Introduction to Analysis
in Several Variables
(Advanced Calculus)

Michael E. Taylor
Contents

0. One-variable calculus
1. The derivative
2. Inverse function and implicit function theorem
3. Systems of differential equations
4. The Riemann integral in $n$ variables
5. Surfaces and surface integrals
6. Differential forms
7. Products and exterior derivatives of forms
8. The general Stokes formula
9. The classical Gauss, Green, and Stokes formulas
10. Holomorphic functions and harmonic functions
11. Homotopies of maps and actions on forms
12. Differential forms and degree theory
13. Fourier series
14. The Fourier transform

A. Metric spaces, convergence, and compactness
B. Partitions of unity
C. Differential forms and the change of variable formula
D. Complements on power series
E. The Weierstrass theorem and the Stone-Weierstrass theorem
F. Sard’s theorem
G. Morse functions
H. Inner product spaces
Introduction

This text was produced for the second semester of a two-semester sequence on advanced calculus, whose aim is to provide a firm logical foundation for analysis, for students who have had 3 semesters of calculus and a course in linear algebra. The first semester treats analysis in one variable, and the text [T4] was written to cover that material. The text at hand treats analysis in several variables. These two texts can be used as companions, but they are written so that they can be used independently, if desired.

We begin with an introductory section, §0, presenting the elements of one-dimensional calculus. We first define the Riemann integral of a class of functions on an interval. We then introduce the derivative, and establish the Fundamental Theorem of Calculus, relating differentiation and integration as essentially inverse operations. Further results are dealt with in the exercises, such as the change of variable formula for integrals, and the Taylor formula with remainder. This section distills material developed in more detail in the companion text [T4]. We have included it here to facilitate the independent use of this text.

In §1 we define the derivative of a function \( F : \mathcal{O} \to \mathbb{R}^m \), where \( \mathcal{O} \) is an open subset of \( \mathbb{R}^n \), as a linear map from \( \mathbb{R}^n \) to \( \mathbb{R}^m \). We establish some basic properties, such as the chain rule. We use the one-dimensional integral as a tool to show that, if the matrix of first order partial derivatives of \( F \) is continuous on \( \mathcal{O} \), then \( F \) is differentiable on \( \mathcal{O} \). We also discuss two convenient multi-index notations for higher derivatives, and derive the Taylor formula with remainder for a smooth function \( F \) on \( \mathcal{O} \subset \mathbb{R}^n \).

In §2 we establish the Inverse Function Theorem, stating that a smooth map \( F : \mathcal{O} \to \mathbb{R}^n \) with an invertible derivative \( DF(p) \) has a smooth inverse defined near \( q = F(p) \). We derive the Implicit Function Theorem as a consequence of this. As a tool in proving the Inverse Function Theorem, we use a fixed point theorem known as the Contraction Mapping Principle.

In §3 we treat systems of differential equations. We establish a basic existence and uniqueness theorem and also study the smooth dependence of a solution on initial data. We interpret the solution operator as a flow generated by a vector field and introduce the concept of the Lie bracket of vector fields.

In §4 we take up the multidimensional Riemann integral. The basic definition is quite parallel to the one-dimensional case, but a number of central results, while parallel in statement to the one-dimensional case, require more elaborate demonstrations in higher dimensions. This section is one of the most demanding in this text, and it is in a sense the heart of the course. One central result is the change of variable formula for multidimensional integrals. Another is the reduction of multiple integrals to iterated integrals.

In §5 we define the notion of a smooth \( m \)-dimensional surface in \( \mathbb{R}^n \) and study properties of these objects. We associate to such a surface a “metric tensor,” and make use of this to define the integral of functions on a surface. This includes the study of surface area. Examples include the computation of areas of higher dimensional spheres. We also explore
integration on the group of rotations on $\mathbb{R}^n$, leading to the notion of “averaging over rotations.”

In §6 we introduce a further class of objects that are defined on surfaces, differential forms. A $k$-form can be integrated over a $k$-dimensional surface, endowed with an extra piece of structure, an “orientation.” The change of variable formula established in §4 plays a central role in establishing this. Properties of differential forms are developed further in the next two sections. In §7 we define exterior products of forms, and interior products of forms with vector fields. Then we define the exterior derivative of a form. Section 8 is devoted to the general Stokes formula, an important integral identity which contains as special cases classical identities of Green, Gauss, and Stokes. These special cases are discussed in some detail in §9.

