2) \wedge dx_{l+1} \\ &= \omega_1 + \omega_2 \wedge dx_{l+1} = \omega\,, \end{aligned}$$

and we are done. If $k = 1$, then $h = \omega_2 - \eta_2$ is a function that does not depend on $x_1, \ldots, x_l$. Write $h(x_{l+1}, \ldots, x_n)$ for it and define

$$H(x) = \int_0^{x_{l+1}} h(t, x_{l+2}, \ldots, x_n)\, dt \quad \text{and} \quad \eta_1' = \eta_1 + H\,.$$

Then

$$d^{(l+1)}\eta_1' = d^{(l+1)}\eta_1 + d^{(l+1)}H = \omega_1 + \eta_2\, dx_{l+1} + h\, dx_{l+1} = \omega_1 + \omega_2\, dx_{l+1} = \omega\,,$$

and we are done again. □

Note that since everything commutes with pull-backs, the same result holds for any manifold M such that M is diffeomorphic to some $\mathbb{R}^n$, for example an open ball. Note also that the proof is *constructive:* following the induction, we can construct a form η such that $\omega = d\eta$, in terms of simple (meaning 1-dimensional) integrals.

10.5. **Interpretation in $\mathbb{R}^3$.** If we consider our interpretations of differential forms and exterior derivatives on $\mathbb{R}^3$, we obtain the following.

- *A vector field* $\mathbf{F}$ *on* $\mathbb{R}^3$ *is the gradient* ∇f *of a function if and only if its curl vanishes:* $\nabla \times \mathbf{F} = 0$.
- *A vector field* $\mathbf{U}$ *on* $\mathbb{R}^3$ *is the curl* $\nabla \times \mathbf{F}$ *of a vector field if and only if its divergence vanishes:* $\nabla \cdot \mathbf{U} = 0$.
- *Any function on* $\mathbb{R}^3$ *is the divergence of a vector field.*

In general, we know that the image of $d : \Omega^{k-1}(M) \to \Omega^k(M)$ (i.e., the exact k-forms) is a subspace of the kernel of $d : \Omega^k(M) \to \Omega^{k+1}(M)$ (i.e., the closed k-forms). The following definition then makes sense.

10.6. **Definition.** Let M be a manifold. For $k \geq 0$, we define the *kth de Rham cohomology group* of M to be the real vector space

$$H^k_{\mathrm{dR}}(M) = \frac{\ker\bigl(d : \Omega^k(M) \to \Omega^{k+1}(M)\bigr)}{\mathrm{im}\bigl(d : \Omega^{k-1}(M) \to \Omega^k(M)\bigr)}\,.$$

Here $\Omega^{-1}(M) = 0$, so $H^0_{\mathrm{dR}}(M)$ is the space of functions f such that $df = 0$. These are exactly the locally constant functions, i.e., they are constant on every connected component of M. Therefore,

$$H^0_{\mathrm{dR}}(M) \cong \mathbb{R}^{\{\text{components of } M\}}\,.$$

Also, if $k > \dim M$, then $H^k_{\mathrm{dR}}(M) = 0$, since already $\Omega^k(M) = 0$.

In general, these cohomology groups contain information about the topology of M. Note that when ω is a closed k-form and S is an oriented $(k+1)$-submanifold of M, then by Stokes,

$$\int_{\partial S} \omega = \int_S d\omega = 0\,.$$

So if T and T' are oriented k-dimensional submanifolds of M such that $T - T' = \partial S$ for a $(k+1)$-dimensional oriented submanifold S (where $T - T'$ means the union of T and T', where the orientation is reversed on T'), then

$$\int_T \omega - \int_{T'} \omega = \int_{\partial S} \omega = 0\,.$$

The condition implies that T and T' have the same boundary, so $\int_T \omega$ only depends on ∂T as long as T does not 'cross a hole' in M. An example of this 'crossing a hole' is when we compare the upper and lower semicircle of $\mathbb{S}^1$; in this case a closed form does not have to give the same integral.

If $\omega = d\eta$ is exact, then for an oriented k-submanifold $T \subset M$,

$$\int_T \omega = \int_T d\eta = \int_{\partial T} \eta\,,$$

and so the integral only depends on ∂T without any conditions. So if M has 'k-dimensional holes', we can expect to find a non-trivial $H^k_{\mathrm{dR}}(M)$.

10.7. **Example.** We have

$$H^k_{\mathrm{dR}}(\mathbb{R}^n) = \begin{cases} \mathbb{R} & \text{if } k = 0, \\ 0 & \text{if } k > 0. \end{cases}$$

This follows from Theorem 10.4 above.

10.8. **Example.** We have

$$H^k_{\mathrm{dR}}(\mathbb{S}^1) = \begin{cases} \mathbb{R} & \text{if } k = 0, 1, \\ 0 & \text{if } k > 1. \end{cases}$$

This follows from Example 10.3.

We mention the following result without proof.

10.9. **Theorem.** *If M is compact, then* $\dim H^k_{\mathrm{dR}}(M) < \infty$ *for all* $k \ge 0$.

This holds more generally for manifolds that can be covered by finitely many charts such that the intersection of the domains of any nonempty subset of this collection of charts is diffeomorphic to $\mathbb{R}^n$ (or to an open ball, which is the same up to diffeomorphism). Here is a sketch for $k = 1$. (The case $k = 0$ follows from the observation made in the definition above, since a compact manifold only has finitely many connected components.)

So let M be compact, with charts $\phi_j : U_j \to V_j$ that have the stated property, and let $\omega \in \Omega^1(M)$ such that $d\omega = 0$. By Thm. 10.4, there are functions $f_j \in \mathcal{C}^\infty(U_j)$ such that $df_j = \omega|_{U_j}$ (since U_j is diffeomorphic to $\mathbb{R}^n$). So on $U_i \cap U_j$, we have $d(f_j - f_i) = df_j - df_i = 0$, so there are constants (since $U_i \cap U_j$ is connected) $c_{ij} \in \mathbb{R}$ with $f_j|_{U_i \cap U_j} - f_i|_{U_i \cap U_j} = c_{ij}$. Any change of f_j will be by a constant γ_j, so c_{ij} can be replaced by $c_{ij} + \gamma_j - \gamma_i$. We obtain a well-defined linear map

$$\ker\bigl(d : \Omega^1(M) \to \Omega^2(M)\bigr) \longrightarrow \frac{\{(c_{ij})_{i<j} : c_{ij} \in \mathbb{R}\}}{\{(\gamma_j - \gamma_i)_{i<j} : \gamma_i \in \mathbb{R}\}}\,.$$

If ω is in the kernel of this map, then there are $\gamma_i \in \mathbb{R}$ such that $f_j - f_i = \gamma_j - \gamma_i$ on $U_i \cap U_j$. So $f_j - \gamma_j = f_i - \gamma_i$, and we can define a function f on M such that $f = f_j - \gamma_j$ on U_j, for all j. Then $\omega = df$ is exact. This gives us an *injective* linear map

$$H^1_{\mathrm{dR}}(M) = \frac{\ker\bigl(d : \Omega^1(M) \to \Omega^2(M)\bigr)}{\operatorname{im}\bigl(d : \Omega^0(M) \to \Omega^1(M)\bigr)} \longrightarrow \frac{\{(c_{ij})_{i<j} : c_{ij} \in \mathbb{R}\}}{\{(\gamma_j - \gamma_i)_{i<j} : \gamma_i \in \mathbb{R}\}}$$

with target a finite-dimensional vector space, so $\dim H^1_{\mathrm{dR}}(M) < \infty$.

This approach can be extended to general k, but gets a bit technical. In the end, we get an isomorphism of the de Rham cohomology group with the *Čech cohomology group* of the cover by charts (this is defined in terms of the combinatorial structure of the collection (U_j)), which computes the *singular cohomology group* (with values in $\mathbb{R}$) of M; this is a topological invariant.

10.10. **Exercise.** Show that the wedge product induces bilinear maps on the de Rham cohomology groups

$$\wedge : H^k_{\mathrm{dR}}(M) \times H^l_{\mathrm{dR}}(M) \longrightarrow H^{k+l}_{\mathrm{dR}}(M)\,.$$

10.11. **Exercise.** Denote by $\Delta_n = \{t_1 e_1 + \cdots + t_n e_n : 0 < t_i < 1\} \subset \mathbb{R}^n$ the n-dimensional standard simplex (where $e_1, \ldots, e_n$ is the canonical basis of $\mathbb{R}^n$). Set $e_0 = 0$. We write $e_j^{(n)}$ to distinguish the vectors in $\mathbb{R}^n$ from the corresponding vectors in other spaces. Define $\iota_m^n : \mathbb{R}^{n-1} \to \mathbb{R}^n$ to be the affine map that sends $e_0^{(n-1)}, \ldots, e_{n-1}^{(n-1)}$ to $e_0^{(n)}, \ldots, e_{m-1}^{(n)}, e_{m+1}^{(n)}, \ldots, e_n^{(n)}$ (i.e., we leave out $e_m^{(n)}$). Let M be a manifold, and let $C^k(M)$ be the set of all formal real linear combinations of smooth maps $f : \mathbb{R}^k \to M$. Define a boundary operator $\partial : C^k(M) \to C^{k-1}(M)$ by

$$\partial f = \sum_{i=0}^{k} (-1)^{i-1} (f \circ \iota_i^k)\,.$$

(1) Show that $\partial\partial f = 0$.

We can therefore define *homology groups*

$$H_k(M) = \frac{\ker\big(\partial : C^k(M) \to C^{k-1}(M)\big)}{\operatorname{im}\big(\partial : C^{k+1}(M) \to C^k(M)\big)}\,.$$

For example, $H_0(M) = \bigoplus_{\text{components of } M} \mathbb{R}$. The idea behind $H_k(M)$ is that it should in some way count the 'k-dimensional holes' in M.

(2) Show that

$$(\omega, f) \longmapsto \int\limits_{f|_{\Delta_k}} \omega$$

induces a bilinear map $\Omega^k(M) \times C^k(M) \to \mathbb{R}$. We denote it by $(\omega, \alpha) \mapsto \langle \omega, \alpha \rangle$.

(3) Show that $\langle \omega, \partial\alpha \rangle = \langle d\omega, \alpha \rangle$ for $\omega \in \Omega^k(M)$ and $\alpha \in C^{k+1}(M)$.

(4) Show that the bilinear map above induces a bilinear map

$$H^k_{\mathrm{dR}}(M) \times H_k(M) \longrightarrow \mathbb{R}\,.$$

One can show that the latter bilinear map induces an isomorphism between $H^k_{\mathrm{dR}}(M)$ and the dual space $H_k(M)^* = \operatorname{Hom}(H_k(M), \mathbb{R})$.

11. Lebesgue Integration

We now turn to the other important topic of this course, which is the *Lebesgue Integral*. Its introduction is motivated by some shortcomings of the Riemann Integral; what it does is to generalize the Riemann integral to a larger class of functions. However, the Riemann integral is still very important, since it provides a way of actually *computing* an integral (as a limit of Riemann sums). Of course, it is also important, because the definition of the Lebesgue integral is based on it.

The following is based on the book [HH].

One of the problems with the Riemann integral is that it does not behave very well with respect to point-wise limits of functions. Ideally, we would like to have a statement of the following type.

If f is a point-wise limit of a sequence of integrable functions (f_n), then f is integrable and $\int f = \lim_{n\to\infty} \int f_n$.