In §10 we use Green’s Theorem to derive fundamental properties of holomorphic functions of a complex variable. Sprinkled throughout earlier sections are some allusions to functions of complex variables, particularly in some of the exercises in §§1–2. Readers with no previous exposure to complex variables might wish to return to these exercises after getting through §10. In this section, we also discuss some results on the closely related study of harmonic functions. One result is Liouville’s Theorem, stating that a bounded harmonic function on all of $\mathbb{R}^n$ must be constant. When specialized to holomorphic functions on $\mathbb{C} = \mathbb{R}^2$, this yields a proof of the Fundamental Theorem of Algebra.

In §11 we define the notion of smoothly homotopic maps and consider the behavior of closed differential forms under pull back by smoothly homotopic maps. This material is then applied in §12, which introduces degree theory and derives some interesting consequences. Applications include the Brouwer fixed point theorem, the non-existence of smooth vector fields tangent to an even-dimensional sphere, and the Jordan-Brouwer separation theorem (in the smooth case). We also show how degree theory yields another proof of the Fundamental Theorem of Algebra.

Section 13 treats Fourier series on the $n$-dimensional torus $T^n$, and §14 treats the Fourier transform for functions on $\mathbb{R}^n$. These two sections are written to be read independently, for the benefit of those who decide to cover only one of these topics. Those who read both sections will therefore notice a bit of repetition, hopefully not too annoying. For those who do cover both sections, §14 ends with a topic that ties them together, known as Poisson’s summation formula. We apply this formula to establish a classical result of Riemann, his functional equation for the Riemann zeta function.

Appendix A at the end of this text covers some basic notions of metric spaces and compactness, used from time to time throughout the text, particularly in the study of the Riemann integral and in the proof of the fundamental existence theorem for ODE. As is the case with §0, this appendix distills material developed at a more leisurely pace in [T4], again serving to make this text independent of the first one.

Further appendices provide additional material, useful as either background or complements to the main body of the text. Appendix B discusses partitions of unity, useful particularly in the proof of the Stokes formula. Appendix C presents a proof of the change of variable formula for multiple integrals, following an idea presented by Lax in [L], which is somewhat shorter than that given in §4. Appendix D discusses the remainder term in the Taylor series of a function. Appendix E gives the Weierstrass theorem on approximating
a continuous function by polynomials, and an extension, known as the Stone-Weierstrass theorem, which is a useful tool in analysis. Applications arise in sections 12 and 13. Appendix F proves Sard’s theorem, of use in the presentation of degree theory in §12. Appendix G gives another application of Sard’s theorem, to the existence of lots of Morse functions, which in turn implies the existence of lots of vector fields with only isolated critical points, also of use in §12. Appendix H gives some basic material on inner product spaces, which arise in sections 13–14.

We mention that the exercises in this text play a particularly important role in developing the material. In addition to exercises providing practice in applying results established in the text, there are also exercises asking the reader to supply details for some of the arguments used in the text. There are also scattered exercises intended to give the reader a fresh perspective on topics familiar from previous study. We mention in particular exercises on determinants and on the cross product in §1, and on trigonometric functions in §3.
0. One variable calculus

In this brief discussion of one variable calculus, we introduce the Riemann integral, and relate it to the derivative. We will define the Riemann integral of a bounded function over an interval $I = [a, b]$ on the real line. For now, we assume $f$ is real valued. To start, we partition $I$ into smaller intervals. A partition $\mathcal{P}$ of $I$ is a finite collection of subintervals \{\(J_k : 0 \leq k \leq N\}\}, disjoint except for their endpoints, whose union is $I$. We can order the $J_k$ so that $J_k = [x_k, x_{k+1}]$, where

(0.1) \hspace{1cm} x_0 < x_1 < \cdots < x_N < x_{N+1}, \hspace{0.5cm} x_0 = a, \hspace{0.5cm} x_{N+1} = b.