Now there are various reasons why this cannot be true in this generality. The main reason is that we can 'lose mass' in the limit, because it gets shifted out to infinity or to the boundary of the domain. For example, let $D =]-1, 0[\cup]0, 1[$, and define f_n to be n times the characteristic function of $D \cap]-1/n, 1/n[$. Then f_n is Riemann integrable on D, and $\int_D f_n(x)\,dx = 2$ for all n. On the other hand, $f_n(x)$ tends to zero for each $x \in D$, so $f_n \to 0$ point-wise, but the integral obviously does not tend to zero. In order to avoid this problem, we have to prevent the mass from escaping. A reasonable way of doing this is to require all the functions to be bounded by a fixed integrable function. This gives the statement of the *Dominated Convergence Theorem:*

Let g be an integrable function, and let (f_n) be a sequence of integrable functions such that $|f_n| \le g$ for all n. If (f_n) converges point-wise to a function f, then f is integrable, and $\int f = \lim_{n\to\infty} \int f_n$.

We will prove this later for Lebesgue integrable functions, with a slightly weakened convergence hypothesis. But let us first see why this does not hold for Riemann integrable functions. For this, consider a sequence (a_n) that enumerates all the rational numbers in the interval $[0,1]$, and define $f_n : [0,1] \to \mathbb{R}$ to be the function that is zero everywhere except that $f_n(a_k) = 1$ for $k \le n$. Then each f_n is integrable, with integral zero. The sequence (f_n) converges point-wise to a function f, which is the characteristic function of $\mathbb{Q} \cap [0,1]$. Now the problem is that this function f is *not* Riemann integrable — it is 'too discontinuous'. In fact, every upper Riemann sum is 1 and every lower Riemann sum is 0. So we will need to extend our notion of 'integrable function' to also include f (which then should have integral zero). Before we can do this, we need to introduce the notion of 'measure zero'.

11.1. **Definition.** An *open box* in $\mathbb{R}^n$ is an open cube

$$B = \,]a_1, a_1 + \delta[\, \times \cdots \times \,]a_n, a_n + \delta[\, \subset \mathbb{R}^n$$

where $(a_1, \dots, a_n) \in \mathbb{R}^n$ and $\delta > 0$.

A subset $X \subset \mathbb{R}^n$ has *measure zero* if for every $\varepsilon > 0$, there exists a sequence (B_k) of open boxes in $\mathbb{R}^n$ such that

$$X \subset \bigcup_k B_k \quad \text{and} \quad \sum_k \operatorname{vol}_n(B_k) \le \varepsilon\,.$$

Note that this differs from the notion of *volume zero*, which requires a *finite* set of cubes covering X of arbitrary small volume. Obviously, every set of volume zero also has measure zero, but the converse is false.

11.2. **Exercise.** Show that in the definition of 'measure zero' we could use arbitrary pavable sets instead of open boxes without changing which sets have measure zero.

11.3. **Example.** The set $\mathbb{Q} \cap [0,1]$ has measure zero (in $\mathbb{R}$), but is not pavable and therefore does not have a defined volume (or length, in this case).

To see that the set has measure zero, enumerate $\mathbb{Q} \cap [0,1] = \{a_1, a_2, \dots\}$. For given $\varepsilon > 0$, let $B_k = \big]a_k - 2^{-k-1}\varepsilon, a_k + 2^{-k-1}\varepsilon\big[$, then $\operatorname{vol}_1 B_k = 2^{-k}\varepsilon$, so $\sum_{k=1}^{\infty} \operatorname{vol}_1 B_k = \varepsilon$, and clearly $\bigcup_k B_k$ contains $\mathbb{Q} \cap [0,1]$.

To see that the set is not pavable, note that any finite union of closed intervals covering $\mathbb{Q} \cap [0,1]$ must also cover $[0,1]$ (since otherwise it would miss a whole open interval which will contain lots of rational numbers). On the other hand, no interval of positive length is contained in $\mathbb{Q} \cap [0,1]$, so the lower volume is 0 and the upper volume is 1.

11.4. **Proposition.**

(1) *If $X \subset \mathbb{R}^n$ has measure zero and $Y \subset X$, then $Y \subset \mathbb{R}^n$ has measure zero.*
(2) *If $X_1, X_2, \dots \subset \mathbb{R}^n$ all have measure zero, then their union $X = \bigcup_{k=1}^{\infty} X_k$ also has measure zero.*

Proof. The first statement is clear from the definition (just take the same boxes as for X). For the second statement, we use the same idea as in the example above. Let $\varepsilon > 0$. Then for each k, we can choose boxes B_{kj}, $j = 1, 2, \dots$ covering X_k and with total volume $\le 2^{-k}\varepsilon$. Then all these boxes together (which can again be arranged in a sequence) cover X and have total volume $\le \varepsilon$. $\square$

The importance of sets of measure zero comes from the fact that they will be what can be safely ignored when (Lebesgue) integrating functions. But they are also important for the Riemann integral.

11.5. **Theorem.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be bounded and have bounded support. Then f is Riemann integrable if and only if f is continuous on $\mathbb{R}^n \setminus X$, where X is a set of measure zero.*

Proof. See [HH], Thm. 4.4.5, pages 440–442. □

Now let us turn back to convergence of integrals. One easy case is the following.

11.6. **Theorem.** *If (f_k) is a sequence of bounded Riemann integrable functions on $\mathbb{R}^n$, all with support contained in a fixed bounded set and converging uniformly to a function f, then f is Riemann integrable, and*

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

Proof. Exercise. □

This is good enough for many applications, but not really satisfying. We can go a bit further and prove a version of the Dominated Convergence Theorem for Riemann integrals.

11.7. **Dominated Convergence Theorem for Riemann Integrals.** *Let $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \dots$ be a sequence of Riemann integrable functions. Suppose that $\operatorname{supp} f_k \subset B$ for a fixed bounded set $B \subset \mathbb{R}^n$ and that $|f_k| \le M$ for all k, with some fixed $M > 0$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be Riemann integrable and such that $(f_k(x))_k$ converges to $f(x)$ except on a set of measure zero. Then*

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

The big weakness of this theorem is that is requires the limit function f to be Riemann integrable. This is hard to show in practice, so we would like to have a notion of integrability where this is automatic. Note also that the requirements on boundedness and support of the f_k is equivalent to saying that all $|f_k|$ are bounded by a fixed Riemann integrable function g, which we can take to be $M\chi_B$.

Proof. Let us make some simplifications of the statement. First of all, we can consider $f_k - f$ instead of f_k and so take $f = 0$ without loss of generality. The next step is to remove the exceptional set of measure zero. So assume the theorem holds when we require point-wise convergence everywhere, and let (f_k), $f = 0$ etc. be as in the statement above. Let $\varepsilon > 0$ and pick a countable union $X_\varepsilon = \bigcup_j B_j$ of open boxes covering the exceptional set and such that $\sum_j \operatorname{vol}_n B_j \le \varepsilon$. Then there is a continuous function $h : \mathbb{R}^n \to [0, 1]$ such that $h(x) = 1$ for $x \notin X_\varepsilon$ and $0 \le h(x) < 1$ for $x \in X_\varepsilon$. Now consider the sequence of functions $g_k = h^k f_k$. Then the g_k also satisfy the assumptions, but additionally, we have that $g_k(x) \to 0$

for all $x \in \mathbb{R}^n$. So we can apply the simplified version to the g_k and find that $\lim_{k\to\infty} \int g_k = 0$. On the other hand,

$$\Big|\int_{\mathbb{R}^n} f_k(x)\,|d^n x| - \int_{\mathbb{R}^n} g_k(x)\,|d^n x|\Big| \le \int_{\mathbb{R}^n} |f_k(x) - g_k(x)|\,|d^n x|$$
$$\le M \int_{R^n} \big(1 - h(x)^k\big)\,|d^n x| \le M \sum_j \operatorname{vol}_n B_j \le M\varepsilon\,,$$

so we see that

$$\limsup_{k\to\infty} \Big|\int_{\mathbb{R}^n} f_k(x)\,|d^n x|\Big| \le M\varepsilon\,.$$

Since $\varepsilon > 0$ was arbitrary, this implies that $\int f_k \to 0$ as desired.

Next, we can write $f_k = f_k^+ - f_k^-$, where $f_k^+ = \max\{f_k, 0\}$ and $f_k^- = \max\{-f_k, 0\}$ are the positive and negative parts of f_k, as usual. Then $f_k^+ \to 0$ and $f_k^- \to 0$ pointwise, and we see that it suffices to prove the theorem for non-negative functions.

Finally, we can scale the variable and the functions in such a way that $0 \le f_k \le 1$ and $\operatorname{supp} f_k \subset Q$ for all k, where Q is the unit cube in $\mathbb{R}^n$. The theorem therefore follows from the result below. □

11.8. **Proposition.** *Let (f_k) be a sequence of functions with support in Q and such that $0 \le f_k \le 1$. Suppose that $f_k(x) \to 0$ for all $x \in Q$. Then*

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x| = 0\,.$$

In order to prove this, we will first look at a special case, where the functions form a decreasing sequence.

11.9. **Proposition.** *Let (f_k) be a sequence of functions with support in Q and such that $1 \ge f_1 \ge f_2 \ge \dots \ge 0$. Suppose that $f_k(x) \to 0$ for all $x \in Q$. Then*

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x| = 0\,.$$

Proof. Since the sequence of functions is decreasing, the sequence given by the values of the integrals is also decreasing; it is also bounded below by zero, so if the conclusion is false, then

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x| = 2K > 0\,.$$

Define $A_k = \{x \in \mathbb{R}^n : f_k(x) \ge K\} \subset Q$. The idea of the proof is that $A_1 \supset A_2 \supset \dots$ is a nested sequence of sets each of which should have volume $\ge K$, so their intersection is non-empty, which contradicts the assumption that $f_k(x) \to 0$ for all x. The problem with that approach is that the A_k need not be pavable. We get around that by considering the *lower volume* $\underline{\operatorname{vol}}_n A_k$, which I claim is $\ge K$:

$$2K \le \int_Q f_k(x)\,|d^n x| \le \int_Q \max\{K, f_k(x)\}\,|d^n x| = L\big(\max\{K, f_k\}\chi_Q\big) \le K + \underline{\operatorname{vol}}_n A_k\,.$$

Here, $L(f) = \lim_{N\to\infty} L_N(f)$ denotes the lower integral of f. For each N, we have

$$\begin{aligned} L_N\big(\max\{K, f_k\}\chi_Q\big) &= \sum_{Q'\in\mathcal{D}_N, Q'\subset Q} 2^{-Nn} \inf\big\{\max\{K, f(x)\} : x \in Q'\big\} \\ &\le K + 2^{-Nn}\#\{Q' \in \mathcal{D}_N : Q' \subset A_k\} \\ &\le K + \underline{\mathrm{vol}}_n A_k \end{aligned}$$

(Note that $f_k \le 1$.) For each k, we can therefore find a finite union A'_k of *closed* dyadic cubes such that $A'_k \subset A_k$ and $\mathrm{vol}_n A'_k \ge \underline{\mathrm{vol}}_n A_k - 2^{-k-1}K$. Let $A''_k = A'_1 \cap \dots \cap A''_k$, then $A''_1 \supset A''_2 \supset \dots$ is a nested sequence of compact sets. We have

$$\begin{aligned} \mathrm{vol}_n A''_k = \mathrm{vol}_n A'_k - \mathrm{vol}_n\Big(A'_k \setminus \bigcap_{j=1}^{k-1} A'_j\Big) \\ \ge \mathrm{vol}_n A'_k - \sum_{j=1}^{k-1} \mathrm{vol}_n(A'_k \setminus A'_j) \\ \ge K - 2^{-k-1}K - \sum_{j=1}^{k-1} 2^{-j-1}K > K/2 > 0\,; \end{aligned}$$

note that $A'_k \setminus A'_j \subset A_j \setminus A'_j$, hence $\mathrm{vol}_n(A'_k \setminus A'_j) \le \underline{\mathrm{vol}}_n A_j - \mathrm{vol}_n A'_j \le 2^{-j-1}K$. This shows in particular that the A''_k are non-empty, so by a standard property of compact sets, it follows that their intersection is non-empty as well. But then,

$$x \in \bigcap_{k=1}^{\infty} A''_k \implies x \in \bigcap_{k=1}^{\infty} A_k \implies f_k(x) \ge K \quad \text{for all } k \ge 1\,,$$

which contradicts the assumption $f_k(x) \to 0$ for all x and therefore finishes the proof. □