We call the points $x_k$ the endpoints of $\mathcal{P}$. We set

(0.2) \hspace{1cm} \ell(J_k) = x_{k+1} - x_k, \hspace{0.5cm} \text{maxsize}(\mathcal{P}) = \max_{0 \leq k \leq N} \ell(J_k)

We then set

(0.3) \hspace{1cm} T_\mathcal{P}(f) = \sum_k \sup_{J_k} f(x) \ell(J_k), \hspace{1cm} L_\mathcal{P}(f) = \sum_k \inf_{J_k} f(x) \ell(J_k).

See (A.10)–(A.11) for the definition of sup and inf. We call $T_\mathcal{P}(f)$ and $L_\mathcal{P}(f)$ respectively the upper sum and lower sum of $f$, associated to the partition $\mathcal{P}$. Note that $L_\mathcal{P}(f) \leq T_\mathcal{P}(f)$. These quantities should approximate the Riemann integral of $f$, if the partition $\mathcal{P}$ is sufficiently “fine.”

To be more precise, if $\mathcal{P}$ and $\mathcal{Q}$ are two partitions of $I$, we say $\mathcal{P}$ refines $\mathcal{Q}$, and write $\mathcal{P} \succ \mathcal{Q}$, if $\mathcal{P}$ is formed by partitioning each interval in $\mathcal{Q}$. Equivalently, $\mathcal{P} \succ \mathcal{Q}$ if and only if all the endpoints of $\mathcal{Q}$ are also endpoints of $\mathcal{P}$. It is easy to see that any two partitions have a common refinement; just take the union of their endpoints, to form a new partition. Note also that refining a partition lowers the upper sum of $f$ and raises its lower sum:

(0.4) \hspace{1cm} \mathcal{P} \succ \mathcal{Q} \implies T_\mathcal{P}(f) \leq T_\mathcal{Q}(f), \hspace{0.5cm} L_\mathcal{P}(f) \geq L_\mathcal{Q}(f).

Consequently, if $\mathcal{P}_1$ are any two partitions and $\mathcal{Q}$ is a common refinement, we have

(0.5) \hspace{1cm} L_{\mathcal{P}_1}(f) \leq L_\mathcal{Q}(f) \leq T_\mathcal{Q}(f) \leq T_{\mathcal{P}_2}(f).

Now, whenever $f : I \to \mathbb{R}$ is bounded, the following quantities are well defined:

(0.6) \hspace{1cm} \bar{I}(f) = \inf_{\mathcal{P} \in \Pi(I)} T_\mathcal{P}(f), \hspace{0.5cm} \underline{I}(f) = \sup_{\mathcal{P} \in \Pi(I)} L_\mathcal{P}(f),
where \( \Pi(I) \) is the set of all partitions of \( I \). We call \( \underline{I}(f) \) the lower integral of \( f \) and \( \overline{I}(f) \) its upper integral. Clearly, by (0.5), \( \underline{I}(f) \leq \overline{I}(f) \). We then say that \( f \) is Riemann integrable provided \( \overline{I}(f) = \underline{I}(f) \), and in such a case, we set

\[
\int_{a}^{b} f(x) \, dx = \int_{I} f(x) \, dx = \underline{I}(f) = \overline{I}(f).
\]

We will denote the set of Riemann integrable functions on \( I \) by \( \mathcal{R}(I) \).

We derive some basic properties of the Riemann integral.

**Proposition 0.1.** If \( f, g \in \mathcal{R}(I) \), then \( f + g \in \mathcal{R}(I) \), and

\[
\int_{I} (f + g) \, dx = \int_{I} f \, dx + \int_{I} g \, dx.
\]

**Proof.** If \( J_k \) is any subinterval of \( I \), then

\[
\sup_{J_k} (f + g) \leq \sup_{J_k} f + \sup_{J_k} g, \quad \text{and} \quad \inf_{J_k} (f + g) \geq \inf_{J_k} f + \inf_{J_k} g,
\]

so, for any partition \( \mathcal{P} \), we have \( \overline{I}_\mathcal{P}(f + g) \leq \overline{I}_\mathcal{P}(f) + \overline{I}_\mathcal{P}(g) \). Also, using common refinements, we can simultaneously approximate \( \overline{I}(f) \) and \( \overline{I}(g) \) by \( \overline{I}_\mathcal{P}(f) \) and \( \overline{I}_\mathcal{P}(g) \), and ditto for \( \overline{I}(f + g) \). Thus the characterization (0.6) implies \( \overline{I}(f + g) \geq \underline{I}(f) + \underline{I}(g) \), and the proposition follows.