11.10. **Corollary.** *Let h and h_k, for $k \ge 1$, be Riemann integrable functions with support in the unit cube Q, such that $0 \le h \le 1$, $h_k \ge 0$ for all k, and $h(x) \le \sum_{k=1}^{\infty} h_k(x)$ for all $x \in Q$ (where we allow the series to diverge to ∞). Then*

$$\int_{\mathbb{R}^n} h(x)\,|d^n x| \le \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} h_k(x)\,|d^n x|\,.$$

Proof. Set $f_k(x) = \max\Big\{0, h(x) - \sum_{j=1}^{k} h_j(x)\Big\}$. Then f_k is Riemann integrable, we have $0 \le f_k(x) \le h(x) \le 1$ for all x, $f_1 \ge f_2 \ge \dots$, and $f_k(x) \to 0$ for all x. So by Prop. 11.9, we know that $\int f_k \to 0$. This implies

$$\begin{aligned} \int_{\mathbb{R}^n} h(x)\,|d^n x| &= \lim_{k\to\infty} \int_{\mathbb{R}^n} \big(h(x) - f_k(x)\big)\,|d^n x| \\ &= \lim_{k\to\infty} \int_{\mathbb{R}^n} \min\Big\{h(x), \sum_{j=1}^{k} h_j(x)\Big\}\,|d^n x| \\ &\le \lim_{k\to\infty} \sum_{j=1}^{k} \int_{\mathbb{R}^n} h_j(x)\,|d^n x| = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} h_j(x)\,|d^n x|\,. \end{aligned}$$

□

Now we come back to the proof of Prop. 11.8.

Proof. Ideally, we would like to consider $g_k = \sup\{f_j : j \geq k\}$; this would be a decreasing sequence of functions converging to zero point-wise, and so we could apply the Monotone Convergence Theorem Prop. 11.9 to it, allowing us to conclude easily. However, the problem is (as usual) that these g_k may not be Riemann integrable. For example, enumerate $\mathbb{Q} \cap [0,1] = \{a_1, a_2, \dots\}$, and let $f_k = \chi_{\{a_k\}}$. These f_k satisfy the assumptions of the proposition, but g_k is the characteristic function of $\mathbb{Q} \cap [0,1] \setminus \{a_1, \dots, a_{k-1}\}$, which is not Riemann integrable. So we have to use some tricks to make the proof work.

Assume the conclusion is false. Since all the integrals are nonnegative, this implies that there is a subsequence of (f_k), which we can assume to be the sequence (f_k) itself, such that

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x| = C > 0\,.$$

Now, as a substitute for the g_k considered above, we will take certain linear combinations of the f_k. For $p \geq 1$, Let

$$K_p = \left\{ \sum_{k=p}^{\infty} a_k f_k : a_k \geq 0, \sum_{k=p}^{\infty} a_k = 1, a_k = 0 \text{ for } k \text{ large} \right\}.$$

If (g_p) is a sequence of functions with $g_p \in K_p$, then we still have that $g_p(x) \to 0$ for all x and $\int g_p \to C$. (For the former, let $\varepsilon > 0$. Then there is K such that for $k \geq K$, we have $0 \leq f_k < \varepsilon$. So if $p \geq K$, we have $0 \leq g_p < \varepsilon$, too. For the latter, given $\varepsilon > 0$ again, there is K such that $\left|\int f_k - C\right| < \varepsilon$ for $k \geq K$, hence

$$\left|\int g_p - C\right| = \left|\int \sum_{k\geq p} a_k f_k - C\right| = \left|\sum_{k\geq p} a_k \Big(\int f_k - C\Big)\right|$$
$$\leq \sum_{k\geq p} a_k \left|\int f_k - C\right| < \sum_{k\geq p} a_k \varepsilon = \varepsilon\,.)$$

Now the idea is to take g_p that are small, so that we have a chance to get a contradiction to $\int g_p \to C$. Therefore, we consider

$$d_p = \inf\left\{ \int_{\mathbb{R}^n} g_p(x)^2\,|d^n x| : g_p \in K_p \right\}.$$

Since $K_{p+1} \subset K_p$, this is an increasing sequence of nonnegative numbers; it is bounded by 1 (since $g_p \leq 1$), so has a limit d. Then for every $p \geq 1$, we can pick a $g_p \in K_p$ such that

$$\int_{\mathbb{R}^n} g_p(x)^2\,|d^n x| \leq d + \frac{1}{p}\,.$$

Now I claim that these g_p are close to one another when p is large: Given $\varepsilon > 0$, there is N such that whenever $p, q \geq N$, we have

$$\int_{\mathbb{R}^n} \big(g_p(x) - g_q(x)\big)^2\,|d^n x| \leq \varepsilon\,.$$

To see this, let N be so large that $d - d_N \leq \varepsilon/8$ and $1/N \leq \varepsilon/8$. By simple algebra, we have that

$$\Big(\frac{1}{2}\big(g_p(x) - g_q(x)\big)\Big)^2 + \Big(\frac{1}{2}\big(g_p(x) + g_q(x)\big)\Big)^2 = \frac{1}{2} g_p(x)^2 + \frac{1}{2} g_q(x)^2\,.$$

This gives

$$\begin{aligned}
\frac{1}{4}\int\limits_{\mathbb{R}^n} &\big(g_p(x)-g_q(x)\big)^2\,|d^n x| \\
&= \frac{1}{2}\int\limits_{\mathbb{R}^n} g_p(x)^2\,|d^n x| + \frac{1}{2}\int\limits_{\mathbb{R}^n} g_q(x)^2\,|d^n x| - \int\limits_{\mathbb{R}^n}\Big(\frac{1}{2}\big(g_p(x)+g_q(x)\big)\Big)^2\,|d^n x| \\
&\le \frac{1}{2}\Big(d+\frac{1}{p}\Big) + \frac{1}{2}\Big(d+\frac{1}{q}\Big) - d_N \\
&\le d - d_N + \frac{1}{N} \le \frac{\varepsilon}{4}\,.
\end{aligned}$$

Note that $p, q \ge N$, so $\frac{1}{2}(g_p + g_q) \in K_N$.

This now allows us to pick a subsequence (h_q) of (g_p) such that

$$\sum_{q=1}^{\infty}\Big(\int\limits_{\mathbb{R}^n}\big(h_q(x)-h_{q+1}(x)\big)^2\,|d^n x|\Big)^{1/2} < \infty\,.$$

To do this, pick a sequence (ε_q) of positive numbers such that $\sum_q \sqrt{\varepsilon_q} < \infty$. Then there is an increasing sequence $N_1 < N_2 < \dots$ such that for $p, p' \ge N_q$, we have $\int (g_p - g_{p'})^2 \le \sqrt{\varepsilon_q}$. We can then set $h_q = g_{N_q}$.

Now observe that $h_q = \sum_{j=q}^{\infty}(h_j - h_{j+1})$ (since the partial sums are $h_q - h_m$, and $h_m \to 0$). This implies

$$h_q(x) \le \sum_{j=q}^{\infty}\big|h_j(x) - h_{j+1}(x)\big| \qquad \text{for all } x \in \mathbb{R}^n.$$

We can then apply Cor. 11.10 and find

$$\begin{aligned}
\int\limits_{\mathbb{R}^n} h_q(x)\,|d^n x| &\le \sum_{j=q}^{\infty}\int\limits_{\mathbb{R}^n}\big|h_j(x)-h_{j+1}(x)\big|\,|d^n x| \\
&\le \sum_{j=q}^{\infty}\Big(\int\limits_{\mathbb{R}^n}\big(h_q(x)-h_{q+1}(x)\big)^2\,|d^n x|\Big)^{1/2}.
\end{aligned}$$

(We use the Cauchy-Schwarz inequality for integrals here.) But the latter tends to zero as $q \to \infty$, since it is the tail end of a converging series. This contradicts the fact that

$$\lim_{q\to\infty}\int\limits_{\mathbb{R}^n} h_q(x)\,|d^n x| = C > 0\,,$$

which was a consequence of our assumption that the conclusion of Prop. 11.8 does not hold. So the Proposition is finally proved. □

Our next goal will be to define the Lebesgue integral. For this, we need the following result.

11.11. **Proposition.** *Let $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \dots$, be Riemann-integrable functions such that*

$$\sum_{k=1}^{\infty}\int\limits_{\mathbb{R}^n}|f_k(x)|\,|d^n x| < \infty\,.$$

Then there is a set $X \subset \mathbb{R}^n$ of measure zero such that the series $\sum_{k=1}^{\infty} f_k(x)$ converges for all $x \in \mathbb{R}^n \setminus X$.

Proof. We will take

$$X = \left\{ x \in \mathbb{R}^n : \sum_{k=1}^{\infty} |f_k(x)| \text{ diverges} \right\}.$$

Then $\sum_k f_k(x)$ converges absolutely on $\mathbb{R}^n \setminus X$. The hard part is to show that X has measure zero. Let

$$A = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|\,|d^n x|.$$

If $A = 0$, then $\int |f_k| = 0$ for all k. Now observe that if f_k is continuous at $x \in \mathbb{R}^n$ and $f_k(x) \neq 0$, then $|f_k|$ is bounded below by a positive constant on some neighborhood of x, hence $\int |f_k| > 0$. So $f_k(x) = 0$ for all x where f_k is continuous. Since f_k is Riemann-integrable, $f_k = 0$ except on a set of measure zero. Since a countable union of sets of measure zero is again a set of measure zero, $f_k = 0$ for all k except on a set of measure zero; in particular, $\sum_k f_k$ converges (to zero) except on a set of measure zero, and the claim is proved in this case.

So assume now that $A > 0$. We first 'normalize convergence' of the sum of integrals: we choose an increasing sequence (k_m) of integers such that setting

$$h_m = \sum_{k=1}^{k_m} |f_k|, \quad \text{we have} \quad \int_{\mathbb{R}^n} h_m(x)\,|d^n x| \geq A\Big(1 - \frac{1}{2^{2m+3}}\Big).$$

Let $\varepsilon > 0$. We will construct a countable union Y of dyadic cubes with total volume $\leq 3\varepsilon$ and such that on $\mathbb{R}^n \setminus Y$, we have $h_m \leq 2A/\varepsilon$ for all m. The latter implies that $X \subset Y$, and since ε can be chosen arbitrarily small, this will show that X has measure zero.

The set Y will be a disjoint union of sets Y_m, $m = 0, 1, \ldots$, each a finite union of dyadic cubes, and such that

$$\sum_{j=0}^{m} \mathrm{vol}_n(Y_j) \leq \varepsilon\Big(3 - \frac{1}{2^m}\Big) \quad \text{and} \quad h_m(x) \leq \frac{A}{\varepsilon}\Big(2 - \frac{1}{2^m}\Big) \text{ for all } x \in \mathbb{R}^n \setminus \bigcup_{j=0}^{m} Y_j.$$

We will construct the Y_m recursively. We begin with Y_0. Since $\int h_0 \leq A$, there is N_0 such that

$$U_{N_0}(h_0) \leq L_{N_0}(h_0) + A \leq 2A$$

(where $U_N(h)$ and $L_N(h)$ denote the upper and lower Riemann sums of the function h with respect to dyadic cubes of size 2^{-N}). We take Y_0 to be the union of those cubes $C \in \mathcal{D}_{N_0}(\mathbb{R}^n)$ such that $\sup_{x \in C} h_0(x) > A/\varepsilon$. Then

$$\frac{A}{\varepsilon}\,\mathrm{vol}_n(Y_0) = \frac{A}{\varepsilon} \sum_{C \in \mathcal{D}_{N_0}, C \subset Y_0} \mathrm{vol}_n(C) \leq \sum_{C \in \mathcal{D}_{N_0}} \sup_{x \in C} h_0(x)\,\mathrm{vol}_n(C) = U_{N_0}(h_0) \leq 2A,$$

and hence $\mathrm{vol}_n(Y_0) \leq 2\varepsilon$.

Now assume, $Y_0, \ldots, Y_m$ have been constructed, together with an increasing sequence $N_0 < N_1 < \cdots < N_m$. Set $g_{m+1} = h_{m+1} - h_m$, then

$$\int_{\mathbb{R}^n} g_{m+1}(x)\,|d^n x| = \Big(A - \int_{R^n} h_m(x)\,|d^n x|\Big) - \Big(A - \int_{\mathbb{R}^n} h_{m+1}(x)\,|d^n x|\Big) \leq \frac{A}{2^{2m+3}}.$$

In the same way as above in the construction of Y_0, we find $N_{m+1} > N_m$ such that

$$U_{N_{m+1}}(g_{m+1}) \leq \frac{A}{2^{2m+2}}.$$

We let Y_{m+1} be the union of the cubes $C \in \mathcal{D}_{N_{m+1}}(\mathbb{R}^n)$ that are not contained in $Y_0 \cup \cdots \cup Y_m$ such that

$$\sup_{x \in C} h_{m+1}(x) > \frac{A}{\varepsilon}\Big(2 - \frac{1}{2^{m+1}}\Big) .$$

On each such C, we must then have $\sup_{x\in C} g_{m+1}(x) > A/(2^{m+1}\varepsilon)$, which implies

$$\frac{A}{2^{m+1}\varepsilon} \operatorname{vol}_n(Y_{m+1}) \le U_{N_{m+1}}(g_{m+1}) \le \frac{A}{2^{2m+2}} , \quad \text{hence} \quad \operatorname{vol}_n(Y_{m+1}) \le \frac{\varepsilon}{2^{m+1}} .$$

The Y_m thus constructed then have the required properties. □

The upshot of this result is that whenever we have a sequence (f_k) of Riemann-integrable functions such that $\sum_k \int |f_k|$ converges, then $\sum_k f_k$ will be defined "almost everywhere", i.e., everywhere except on a set of measure zero. Based on the intuition that sets of measure zero "don't matter" with respect to integration, we can then make the following definition.

11.12. **Definition.** Let $f_k, g_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \ldots$ be Riemann-integrable functions such that

$$\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|\, |d^n x| < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |g_k(x)|\, |d^n x| < \infty .$$

We write

$$\sum_{k=1}^{\infty} f_k \underset{L}{=} \sum_{k=1}^{\infty} g_k$$

if there is a set $X \subset \mathbb{R}^n$ of measure zero such that for all $x \in \mathbb{R}^n \setminus X$, both series $\sum_k f_k(x)$ and $\sum_k g_k(x)$ converge and have the same value. More generally, if f and g are functions defined on $\mathbb{R}^n$ except a set of measure zero, then we write $f \underset{L}{=} g$ if $f(x) = g(x)$ for all $x \in \mathbb{R}^n \setminus X$, where X is a set of measure zero such that f and g are both defined on $\mathbb{R}^n \setminus X$.

The 'L' in '$\underset{L}{=}$' stands for 'Lebesgue'.

We need another theorem to make sure that the definition we are aiming at will make sense.

11.13. **Theorem.** *Let $f_k, g_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \ldots$ be Riemann-integrable functions such that*

$$\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|\, |d^n x| < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |g_k(x)|\, |d^n x| < \infty$$

and assume that $\sum_{k=1}^{\infty} f_k \underset{L}{=} \sum_{k=1}^{\infty} g_k$. Then

$$\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} f_k(x)\, |d^n x| = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} g_k(x)\, |d^n x| .$$

Proof. Let $h_m = \sum_{k=1}^{m} (f_k - g_k)$, then (h_m) is a sequence of Riemann-integrable functions that converges to zero except on a set of measure zero. If the h_m are all bounded by some constant M and all have support in a fixed bounded set B, then the claim follows from the Dominated Convergence Theorem for Riemann

Integrals 11.7. In the general case, we use a suitable 'truncation': for any function h on $\mathbb{R}^n$ and $R > 0$, define

$$[h]_R(x) = \begin{cases} h(x) & \text{if } \|x\| \le R \text{ and } |h(x)| \le R \\ 0 & \text{otherwise.} \end{cases}$$

Pick $\varepsilon > 0$, then there is m such that both 'tail ends' of the sums of integrals are small:

$$\sum_{k=m+1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|\,|d^n x| < \frac{\varepsilon}{7} \quad \text{and} \quad \sum_{k=m+1}^{\infty} \int_{\mathbb{R}^n} |g_k(x)|\,|d^n x| < \frac{\varepsilon}{7}\,.$$

Also, h_m is Riemann-integrable, hence bounded and with bounded support, so there is some $R > 0$ such that $h_m = [h_m]_R$. I claim that for $k > m$, we have

$$\big|h_k - [h_k]_{2R}\big| \le 3 \sum_{j=m+1}^{k} \big(|f_j - g_j|\big) \le 3 \sum_{j=m+1}^{k} \big(|f_j| + |g_j|\big)\,.$$

Indeed, if a and b are functions such that $[a]_R = a$, then for $\|x\| \le 2R$,

$$\big|[a+b]_{2R}(x) - a(x) - [b]_{2R}(x)\big| = \begin{cases} 0 & \text{if } |b(x)| \le 2R,\ |a(x)+b(x)| \le 2R, \\ |b(x)| & \text{if } |b(x)| > 2R,\ |a(x)+b(x)| \le 2R, \\ |a(x)+b(x)| & \text{if } |b(x)| \le 2R,\ |a(x)+b(x)| > 2R, \\ |a(x)| & \text{if } |b(x)| > 2R,\ |a(x)+b(x)| > 2R. \end{cases}$$

In each case, this is $\le 2|b(x)|$ (note that in the third line, we must have

$$|b(x)| \ge |a(x)+b(x)| - |a(x)| > 2R - R = R \ge |a(x)|\,).$$

So with $a = h_m$ and $b = h_k - h_m$, we find that

$$\big|[h_k]_{2R} - h_m - [h_k - h_m]_{2R}\big| \le 2\big|[h_k - h_m]_{2R}\big| \le 2|h_k - h_m|$$

The claim then follows, since

$$\big|h_k - [h_k]_{2R}\big| \le \big|h_k - h_m - [h_k - h_m]_{2R}\big| + 2|h_k - h_m| \le 3|h_k - h_m| \le 3 \sum_{j=m+1}^{k} |f_j - g_j|\,.$$

For $k > m$, we then have

$$\left|\int_{\mathbb{R}^n} \big(h_k(x) - [h_k]_{2R}(x)\big)\,|d^n x|\right| \le 3 \sum_{j=m+1}^{k} \int_{\mathbb{R}^n} \big(|f_j(x)| + |g_j(x)|\big)\,|d^n x| < \frac{6}{7}\varepsilon\,.$$

Now we apply Thm. 11.7 to the sequence $([h_k]_{2R})$. This tells us that there is an index $K \ge m$ such that for $k > K$, we have

$$\left|\int_{\mathbb{R}^n} [h_k]_{2R}(x)\,|d^n x|\right| < \frac{\varepsilon}{7}\,.$$

So for $p > K$, we obtain

$$\begin{aligned}\Big|\sum_{k=1}^{p}\int_{\mathbb{R}^n} f_k(x)\,|d^n x| - \sum_{k=1}^{p}\int_{\mathbb{R}^n} g_k(x)\,|d^n x|\Big| &= \Big|\int_{\mathbb{R}^n} h_p(x)\,|d^n x|\Big| \\ &\le \Big|\int_{\mathbb{R}^n} [h_p]_{2R}(x)\,|d^n x|\Big| + \Big|\int_{\mathbb{R}^n} \big(h_p(x) - [h_p]_{2R}(x)\big)\,|d^n x|\Big| \\ &< \frac{\varepsilon}{7} + \frac{6}{7}\varepsilon = \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this proves the claim. □

We can now (at last!) make the following definition.

11.14. **Definition.** Let $f : \mathbb{R}^n \setminus X \to \mathbb{R}$ be a function, where X is a set of measure zero. We say that f is *(Lebesgue-)integrable* if there is a sequence of Riemann-integrable functions $f_k : \mathbb{R}^n \to \mathbb{R}$ such that

$$\sum_{k=1}^{\infty}\int_{\mathbb{R}^n} |f_k(x)|\,|d^n x| < \infty$$

and such that

$$f \underset{L}{=} \sum_{k=1}^{\infty} f_k\,.$$

In this case, the *(Lebesgue) integral* of f is

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| = \sum_{k=1}^{\infty}\int_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

By the preceding theorem, the Lebesgue integral is well-defined.

11.15. **Example.** If $f : \mathbb{R}^n \to \mathbb{R}$ is Riemann-integrable, then f is Lebesgue-integrable, and the Lebesgue integral of f equals the Riemann integral of f. (And thus we are justified to use the same notation for both integrals!)

Indeed, we can just take $f_1 = f$ and $f_k = 0$ for $k \ge 2$ in the definition above.

11.16. **Corollary.** *If f, g are functions defined almost everywhere on $\mathbb{R}^n$, $f \underset{L}{=} g$, and f is Lebesgue integrable, then so is g, and*

$$\int_{R^n} f(x)\,|d^n x| = \int_{\mathbb{R}^n} g(x)\,|d^n x|\,.$$

Proof. Take any sequence (f_k) as in Def. 11.14 such that $f \underset{L}{=} \sum_k f_k$. Then we also have $g \underset{L}{=} \sum_k f_k$, and the result follows. □

It thus makes sense to talk about Lebesgue integrability and Lebesgue integrals of functions that are only defined on $\mathbb{R}^n \setminus X$, where X has measure zero.

11.17. **Example.** The standard example that shows that Lebesgue integration theory is richer than Riemann integration theory is the function $f = \chi_{[0,1]\cap\mathbb{Q}}$. As we have seen, f is not Riemann integrable (all upper Riemann sums $U_N(f)$ are 1, all lower Riemann sums $L_N(f)$ are zero), but f is certainly Lebesgue integrable, since $f \underset{L}{=} 0$ (and so f has integral zero).

11.18. **Examples.** Another situation where the Lebesgue integral is of advantage is when dealing with unbounded functions or functions with unbounded support. For example, $f(x) = 1/(x^2+1)$ is not Riemann integrable, but it is Lebesgue integrable: we can write $f = \sum_{k\in\mathbb{Z}} f\chi_{[k,k+1[}$, and every $f_k = f\chi_{[k,k+1[}$ is Riemann integrable, with

$$\int_{\mathbb{R}} |f_k(x)|\,|dx| = \int_{\mathbb{R}} f_k(x)\,|dx| = \int_k^{k+1} \frac{dx}{x^2+1} = \arctan(k+1) - \arctan(k)\,,$$

so

$$\sum_{k\in\mathbb{Z}} \int_{\mathbb{R}} |f_k(x)|\,|dx| = \sum_{k=-\infty}^{\infty} \big(\arctan(k+1) - \arctan(k)\big) = \pi\,,$$

and since everything is nonnegative, this is also the (Lebesgue) integral of f.

As an example of an unbounded function, consider $f(x) = 1/\sqrt{x}$ for $0 < x < 1$, $f(x) = 0$ else. Again, we can write f as a series $f = \sum_{k=1}^{\infty} f_k$ with $f_k = f\chi_{]1/(k+1),1/k]}$, and we find that f is Lebesgue integrable with integral 2.