Next, there is a fair supply of Riemann integrable functions.

**Proposition 0.2.** If \( f \) is continuous on \( I \), then \( f \) is Riemann integrable.

**Proof.** Any continuous function on a compact interval is bounded and uniformly continuous (see Propositions A.14 and A.15). Let \( \omega(\delta) \) be a modulus of continuity for \( f \), so

\[
|x - y| \leq \delta \implies |f(x) - f(y)| \leq \omega(\delta), \quad \omega(\delta) \to 0 \quad \text{as} \quad \delta \to 0.
\]

Then

\[
\max \text{size}(\mathcal{P}) \leq \delta \implies \overline{I}_\mathcal{P}(f) - \underline{I}_\mathcal{P}(f) \leq \omega(\delta) \cdot \ell(I),
\]

which yields the proposition.

We denote the set of continuous functions on \( I \) by \( C(I) \). Thus Proposition 0.2 says

\[
C(I) \subset \mathcal{R}(I).
\]

The proof of Proposition 0.2 provides a criterion on a partition guaranteeing that \( \overline{I}_\mathcal{P}(f) \) and \( \underline{I}_\mathcal{P}(f) \) are close to \( \int_{I} f \, dx \) when \( f \) is continuous. We produce an extension, giving a condition under which \( \overline{I}_\mathcal{P}(f) \) and \( \overline{I}(f) \) are close, and \( \underline{I}_\mathcal{P}(f) \) and \( \underline{I}(f) \) are close, given \( f \) bounded on \( I \). Given a partition \( \mathcal{P}_0 \) of \( I \), set

\[
\min \text{size}(\mathcal{P}_0) = \min\{\ell(J_k) : J_k \in \mathcal{P}_0\}.
\]
Lemma 0.3. Let $P$ and $Q$ be two partitions of $I$. Assume

\[ \text{maxsize}(P) \leq \frac{1}{k} \text{minsize}(Q). \]  

Let $|f| \leq M$ on $I$. Then

\[ I_P(f) \leq I_Q(f) + \frac{2M}{k} \ell(I), \]

\[ L_P(f) \geq L_Q(f) - \frac{2M}{k} \ell(I). \]  

Proof. Let $P_1$ denote the minimal common refinement of $P$ and $Q$. Consider on the one hand those intervals in $P$ that are contained in intervals in $Q$ and on the other hand those intervals in $P$ that are not contained in intervals in $Q$. Each interval of the first type is also an interval in $P_1$. Each interval of the second type gets partitioned, to yield two intervals in $P_1$. Denote by $P_1^b$ the collection of such divided intervals. By (0.12), the lengths of the intervals in $P_1^b$ sum to $\leq \ell(I)/k$. It follows that

\[ |I_P(f) - I_{P_1}(f)| \leq \sum_{J \in P_1^b} 2M\ell(J) \leq 2M\frac{\ell(I)}{k}, \]

and similarly $|L_P(f) - L_{P_1}(f)| \leq 2M\ell(I)/k$. Therefore

\[ I_P(f) \leq I_{P_1}(f) + \frac{2M}{k} \ell(I), \quad L_P(f) \geq L_{P_1}(f) - \frac{2M}{k} \ell(I). \]

Since also $I_{P_1}(f) \leq I_Q(f)$ and $L_{P_1}(f) \geq L_Q(f)$, we obtain (0.13).

The following consequence is sometimes called Darboux’s Theorem.