In both cases, the result agrees with the improper Riemann integral

$$\lim_{a\to-\infty,b\to\infty} \int_a^b \frac{dx}{x^2+1} \qquad \text{or} \qquad \lim_{\varepsilon\searrow 0} \int_\varepsilon^1 \frac{dx}{\sqrt{x}}\,.$$

However, it is not always the case that a function, for which an improper integral like the above exists, is Lebesgue integrable. As an example, consider $f(x) = (\sin x)/x$ for $x > 0$, $f(x) = 0$ otherwise. The improper integral

$$\lim_{a\to\infty} \int_0^a \frac{\sin x}{x}\,dx$$

exists, but f is not Lebesgue integrable. Indeed, if it were, then $|f|$ would also be Lebesgue integrable (see below), which would imply that

$$\lim_{a\to\infty} \int_0^a \frac{|\sin x|}{x}\,dx < \infty\,,$$

which is not the case.

We can now generalize the notion of pavable sets. First, we define measurable functions; these are functions that are 'locally integrable'.

11.19. **Definition.**

(1) A function $f : \mathbb{R}^n \to \mathbb{R}$ is *measurable* if for every $R > 0$, $[f]_R$ is Lebesgue integrable, where $[f]_R$ is the 'R-truncation of f' that was used in the proof of Thm. 11.13.

(2) A subset $A \subset \mathbb{R}^n$ is *measurable* if its characteristic function χ_A is a measurable function. A has *finite measure* if χ_A is Lebesgue integrable; then the *measure* of A is

$$\operatorname{meas}(A) = \int\limits_{\mathbb{R}^n} \chi_A(x)\,|d^n x|\,.$$

Otherwise, A has *infinite measure*, and we write $\operatorname{meas}(A) = \infty$.

(3) If $A \subset \mathbb{R}^n$ is measurable and $f : A \to \mathbb{R}$ is a function, then we say that f is *(Lebesgue) integrable on* A if $\tilde{f}$ is Lebesgue integrable (on $\mathbb{R}^n$), where $\tilde{f}(x) = f(x)$ for $x \in A$ and $\tilde{f}(x) = 0$ else. We then write

$$\int\limits_A f(x)\,|d^n x| := \int\limits_{\mathbb{R}^n} \tilde{f}(x)\,|d^n x|\,.$$

Note that a Lebesgue integrable function is measurable, and a measurable function f is Lebesgue integrable if and only if

$$\sup_{R>0} \int\limits_{\mathbb{R}^n} \big|[f]_R(x)\big|\,|d^n x| < \infty\,.$$

This will follow from the Monotone Convergence Theorem 11.24 below (applied to $|[f]_k|$) and the discussion following it.

Also, a set of measure zero is measurable and has finite measure zero.

We will now prove some simple properties of the Lebesgue integral: it is linear and monotonic.

11.20. **Proposition.** *If $f, g : \mathbb{R}^n \to \mathbb{R}$ are Lebesgue integrable, $a, b \in \mathbb{R}$, then $af + bg$ is Lebesgue integrable, and*

$$\int\limits_{\mathbb{R}^n} \big(af(x) + bg(x)\big)\,|d^n x| = a \int\limits_{\mathbb{R}^n} f(x)\,|d^n x| + b \int\limits_{\mathbb{R}^n} g(x)\,|d^n x|\,.$$

Proof. Let $f \underset{L}{=} \sum_k f_k$ and $g \underset{L}{=} \sum_k g_k$ as in Def. 11.14. Then

$$\sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^n} \big|af_k(x) + bg_k(x)\big|\,|d^n x| \le |a| \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^n} \big|f_k(x)\big|\,|d^n x| + |b| \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^n} \big|g_k(x)\big|\,|d^n x| < \infty$$

and $af + bg \underset{L}{=} \sum_k (af_k + bg_k)$. So $af + bg$ is Lebesgue integrable, and

$$\begin{aligned}
\big(af(x) + bg(x)\big)\,|d^n x| &= \sum_{k=1}^{\infty} \int\limits_{R^n} \big(af_k(x) + bg_k(x)\big)\,|d^n x| \\
&= a \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^n} f_k(x)\,|d^n x| + b \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^n} g_k(x)\,|d^n x| \\
&= a \int\limits_{\mathbb{R}^n} f(x)\,|d^n x| + b \int\limits_{\mathbb{R}^n} g(x)\,|d^n x|\,.
\end{aligned}$$

□

In the following, we write $f \underset{L}{\leq} g$, if $f(x) \leq g(x)$ for all x except on a set of measure zero, and similarly for $f \underset{L}{\geq} g$.

11.21. **Proposition.** *If $f, g : \mathbb{R}^n \to \mathbb{R}$ are Lebesgue integrable and $f \underset{L}{\leq} g$, then*

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| \leq \int_{\mathbb{R}^n} g(x)\,|d^n x|\,.$$

Proof. Write $g = f + h$ where $h = g - f \underset{L}{\geq} 0$. It suffices to show that $\int h \geq 0$. Write $h \underset{L}{=} \sum_k h_k$, and let $H_k = \left(\sum_{j=1}^k h_j\right)_+$ (where, as usual, $F_+ = \max\{0, F\}$). Set $\tilde{h}_k = H_k - H_{k-1}$ (with $H_0 = 0$). Then $|\tilde{h}_k| \leq |h_k|$, $h \underset{L}{=} \sum_k \tilde{h}_k$, and $\sum_{j=1}^k \tilde{h}_j = H_k \geq 0$. We find

$$\int_{\mathbb{R}^n} h(x)\,|d^n x| = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} \tilde{h}_k(x)\,|d^n x| = \lim_{k\to\infty} \int_{\mathbb{R}^n} H_k(x)\,|d^n x| \geq 0\,.$$

□

11.22. **Proposition.** *If $f, g : \mathbb{R}^n \to \mathbb{R}$ are Lebesgue integrable, then* $\max\{f, g\}$ *and* $\min\{f, g\}$ *are Lebesgue integrable as well. In particular, $|f|$ is Lebesgue integrable.*

Proof. The definition of Lebesgue integrability can be formulated as follows: f is Lebesgue integrable if there is a sequence (f_k) of Riemann integrable functions with $f_k \to f$ almost everywhere (i.e., except on a set of measure zero) and such that $\sum_{k=1}^{\infty} \int |f_{k+1} - f_k| < \infty$. Now consider $\max\{f_k, g_k\}$ (or $\min\{f_k, g_k\}$) and observe that (for example)

$$\big|\max\{f_{k+1}, g_{k+1}\} - \max\{f_k, g_k\}\big| \leq |f_{k+1} - f_k| + |g_{k+1} - g_k|\,.$$

The last statement follows from $|f| = \max\{f, -f\}$. □

The next theorem is the key for the important results that follow. It generalizes the formula that we used to define the Lebesgue integral in Def. 11.14 to series of Lebesgue integrable functions.

11.23. **Theorem.** *Let $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \ldots$, be Lebesgue integrable and assume that*

$$\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|\,|d^n x| < \infty\,.$$

Then the series $\sum\limits_{k=1}^{\infty} f_k$ converges almost everywhere. Let $f \underset{L}{=} \sum\limits_{k=1}^{\infty} f_k$, then f is Lebesgue integrable, and

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

Proof. We write $f_k \underset{L}{=} \sum_{j=1}^{\infty} f_{k,j}$ with $f_{k,j}$ Riemann integrable and $\sum_j \int |f_{k,j}| < \infty$. We would like to write $f \underset{L}{=} \sum_{k,j} f_{k,j}$, but $\sum_{k,j} \int |f_{k,j}|$ may not converge. So we need to modify our representations of the f_k in such a way that the double sum converges. We do this by putting most of each series into the first term.

For every k, there is an index $m(k)$ such that

$$\sum_{j>m(k)} \int_{\mathbb{R}^n} |f_{k,j}(x)|\,|d^n x| < 2^{-k}\,.$$

We now set

$$g_{k,1} = \sum_{j=1}^{m(k)} f_{k,j} \qquad \text{and for } j > 1, \quad g_{k,j} = f_{k,m(k)+j-1}\,.$$

Then $f_k \underset{L}{=} \sum_{j=1}^{\infty} g_{k,j}$ and

$$\begin{aligned}
\sum_{j,k=1}^{\infty} \int_{\mathbb{R}^n} |g_{k,j}(x)|\,|d^n x| &= \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |g_{k,1}(x)|\,|d^n x| + \sum_{k=1}^{\infty}\sum_{j=2}^{\infty} \int_{\mathbb{R}^n} |g_{k,j}(x)|\,|d^n x| \\
&= \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} \Big| f_k(x) - \sum_{j>m(k)} g_{k,j}(x) \Big|\,|d^n x| + \sum_{k=1}^{\infty}\sum_{j>m(k)} \int_{\mathbb{R}^n} |f_{k,j}(x)|\,|d^n x| \\
&\le \sum_{k=1}^{\infty} \Big(|f_k(x)|\,|d^n x| + 2^{-k} \Big) + \sum_{k=1}^{\infty} 2^{-k} < \infty\,.
\end{aligned}$$

So the series

$$\sum_{k=1}^{\infty} f_k = \sum_{k,j=1}^{\infty} g_{k,j}$$

converges almost everywhere (by Prop. 11.11), and if $f \underset{L}{=} \sum_{k=1}^{\infty} f_k$, then f is Lebesgue integrable, and

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \int_{\mathbb{R}^n} g_{k,j}(x)\,|d^n x| = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

□

An immediate corollary is the following.

11.24. **Monotone Convergence Theorem.** *Let $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \dots,$ be Lebesgue integrable such that $0 \underset{L}{\le} f_1 \underset{L}{\le} f_2 \underset{L}{\le} \dots$. If*

$$\sup_k \int_{\mathbb{R}^n} f_k(x)\,|d^n x| < \infty\,,$$

then $\lim_{k\to\infty} f_k$ exists almost everywhere. If $f \underset{L}{=} \lim_{k\to\infty} f_k \underset{L}{=} \sup_k f_k$, then f is Lebesgue integrable, and

$$\int_{\mathbb{R}^n} f(x)\,|d^n x| = \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\,|d^n x| = \sup_k \int_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

Proof. Apply Thm. 11.23 to the series $f_1 + (f_2 - f_1) + (f_3 - f_2) + \dots$. All of its terms are nonnegative (almost everywhere), therefore the assumption here implies the assumption in that theorem. The limits equal the suprema since the sequences are increasing (almost everywhere in case of the functions). □

If the assumption $\sup_k \int f_k < \infty$ is not satisfied, then either $\sup_k f_k$ does not exist almost everywhere (i.e., the subset on which $\sup_k f_k(x) = \infty$ does not have measure zero), or else the resulting function f is not Lebesgue integrable (since $\infty > \int f \geq \sup_k \int f_k$ otherwise, which is not possible).