Theorem 0.4. Let $P_\nu$ be a sequence of partitions of $I$ into $\nu$ intervals $J_{\nu k}, \ 1 \leq k \leq \nu$, such that

\[ \text{maxsize}(P_\nu) \longrightarrow 0. \]

If $f : I \to \mathbb{R}$ is bounded, then

\[ \bar{I}_{P_\nu}(f) \to \bar{I}(f) \quad \text{and} \quad \underline{I}_{P_\nu}(f) \to \underline{I}(f). \]  

Consequently,

\[ f \in \mathcal{R}(I) \iff \bar{I}(f) = \lim_{\nu \to \infty} \sum_{k=1}^\nu f(\xi_{\nu k})\ell(J_{\nu k}), \]

for arbitrary $\xi_{\nu k} \in J_{\nu k}$, in which case the limit is $\int_I f \, dx$. 

Proof. As before, assume $|f| \leq M$. Pick $\varepsilon = 1/k > 0$. Let $Q$ be a partition such that

$$
\mathcal{T}(f) \leq \mathcal{T}_Q(f) \leq \mathcal{T}(f) + \varepsilon,
$$

$$
\mathcal{I}(f) \geq \mathcal{I}_Q(f) \geq \mathcal{I}(f) - \varepsilon.
$$

Now pick $N$ such that

$$
\nu \geq N \implies \max \mathcal{P}_\nu \leq \varepsilon \min \mathcal{Q}.
$$

Lemma 0.3 yields, for $\nu \geq N$,

$$
\mathcal{T}_{\mathcal{P}_{\nu}}(f) \leq \mathcal{T}_Q(f) + 2M\ell(I)\varepsilon,
$$

$$
\mathcal{I}_{\mathcal{P}_{\nu}}(f) \geq \mathcal{I}_Q(f) - 2M\ell(I)\varepsilon.
$$

Hence, for $\nu \geq N$,

$$
\mathcal{T}(f) \leq \mathcal{T}_{\mathcal{P}_{\nu}}(f) \leq \mathcal{T}(f) + [2M\ell(I) + 1]\varepsilon,
$$

$$
\mathcal{I}(f) \geq \mathcal{I}_{\mathcal{P}_{\nu}}(f) \geq \mathcal{I}(f) - [2M\ell(I) + 1]\varepsilon.
$$

This proves (0.14).

Remark. The sums on the right side of (0.15) are called Riemann sums, approximating $\int_I f \, dx$ (when $f$ is Riemann integrable).

Remark. A second proof of Proposition 0.1 can readily be deduced from Theorem 0.4.

One should be warned that, once such a specific choice of $\mathcal{P}_\nu$ and $\xi_{\nu k}$ has been made, the limit on the right side of (0.15) might exist for a bounded function $f$ that is not Riemann integrable. This and other phenomena are illustrated by the following example of a function which is not Riemann integrable. For $x \in I$, set

$$
\vartheta(x) = 1 \text{ if } x \in \mathbb{Q}, \quad \vartheta(x) = 0 \text{ if } x \notin \mathbb{Q},
$$

where $\mathbb{Q}$ is the set of rational numbers. Now every interval $J \subset I$ of positive length contains points in $\mathbb{Q}$ and points not in $\mathbb{Q}$, so for any partition $\mathcal{P}$ of $I$ we have $\mathcal{T}_\mathcal{P}(\vartheta) = \ell(I)$ and $\mathcal{I}_\mathcal{P}(\vartheta) = 0$, hence

$$
\mathcal{T}(\vartheta) = \ell(I), \quad \mathcal{I}(\vartheta) = 0.
$$

Note that, if $\mathcal{P}_\nu$ is a partition of $I$ into $\nu$ equal subintervals, then we could pick each $\xi_{\nu k}$ to be rational, in which case the limit on the right side of (0.15) would be $\ell(I)$, or we could pick each $\xi_{\nu k}$ to be irrational, in which case this limit would be zero. Alternatively, we could pick half of them to be rational and half to be irrational, and the limit would be $\ell(I)/2$. 