11.25. **Proposition.** *Let $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \dots$, be Lebesgue integrable and assume that there is a Lebesgue integrable function $F : \mathbb{R}^n \to \mathbb{R}$ such that $|f_k| \underset{L}{\leq} F$ for all k. Then $\sup_k f_k$ and $\inf_k f_k$ exist almost everywhere and are Lebesgue integrable, and we have*

$$-\int\limits_{\mathbb{R}^n} F(x)\,|d^n x| \leq \int\limits_{\mathbb{R}^n} \inf_k f_k(x)\,|d^n x| \leq \inf_k \int\limits_{\mathbb{R}^n} f_k(x)\,|d^n x|$$
$$\leq \sup_k \int\limits_{\mathbb{R}^n} f_k(x)\,|d^n x| \leq \int\limits_{\mathbb{R}^n} \sup_k f_k(x)\,|d^n x| \leq \int\limits_{\mathbb{R}^n} F(x)\,|d^n x|\,.$$

Proof. Once we prove that $\sup_k f_k$ and $\inf_k f_k$ are Lebesgue integrable, the inequalities will follow from Prop. 11.21. It suffices to consider $\sup_k f_k$ (apply the statement to $(-f_k)$ to get the inf). Also, if we replace f_k by $\tilde{f}_k = f_k + F$, we have $0 \underset{L}{\leq} \tilde{f}_k \underset{L}{\leq} 2F$, so we can assume without loss of generality that $f_k \underset{L}{\geq} 0$. Now consider the sequence

$$g_1 = f_1, \quad g_2 = \max\{g_1, f_2\}, \quad g_3 = \max\{g_2, f_3\}, \quad \dots.$$

We have $0 \underset{L}{\leq} g_1 \leq g_2 \leq g_3 \leq \dots \underset{L}{\leq} F$, the g_k are Lebesgue integrable, $\sup_k f_k = \sup_k g_k$, and

$$\sup_k \int\limits_{\mathbb{R}^n} g_k(x)\,|d^n x| \leq \int\limits_{\mathbb{R}^n} F(x)\,|d^n x| < \infty\,.$$

By the Monotone Convergence Theorem 11.24, $\sup_k f_k$ is then defined almost everywhere and Lebesgue integrable. □

11.26. **Dominated Convergence Theorem.** *Let $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, 2, \dots$, be Lebesgue integrable and assume that there is a Lebesgue integrable function $F : \mathbb{R}^n \to \mathbb{R}$ such that $|f_k| \underset{L}{\leq} F$ for all k. If the sequence (f_k) converges almost everywhere to a function f, then f is Lebesgue integrable with integral*

$$\int\limits_{\mathbb{R}^n} f(x)\,|d^n x| = \lim_{k\to\infty} \int\limits_{\mathbb{R}^n} f_k(x)\,|d^n x|\,.$$

Proof. Let $g_k = \inf\{f_j : j \geq k\}$ and $h_k = \sup\{f_j : j \geq k\}$. By the preceding proposition, g_k and h_k are Lebesgue integrable, and we have that

$$\sup_k g_k \underset{L}{=} f \underset{L}{=} \inf_k h_k \quad \text{and} \quad g_k \underset{L}{\leq} f_k \underset{L}{\leq} h_k\,.$$

This already shows that f is Lebesgue integrable, again by the preceding proposition. It also implies that (using Thm. 11.24; note that (g_k) is increasing and (h_k) is decreasing)

$$\int\limits_{\mathbb{R}^n} f(x)\,|d^nx| = \int\limits_{\mathbb{R}^n} \inf_k h_k(x)\,|d^nx| = \lim_{k\to\infty}\int\limits_{\mathbb{R}^n} h_k(x)\,|d^nx| \ge \limsup_{k\to\infty}\int\limits_{\mathbb{R}^n} f_k(x)\,|d^nx|$$

and similarly

$$\int\limits_{\mathbb{R}^n} f(x)\,|d^nx| = \int\limits_{\mathbb{R}^n} \sup_k g_k(x)\,|d^nx| = \lim_{k\to\infty}\int\limits_{\mathbb{R}^n} g_k(x)\,|d^nx| \le \liminf_{k\to\infty}\int\limits_{\mathbb{R}^n} f_k(x)\,|d^nx|$$

These two statements together show that $\lim\limits_{k\to\infty}\int f_k$ exists and is equal to $\int f$. □

The Dominated Convergence Theorem is very strong, since it only requires a minimal amount of assumptions (compare with the version for Riemann integrals!). As an example of its many applications, we can prove the following result on integrals depending on a parameter.

11.27. **Theorem.** *Let $I \subset \mathbb{R}$ be an open interval, $a \in I$, and let $f : I \times \mathbb{R}^n \to \mathbb{R}$ be a function such that for every $t \in I$, $f_t : x \mapsto f(t,x)$ is Lebesgue integrable with $|f_t| \underset{L}{\le} F$ for some fixed Lebesgue integrable function $F : \mathbb{R}^n \to \mathbb{R}$. Assume that $I \ni t \mapsto f(t,x)$ is continuous at a for all $x \in \mathbb{R}^n$ except on a set of measure zero. Then*

$$\lim_{t\to a}\int\limits_{\mathbb{R}^n} f(t,x)\,|d^nx| = \int\limits_{\mathbb{R}^n} f(a,x)\,|d^nx|\,,$$

i.e., the function $I \to \mathbb{R}$, $t \mapsto \int f_t$, is continuous at a.

Proof. Take any sequence (t_k) such that $t_k \to a$. Then we can apply Thm. 11.26 to the sequence of functions $(f_{t_k})_{k\ge 1}$. Since $f_a \underset{L}{=} \lim\limits_{k\to\infty} f_{t_k}$, we obtain that

$$\lim_{k\to\infty}\int\limits_{\mathbb{R}^n} f(t_k,x)\,|d^nx| = \int\limits_{\mathbb{R}^n} f(a,x)\,|d^nx|\,.$$

Since this holds for all sequences as above, this proves the claim. □

We can extend this to differentiability instead of continuity. 'Almost all' means 'all except on a set of measure zero'.

11.28. **Theorem.** *Let $I \subset \mathbb{R}$ be an open interval, let $f : I \times \mathbb{R}^n \to \mathbb{R}$ be a function such that for almost all $x \in \mathbb{R}^n$, $\frac{\partial f}{\partial t}(t,x)$ exists. Assume that for all $f \in I$, $x \mapsto f(t,x)$ is Lebesgue integrable and that there is a Lebesgue integrable function $F : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\left|\frac{f(s,x)-f(t,x)}{s-t}\right| \le F(x) \quad \text{for almost all } x \in \mathbb{R}^n$$

whenever $s,t \in I$, $s \ne t$. Then

$$g : I \longrightarrow \mathbb{R}\,, \quad t \longmapsto \int\limits_{\mathbb{R}^n} f(t,x)\,|d^nx|$$

is differentiable on I, and

$$g'(t) = \int\limits_{\mathbb{R}^n} \frac{\partial f}{\partial t}(t,x)\,|d^n x|\,.$$

Proof. Fix $t \in I$ and apply the previous theorem to

$$\tilde{f}(s,x) = \begin{cases} \frac{f(s,x)-f(t,x)}{s-t} & \text{if } s \neq t \\ \frac{\partial f}{\partial t}(t,x) & \text{if } s = t. \end{cases}$$

□

11.29. **Are There Non-Measurable Sets?** The Monotone Convergence Theorem 11.24 implies that a countable union of measurable sets is again measurable (Exercise). Pavable sets like open balls are measurable. Since every open set is a countable union of open balls, all open sets are measurable. The complement of a measurable set is measurable, and sets of measure zero are measurable. So any set that can be constructed from sets of measure zero and open (or closed) sets by taking complements or countable unions (or intersections) will be measurable. It is hard to image a set that cannot be obtained in this way. So it is a natural question to ask if there are any subsets of $\mathbb{R}$ (say) that are *not* measurable.

It turns out that this is tied up with fundamental questions in Set Theory. It is not possible to construct a non-measurable set in an explicit way. However, if we allow ourselves to use the Axiom of Choice, then non-measurable sets can be constructed. Here is a standard example. For each coset $a + \mathbb{Q} \subset \mathbb{R}$, pick a representative in $[0,1]$, and let X be the set whose elements are these representatives. Let M be the set of rational numbers in $[-1,1]$, then we have

$$[0,1] \subset Y = \coprod_{r \in M} (X + r) \subset [-1,2]$$

(the set Y is the disjoint union of the sets $X + r$). If X were measurable, then we would have $0 \leq \text{meas}(X) \leq 1$ and $\text{meas}(X+r) = \text{meas}(X)$. Since M is countable, the set Y above would be measurable, with measure ≤ 3 and ≥ 1. So X cannot have measure zero (otherwise $\text{meas}(Y) = 0$), but likewise, X cannot have positive measure (otherwise $\text{meas}(Y) = \infty$), and we get a contradiction.

On the other hand, adding the axiom that all sets are measurable does not lead to a contradiction in Set Theory *without* the Axiom of Choice. So we have a choice here as to which assumptions we would like to make. It should be said that the Axiom of Choice is very important in some areas of mathematics (Functional Analysis, for example — one of its basic results, the *Hahn-Banach Theorem*, is in fact equivalent to the Axiom of Choice), so the general attitude today is to accept it, and with it accept its somewhat counter-intuitive consequences like the *Banach-Tarski Paradox*. It says that you can partition the unit ball in $\mathbb{R}^3$ into a finite number (e.g., seven) pieces, that after moving them around by euclidean motions will form two disjoint unit balls. Of course, this implies that at least some of the sets cannot be measurable.

There are two more 'big theorems' on the Lebesgue integral: Fubini's Theorem and the change-of-variables formula. Both follow from their counterparts that hold for Riemann integrals.

11.30. **Fubini's Theorem.** *Let $f : \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be Lebesgue integrable. Then $f_y : x \mapsto f(x,y)$ is Lebesgue integrable for almost all y, the function*

$$\mathbb{R}^m \longrightarrow \mathbb{R}\,, \quad y \longmapsto \int\limits_{\mathbb{R}^n} f(x,y)\,|d^n x|$$

(which is defined almost everywhere) is Lebesgue integrable, and we have

$$\int\limits_{\mathbb{R}^{n+m}} f(x,y)\,|d^{n+m}(x,y)| = \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} f(x,y)\,|d^n x| \Big)\,|d^m y|\,.$$

Conversely, if $f : \mathbb{R}^{n+m} \to \mathbb{R}$ is measurable such that $f_y : x \mapsto f(x,y)$ is Lebesgue integrable for almost all y, and $y \mapsto \int_{\mathbb{R}^n} |f(x,y)|\,|d^n x|$ is Lebesgue integrable, then f is Lebesgue integrable, and the relation above holds.

Proof. Note that the second statement follows from the first: if f is measurable, then $[f]_R$ and therefore also $|[f]_R|$ are Lebesgue integrable for every $R > 0$. From the first statement, we find

$$\begin{aligned}\int\limits_{\mathbb{R}^{n+m}} \big|[f]_R(x,y)\big|\,|d^{n+m}(x,y)| &= \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} \big|[f]_R(x,y)\big|\,|d^n x| \Big)\,|d^m y| \\ &\le \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} |f(x,y)|\,|d^n x| \Big)\,|d^m y| < \infty\,.\end{aligned}$$

By the Monotone Convergence Theorem 11.24, applied to the sequence $\big(|[f]_k|\big)$, it follows that $|f|$, and hence f, are Lebesgue integrable. The relation between the integrals then also follows from the first part of the theorem.

So it suffices to prove the first part. Write $f \underset{L}{=} \sum\limits_{k=1}^{\infty} f_k$ with Riemann integrable functions f_k such that $\sum_k \int |f_k| < \infty$. We know that for each f_k, the function $x \mapsto f_k(x,y)$ is Riemann integrable for all y outside a set X_k of volume zero, that $y \mapsto \int_{\mathbb{R}^n} f_k(x,y)\,|d^n x|$ for $y \notin X_k$, $y \mapsto 0$ for $y \in X_k$ is Riemann integrable, and that

$$\int\limits_{\mathbb{R}^{n+m}} f_k(x,y)\,|d^{n+m}(x,y)| = \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} f_k(x,y)\,|d^n x| \Big)\,|d^m y|\,.$$

This gives

$$\int\limits_{\mathbb{R}^{n+m}} f(x,y)\,|d^{n+m}(x,y)| = \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} f_k(x,y)\,|d^n x| \Big)\,|d^m y|\,.$$

We have the following estimate (using Fubini for $|f_k|$).

$$\begin{aligned}\sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^m} \Big| \int\limits_{\mathbb{R}^n} f_k(x,y)\,|d^n x| \Big|\,|d^m y| &\le \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} |f_k(x,y)|\,|d^n x| \Big)\,|d^m y| \\ &= \sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^{n+m}} |f_k(x,y)|\,|d^{n+m}(x,y)| < \infty\,.\end{aligned}$$

Thm. 11.23 then allows us to conclude that

$$\sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^m} \Big(\int\limits_{\mathbb{R}^n} f_k(x,y)\,|d^n x| \Big)\,|d^m y| = \int\limits_{\mathbb{R}^m} \Big(\sum_{k=1}^{\infty} \int\limits_{\mathbb{R}^n} f_k(x,y)\,|d^n x| \Big)\,|d^m y|\,,$$

with the inner sum converging almost everywhere. Using $g = \sum_k |f_k|$ (which is also Lebesgue integrable) instead of f, we see that

$$\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x,y)|\,|d^n x|$$

converges for almost all $y \in \mathbb{R}^m$. This shows that f_y is Lebesgue integrable for these almost all y. Another application of Thm. 11.23 then leads to

$$\begin{aligned}\int_{\mathbb{R}^m} \Big(\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} f_k(x,y)\,|d^n x|\Big)\,|d^m y| &= \int_{\mathbb{R}^m} \Big(\int_{\mathbb{R}^n} \sum_{k=1}^{\infty} f_k(x,y)\,|d^n x|\Big)\,|d^m y| \\ &= \int_{\mathbb{R}^m} \Big(\int_{\mathbb{R}^n} f(x,y)\,|d^n x|\Big)\,|d^m y|\,,\end{aligned}$$

thus finishing the proof. □

Our next result is the Change-of-Variables Formula.

11.31. **Theorem.** *Let $U, V \subset \mathbb{R}^n$ be open, $\Phi : U \to V$ a $\mathcal{C}^1$-diffeomorphism, and let $f : V \to \mathbb{R}$ be a function. Then f is Lebesgue integrable on V if and only if $(f \circ \Phi)|\det D\Phi|$ is Lebesgue integrable on U, in which case we have*

$$\int_V f(v)\,|d^n v| = \int_U f\big(\Phi(u)\big)\,|\det D\Phi_u|\,|d^n u|\,.$$

Proof. We will reduce this to the corresponding statement for Riemann integrals, which was proved for sets with compact closure and nice boundary. So we write V as a countable disjoint union of dyadic cubes $C \in \mathcal{Q}$, then $f = \sum_{C\in\mathcal{Q}} f\chi_C$. Assume that f is Lebesgue integrable, hence we can write, as usual, $f \underset{L}{=} \sum_k f_k$ with Riemann integrable f_k such that $\sum_k \int |f_k| < \infty$. Then $f \underset{L}{=} \sum_{k,C} f_k\chi_C$, and

$$\sum_{k=1}^{\infty} \sum_{C\in\mathcal{Q}} \int_{\mathbb{R}^n} |f_k(x)\chi_C(x)|\,|d^n x| = \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|\,|d^n x| < \infty\,.$$

We know that each $(f_k\chi_C \circ \Phi)|\det D\Phi|$ is Riemann integrable, and the change-of-variables formula holds for $f_k\chi_C$. Note that

$$\sum_{k=1}^{\infty} \sum_{C\in\mathcal{Q}} \int_{\mathbb{R}^n} |(f_k\chi_C \circ \Phi)(x)|\,|\det D\Phi_x|\,|d^n x| = \sum_{k=1}^{\infty} \sum_{C\in\mathcal{Q}} \int_{\mathbb{R}^n} |f_k(x)\chi_C(x)|\,|d^n x| < \infty\,,$$

so $(f \circ \Phi)|\det D\Phi| \underset{L}{=} \sum_{k,C} (f_k\chi_C \circ \Phi)|\det D\Phi|$ is Lebesgue integrable, and we find that

$$\begin{aligned}\int_V f(v)\,|d^n v| &= \sum_{k=1}^{\infty} \sum_{C\in\mathcal{Q}} \int_C f_k(v)\,|d^n v| \\ &= \sum_{k=1}^{\infty} \sum_{C\in\mathcal{Q}} \int_{\Phi^{-1}(C)} f_k\big(\Phi(u)\big)\,|\det D\Phi_u|\,|d^n u| \\ &= \int_U f\big(\Phi(u)\big)\,|\det D\Phi_u|\,|d^n u|\,.\end{aligned}$$

This proves the 'only if' part of the statement. The 'if' part follows by looking at $\Phi^{-1} : V \to U$ and using the chain rule. $\square$

11.32. **Example.** The classical sample application of Fubini's Theorem and the Change-of-Variables Formula is the evaluation of the *Gaussian integral*

$$A = \int_{-\infty}^{\infty} e^{-x^2}\, dx\,.$$

The function $x \mapsto e^{-x^2}$ has no 'elementary' antiderivative, so we cannot simply use the improper Riemann integral to find the value.

The trick is to compute A^2 instead, which can be done explicitly using Fubini and polar coordinates:

$$\begin{aligned} A^2 &= \Big(\int_{\mathbb{R}} e^{-x^2}\,|dx|\Big)\Big(\int_{\mathbb{R}} e^{-y^2}\,|dy|\Big) = \int_{\mathbb{R}^2} e^{-x^2-y^2}\,|d^2(x,y)| \\ &= \int_{]0,\infty[\times]0,2\pi[} e^{-r^2}\, r\,|d^2(r,\theta)| = \int_0^{2\pi}\Big(\int_0^{\infty} re^{-r^2}\,dr\Big)\,d\theta \\ &= 2\pi \lim_{r\to\infty}\Big(\frac{1-e^{-r^2}}{2}\Big) = \pi\,. \end{aligned}$$

So $A = \sqrt{\pi}$. (Strictly speaking, we have computed the integral over $\mathbb{R}^2$ minus the nonnegative x-axis, but since we remove a set of measure zero, this has no effect on the value.)

More on Measurable Functions. We will now prove a characterization of measurable functions in terms of measurable sets. First note the following.

11.33. **Lemma.** *If (f_k) is a sequence of measurable functions that converges almost everywhere to a bounded function f, then f is measurable.*

Proof. Let $R > 0$ such that $|f| \le 2R$, then $[f_k]_R \to [f]_R$ almost everywhere. By the Dominated Convergence Theorem 11.26 (note that $|[f_k|_R| \le R\chi_{B_R(0)}$), $[f]_R$ is integrable. $\square$

As we will see later, the hypothesis that f is bounded is unnecessary.

11.34. **Proposition.** *Let f be a measurable function, $a \in \mathbb{R}$. Then the set $M_a = f^{-1}(]a,\infty[)$ is measurable.*

Proof. We can write

$$\chi_{M_a} = \lim_{k\to\infty} \min\{1, k\max\{f-a, 0\}\}\,.$$

By the preceding lemma, χ_{M_a} is measurable. $\square$

11.35. **Definition.** The *σ-algebra of Borel sets* on $\mathbb{R}^n$ is the smallest set of subsets of $\mathbb{R}^n$ that contains all open sets and is closed under taking complements and countable unions.

Note that every Borel set is measurable, but not every measurable set is Borel (not even in $\mathbb{R}$ — Exercise!). This can be used to construct a continuous function f and a measurable function g such that $g \circ f$ is *not* measurable.

11.36. **Corollary.** *Let f be measurable and $B \subset \mathbb{R}$ a Borel set. Then $f^{-1}(B)$ is measurable.*

Proof. Any Borel of $\mathbb{R}$ set can be obtained from open sets (in fact, from intervals $]a, \infty[$ by complements and countable unions. The claim follows from the preceding proposition and the fact that complements and countable unions of measurable sets are measurable. □

There is a converse to Prop. 11.34. We can combine both statements into the following.

11.37. **Proposition.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is measurable if and only if for every $a \in \mathbb{R}$, the set $f^{-1}(]a, \infty[)$ is measurable.*

Proof. We only need to show the 'if' part. It suffices to consider $f \geq 0$. (The condition on the right hand side implies the same condition for $|f|$.) So assume that $f \geq 0$ and $f^{-1}(]a, \infty[)$ is measurable for all $a \in \mathbb{R}$. Then $f^{-1}([a, b[)$ is also measurable for all $a < b$. Let $R > 0$. Then

$$[f]_R = \lim_{n\to\infty} \sum_{k=0}^{\lfloor Rn \rfloor} \frac{k}{n} \chi_{\overline{B_R(0)} \cap f^{-1}([\frac{k}{n}, \frac{k+1}{n}[)}$$

is a bounded limit of integrable functions, hence integrable. □

11.38. **Corollary.** *If f is almost everywhere the point-wise limit of a sequence (f_k) of measurable functions, then f is measurable.*

Proof. Let $a \in \mathbb{R}$. Outside a set of measure zero, we have

$$\begin{aligned} f(x) > a &\iff \exists b > a \; \exists K \; \forall k \geq K : f_k(x) > b \\ &\iff x \in \bigcup_{n=1}^{\infty} \bigcup_{K=1}^{\infty} \bigcap_{k=K}^{\infty} f_k^{-1}\Big(\big]a + \frac{1}{n}, \infty\big[\Big), \end{aligned}$$

and so $f^{-1}(]a, \infty[)$ is measurable. By the preceding proposition, this implies that f is measurable. □

12. L^p Spaces

The theory of the Lebesgue integral is indispensable for the introduction of L^p spaces, which are very important function spaces used in Real and Functional Analysis. We first recall the following definition.

12.1. **Definition.** A *norm* on a real vector space V is a function $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ with the following properties.

1. For all $x \in V$, $\|x\| = 0$ if and only if $x = 0$.
2. For all $x, y \in V$, $\|x + y\| \leq \|x\| + \|y\|$.
3. For all $\lambda \in \mathbb{R}$, $x \in V$, $\|\lambda x\| = |\lambda| \, \|x\|$.

This implies that $d(x, y) = \|x - y\|$ is a metric on V, hence V can be considered as a metric space. A pair $(V, \|\cdot\|)$, where $\|\cdot\|$ is a norm on V, is called a *normed space*.

12.2. **Definition.** A *Banach space* is a normed space that is complete as a metric space. (This means that every Cauchy sequence has a limit.)

We will now define the L^p spaces with their norms. Our goal will then be to show that they are Banach spaces.

12.3. **Definition.** Let $X \subset \mathbb{R}^n$ be a measurable set, not of measure zero. Let $Z(X)$ be the set of all functions f on X such that $f \underset{L}{=} 0$. If $1 \le p < \infty$, we set

$$L^p(X) = \Big\{ f : X \to \mathbb{R} \text{ measurable } \Bigm| \int_X |f|^p < \infty \Big\} \,\big/\, Z(X)\,,$$

and for $f \in L^p(X)$, we set

$$\|f\|_p = \Big(\int_X |f(x)|^p \, |dx| \Big)^{1/p} .$$

(Note that the integral does not depend on the representative function.) It is easy to see that $L^p(X)$ is a vector space.

In addition, we define

$$L^\infty(X) = \{ f : X \to \mathbb{R} \text{ measurable } \mid |f| \underset{L}{\le} M \text{ for some } M > 0 \} \,\big/\, Z(X)\,,$$

and for $f \in L^\infty(X)$, we set

$$\|f\|_\infty = \inf\{ M \in \mathbb{R} : |f| \underset{L}{\le} M \} .$$

Note that for all $1 \le p \le \infty$, we then have that

$$\|f\|_p = 0 \iff f = 0 \quad \text{and} \quad \|\lambda f\|_p = |\lambda| \, \|f\|_p .$$

"$f = 0$" refers to the quotient space; it means "$f \underset{L}{=} 0$" for any representative function.