Associated to the Riemann integral is a notion of size of a set \( S \), called *content*. If \( S \) is a subset of \( I \), define the “characteristic function”

\[
\chi_S(x) = 1 \text{ if } x \in S, \ 0 \text{ if } x \notin S.
\]

We define “upper content” \( \text{cont}^+ \) and “lower content” \( \text{cont}^- \) by

\[
\text{cont}^+(S) = T(\chi_S), \quad \text{cont}^-(S) = I(\chi_S).
\]

We say \( S \) “has content,” or “is contented” if these quantities are equal, which happens if and only if \( \chi_S \in \mathcal{R}(I) \), in which case the common value of \( \text{cont}^+(S) \) and \( \text{cont}^-(S) \) is

\[
m(S) = \int_I \chi_S(x) \, dx.
\]

It is easy to see that

\[
\text{cont}^+(S) = \inf \left\{ \sum_{k=1}^N \ell(J_k) : S \subset J_1 \cup \cdots \cup J_N \right\},
\]

where \( J_k \) are intervals. Here, we require \( S \) to be in the union of a finite collection of intervals.

See the appendix at the end of this section for a generalization of Proposition 0.2, giving a sufficient condition for a bounded function to be Riemann integrable on \( I \), in terms of the upper content of its set of discontinuities.

There is a more sophisticated notion of the size of a subset of \( I \), called Lebesgue measure. The key to the construction of Lebesgue measure is to cover a set \( S \) by a countable (either finite or infinite) set of intervals. The *outer measure* of \( S \subset I \) is defined by

\[
m^*(S) = \inf \left\{ \sum_{k \geq 1} \ell(J_k) : S \subset \bigcup_{k \geq 1} J_k \right\}.
\]

Here \( \{ J_k \} \) is a finite or countably infinite collection of intervals. Clearly

\[
m^*(S) \leq \text{cont}^+(S).
\]

Note that, if \( S = I \cap \mathbb{Q} \), then \( \chi_S = \vartheta \), defined by (0.16). In this case it is easy to see that \( \text{cont}^+(S) = \ell(I) \), but \( m^*(S) = 0 \). Zero is the “right” measure of this set. More material on the development of measure theory can be found in a number of books, including [Fol] and [T2].

It is useful to note that \( \int_I f \, dx \) is additive in \( I \), in the following sense.
Proposition 0.5. If $a < b < c$, $f : [a, c] \to \mathbb{R}$, $f_1 = f|_{[a, b]}$, $f_2 = f|_{[b, c]}$, then

\begin{equation}
 f \in \mathcal{R}([a, c]) \iff f_1 \in \mathcal{R}([a, b]) \quad \text{and} \quad f_2 \in \mathcal{R}([b, c]),
\end{equation}

and, if this holds,

\begin{equation}
 \int_a^c f \, dx = \int_a^b f_1 \, dx + \int_b^c f_2 \, dx.
\end{equation}

Proof. Since any partition of $[a, c]$ has a refinement for which $b$ is an endpoint, we may as well consider a partition $P = P_1 \cup P_2$, where $P_1$ is a partition of $[a, b]$ and $P_2$ is a partition of $[b, c]$. Then

\begin{equation}
 I_P(f) = I_{P_1}(f_1) + I_{P_2}(f_2), \quad I_P(f) = I_{P_1}(f_1) + I_{P_2}(f_2),
\end{equation}

so

\begin{equation}
 I_P(f) - I_P(f) = \{I_{P_1}(f_1) - I_{P_1}(f_1)\} + \{I_{P_2}(f_2) - I_{P_2}(f_2)\}.
\end{equation}

Since both terms in braces in (0.27) are $\geq 0$, we have equivalence in (0.24). Then (0.25) follows from (0.26) upon taking sufficiently fine partitions.

Let $I = [a, b]$. If $f \in \mathcal{R}(I)$, then $f \in \mathcal{R}([a, x])$ for all $x \in [a, b]$, and we can consider the function

\begin{equation}
 g(x) = \int_a^x f(t) \, dt.
\end{equation}

If $a \leq x_0 \leq x_1 \leq b$, then

\begin{equation}
 g(x_1) - g(x_0) = \int_{x_0}^{x_1} f(t) \, dt,
\end{equation}

so, if $|f| \leq M$,

\begin{equation}
 |g(x_1) - g(x_0)| \leq M|x_1 - x_0|.
\end{equation}

In other words, if $f \in \mathcal{R}(I)$, then $g$ is Lipschitz continuous on $I$.