In order to see that we have really defined a norm, we need to prove the triangle inequality.

12.4. **Minkowski's Inequality.** *Let $1 \le p \le \infty$, $f, g \in L^p(X)$. Then we have $\|f+g\|_p \le \|f\|_p + \|g\|_p$.*

Proof. This is clear when $p = \infty$, or when f or g are zero. So we can assume that $1 \le p < \infty$ and that $\|f\|_p = \alpha > 0$ and $\|g\|_p = \beta > 0$. Then we can write

$$|f| = \alpha f_0 \,, \quad |g| = \beta g_0$$

with $f_0, g_0 \ge 0$ and $\|f_0\|_p = \|g_0\|_p = 1$. Set $\lambda = \alpha/(\alpha+\beta)$. We obtain

$$\begin{aligned} |f+g|^p &\le \big(|f| + |g|\big)^p = (\alpha f_0 + \beta g_0)^p = (\alpha+\beta)^p \big(\lambda f_0 + (1-\lambda) g_0\big)^p \\ &\le (\alpha+\beta)^p \big(\lambda f_0^p + (1-\lambda) g_0^p\big) . \end{aligned}$$

In the last inequality, we have used the fact that the function $t \mapsto t^p$ is convex (here we need that $p \ge 1$). If we now integrate, we find

$$\|f+g\|_p^p \le (\alpha+\beta)^p \big(\lambda \|f_0\|_p + (1-\lambda)\|g_0\|_p\big) = \big(\|f\|_p + \|g\|_p\big)^p ,$$

since $\|f_0\|_p = \|g_0\|_p = 1$. □

We conclude that $(L^p(X), \|\cdot\|_p)$ is a normed space for all $1 \le p \le \infty$.

There is another important inequality that relates norms for different p.

12.5. **Hölder's Inequality.** *Let* $1 \le p, q \le \infty$ *such that* $\dfrac{1}{p} + \dfrac{1}{q} = 1$. *Let* $f \in L^p(X)$, $g \in L^q(X)$. *Then* $fg \in L^1(X)$, *and*

$$\|fg\|_1 = \int_X |f(x)g(x)|\,|dx| \le \|f\|_p \, \|g\|_q \,.$$

Note that for $p = 1$, $q = \infty$, this specializes to the standard estimate for an integral, and for $p = q = 2$, it specializes to the Cauchy-Schwarz inequality.

Proof. The case $p = 1$, $q = \infty$ is easy, so we can assume that $1 < p, q < \infty$. Also, we can assume that $f, g \ge 0$. Set $h = g^{q-1} = g^{q/p}$; then $g = h^{p/q} = h^{p-1}$. For $t \ge 0$, we have

$$ptfg = ptfh^{p-1} \le (h + tf)^p - h^p \,;$$

the latter comes from the standard inequality $(1 + x)^p \ge 1 + px$, valid for $x \ge -1$ and $p \ge 1$. Integrating, we find (using Minkowski's inequality)

$$pt\|fg\|_1 \le \|h + tf\|_p^p - \|h\|_p^p \le \big(\|h\|_p + t\|f\|_p\big)^p - \|h\|_p^p \,.$$

We have equality at $t = 0$. Dividing by pt, and letting $t \searrow 0$, we get

$$\|fg\|_1 \le \|f\|_p \, \|h\|_p^{p-1} = \|f\|_p \, \|g\|_q \,.$$

For this last equality, note that

$$\|h\|_p^{p-1} = \Big(\int_X h^p\Big)^{(p-1)/p} = \Big(\int_X g^q\Big)^{1/q} = \|g\|_q \,.$$

□

The importance of this inequality comes from the fact that it tells us that every $g \in L^q(X)$ provides us with a *continuous linear functional* on $L^p(X)$ via

$$L^p(X) \ni f \longmapsto \int_X f(x)g(g)\,|dx| \in \mathbb{R} \,.$$

Note that a linear functional (or linear form) ϕ on a normed space V is continuous if and only if it is *bounded*, i.e., if there is $M \ge 0$ such that $|\phi(v)| \le M \, \|v\|$ for all $v \in V$. The bound here is given by $\|g\|_q$, according to Hölder's inequality.

12.6. **Theorem (Fischer-Riesz).** $L^p(X)$ *is a Banach space.*

Proof. We have seen that $(L^p(X), \| \cdot \|_p)$ is a normed vector space. We have to show that it is complete. So let (f_n) be a Cauchy sequence in $L^p(X)$. It suffices to show that there is a convergent subsequence. We pick a subsequence (f_{n_k}) such that $\|f_{n_{k+1}} - f_{n_k}\|_p \le 2^{-k}$ for all k.

If $p = \infty$, this means that $|f_{n_{k+1}} - f_{n_k}| \underset{L}{\le} 2^{-k}$ for all k, hence (f_{n_k}) converges point-wise almost everywhere to a function f, and $|f| \underset{L}{\le} \|f_{n_1}\|_\infty + \sum_k 2^{-k} < \infty$, so $f \in L^\infty(X)$. Also, $\|f - f_{n_k}\|_\infty \le 2 \cdot 2^{-k}$, so (f_{n_k}) converges to f in the metric of $L^\infty(X)$.

Now assume that $1 \le p < \infty$. We have

$$|f_{n_k}| \underset{L}{\le} g_k = |f_{n_1}| + |f_{n_2} - f_{n_1}| + \cdots + |f_{n_k} - f_{n_{k-1}}| \,,$$

and by Minkowski's inequality,

$$\|g_k\|_p \le \|f_{n_1}\|_p + \|f_{n_2} - f_{n_1}\|_p + \cdots + \|f_{n_k} - f_{n_{k-1}}\|_p \le \|f_{n_1}\|_p + 1 \, .$$

The sequence of functions (g_k) is increasing, and $\int_X g_k^p$ is bounded, hence by the Monotone Convergence Theorem 11.24, (g_k) converges point-wise almost everywhere to a limit function g, and g^p is integrable on X. We therefore see that $|f_{n_k}|^p \underset{L}{\le} g^p$, so by the Dominated Convergence Theorem 11.26, the sequence (f_{n_k}) converges point-wise almost everywhere to a function f such that $|f|^p$ is integrable on X, so $f \in L^p(X)$. The sequence $|f - f_{n_k}|^p$ then converges point-wise to zero almost everywhere and is bounded by the integrable function $2^p g^p$, so by the Dominated Convergence Theorem again,

$$\|f - f_{n_k}\|_p = \int_X \left|f(x) - f_{n_k}(x)\right|^p |dx| \to 0 \qquad \text{as } k \to \infty.$$

This means that f_{n_k} converges to f in the metric of $L^p(X)$, which was to be shown. □

12.7. **Hilbert Space.** In particular, $L^2(X)$ is a *Hilbert space*. This is a Banach space such that the norm comes from an inner product: $\|x\| = \sqrt{\langle x, x\rangle}$. In the case of $L^2(X)$, the inner product is

$$\langle f, g\rangle = \int_X f(x)g(x) \, |dx| \, .$$

Note that Hölder's inequality tells us that this makes sense, i.e., fg is integrable on X when $f, g \in L^2(X)$.

Recall that a family (f_j) is said to be *orthonormal* if $\langle f_i, f_j\rangle = 0$ whenever $i \ne j$ and $\|f_j\|_2 = 1$ for all j. By Zorn's Lemma, there are maximal orthonormal families. It can be shown that $L^2(X)$ is *separable*, i.e., it has a countable dense subset. This implies that every orthonormal family must be countable (otherwise it would provide an uncountable discrete subset), so we can write it as $(f_n)_{n\in\mathbb{N}}$. If the family is maximal, it is then true that

$$f = \sum_{n=1}^{\infty} \langle f, f_n\rangle f_n$$

for all $f \in L^2(X)$, in the sense of convergence in the metric of $L^2(X)$:

$$\lim_{n\to\infty} \left\| f - \sum_{k=1}^{n} f_k \right\|_2 = 0 \, .$$

Note that we have $\sum_n \langle f, f_n\rangle^2 = \|f\|_2^2 < \infty$. Conversely, if (a_n) is a square-summable sequence of real numbers, i.e., $\sum_n a_n^2 < \infty$, then $\sum_n a_n f_n$ converges. In this way, we obtain an isomorphism between $L^2(X)$ and the Hilbert space ℓ^2 of square-summable sequences.

For example, when X is the unit interval $[0, 1]$, then the constant function 1, together with $\sqrt{2}\sin 2\pi nx$ and $\sqrt{2}\cos 2\pi nx$ for $n \ge 1$ are a maximal orthonormal family, and the expression for f given above is nothing else than its *Fourier series*.

Let $F : L^2(X) \to \mathbb{R}$ be a bounded (i.e., continuous) linear functional (so $|F(f)| \le M\|f\|_2$ for some $M \ge 0$). Then we have

$$F(f) = F\Big(\sum_n \langle f, f_n\rangle f_n\Big) = \sum_n \langle f, f_n\rangle F(f_n) = \Big\langle f, \sum_n F(f_n) f_n \Big\rangle .$$

For any n, we have

$$\sum_{k=1}^n F(f_k)^2 = F\Big(\sum_{k=1}^n F(f_k) f_k\Big) \le M \Big\|\sum_{k=1}^n F(f_k) f_k\Big\|_2 = M \sqrt{\sum_{k=1}^n F(f_k)^2}\,,$$

so $\displaystyle\sum_{n=1}^\infty F(f_n)^2 \le M^2 < \infty$. This implies that

$$g = \sum_n F(f_n) f_n$$

exists in $L^2(X)$, and $F(f) = \langle f, g\rangle$ for all $f \in L^2(X)$:
Every bounded linear functional on $L^2(X)$ is obtained as inner product with some fixed element $g \in L^2(X)$.

This is a special case of the following more general result.

12.8. **Riesz Representation Theorem.** *Let $1 \le p < \infty$, and let $F : L^p(X) \to \mathbb{R}$ be a bounded linear functional. Then there is a unique $g \in L^q(X)$ such that $F(f) = \int_X f(x)g(x)\,|dx|$ for all $f \in L^p(X)$.*

The proof is beyond the scope of this course. See the Real Analysis and Functional Analysis graduate courses.

Note that $p = \infty$ is excluded: there are more bounded linear functionals on $L^\infty(X)$ than those coming from integrable functions. Here is a sketch. One version of the *Hahn-Banach Theorem* (which relies on the axiom of choice) says that a bounded linear functional on a closed subspace of a Banach space can be extended to a bounded linear functional on the whole space. Let $X = [-1, 1]$ (for concreteness). The space $\mathcal{C}(X)$ of continuous functions on X, together with the maximum norm, is a Banach space (Exercise!), which can be identified with a closed subspace of $L^\infty(X)$. On $\mathcal{C}(X)$, there is the bounded linear functional $f \mapsto f(0)$, which we then can extend to a bounded linear functional on $L^\infty(X)$. If this were represented by a function $g \in L^1(X)$, then we would need to have

$$\int_{-1}^{1} f(x)g(x)\,dx = f(0) \qquad \text{for all } f \in \mathcal{C}(X).$$

But such a function g cannot exist (Exercise — use that continuous functions are dense in $L^1(X)$).

References

[HH] J.H. Hubbard and B. Burke Hubbard: *Vector calculus, linear algebra and differential forms: a unified approach.* Prentice Hall, 2002.

[R] H.L. Royden: *Real analysis*, third edition. Macmillan Publishing Co., New York, 1988.

[S] M. Spivak: *Calculus on manifolds.* W.A. Benjamin, 1965.

[W] F.W. Warner: *Foundations of differentiable manifolds and Lie groups.* Springer Graduate Texts in Mathematics **94**, 1983.