A function $g : (a, b) \to \mathbb{R}$ is said to be differentiable at $x \in (a, b)$ provided there exists the limit

\begin{equation}
 \lim_{h \to 0} \frac{1}{h} [g(x + h) - g(x)] = g'(x).
\end{equation}

When such a limit exists, $g'(x)$, also denoted $dg/dx$, is called the derivative of $g$ at $x$. Clearly $g$ is continuous wherever it is differentiable.

The next result is part of the Fundamental Theorem of Calculus.
Theorem 0.6. If \( f \in C([a, b]) \), then the function \( g \), defined by (2.28), is differentiable at each point \( x \in (a, b) \), and

\[
g'(x) = f(x).
\]

Proof. Parallel to (2.29), we have, for \( h > 0 \),

\[
\frac{1}{h} [g(x + h) - g(x)] = \frac{1}{h} \int_x^{x+h} f(t) \, dt.
\]

If \( f \) is continuous at \( x \), then, for any \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that \( |f(t) - f(x)| \leq \varepsilon \) whenever \( |t - x| \leq \delta \). Thus the right side of (0.33) is within \( \varepsilon \) of \( f(x) \) whenever \( h \in (0, \delta] \). Thus the desired limit exists as \( h \downarrow 0 \). A similar argument treats \( h \uparrow 0 \).

The next result is the rest of the Fundamental Theorem of Calculus.

Theorem 0.7. If \( G \) is differentiable and \( G'(x) \) is continuous on \([a, b] \), then

\[
\int_a^b G'(t) \, dt = G(b) - G(a).
\]

Proof. Consider the function

\[
g(x) = \int_a^x G'(t) \, dt.
\]

We have \( g \in C([a, b]) \), \( g(a) = 0 \), and, by Theorem 0.6,

\[
g'(x) = G'(x), \quad \forall \ x \in (a, b).
\]

Thus \( f(x) = g(x) - G(x) \) is continuous on \([a, b] \), and

\[
f'(x) = 0, \quad \forall \ x \in (a, b).
\]

We claim that (0.36) implies \( f \) is constant on \([a, b] \). Granted this, since \( f(a) = g(a) - G(a) = -G(a) \), we have \( f(x) = -G(a) \) for all \( x \in [a, b] \), so the integral (2.35) is equal to \( G(x) - G(a) \) for all \( x \in [a, b] \). Taking \( x = b \) yields (2.34).

The fact that (0.36) implies \( f \) is constant on \([a, b] \) is a consequence of the following result, known as the Mean Value Theorem.

Theorem 0.8. Let \( f : [a, \beta] \rightarrow \mathbb{R} \) be continuous, and assume \( f \) is differentiable on \((a, \beta) \). Then \( \exists \ \xi \in (a, \beta) \) such that

\[
f' (\xi) = \frac{f(\beta) - f(a)}{\beta - a}.
\]
Proof. Set \( g(x) = f(x) - \kappa(x - a) \), where \( \kappa \) is the right side of (0.37). Note that \( g'(\xi) = f'(\xi) - \kappa \), so it suffices to show that \( g'(\xi) = 0 \) for some \( \xi \in (a, \beta) \). Note also that \( g(a) = g(\beta) \). Since \([a, \beta]\) is compact, \( g \) must assume a maximum and a minimum on \([a, \beta]\). Since \( g(a) = g(\beta) \), one of these must be assumed at an interior point, at which \( g' \) vanishes.

Now, to see that (0.36) implies \( f \) is constant on \([a, b]\); if not, \( \exists \beta \in (a, b) \) such that \( f(\beta) \neq f(a) \). Then just apply Theorem 0.8 to \( f \) on \([a, \beta]\). This completes the proof of Theorem 0.7.

We now extend Theorems 0.6–0.7 to the setting of Riemann integrable functions.

**Proposition 0.9.** Let \( f \in \mathcal{R}([a, b]) \), and define \( g \) by (0.28). If \( x \in [a, b] \) and \( f \) is continuous at \( x \), then \( g \) is differentiable at \( x \), and \( g'(x) = f(x) \).

The proof is identical to that of Theorem 0.6.

**Proposition 0.10.** Assume \( G \) is differentiable on \([a, b]\) and \( G' \in \mathcal{R}([a, b]) \). Then (0.34) holds.

**Proof.** We have

\[
G(b) - G(a) = \sum_{k=0}^{n-1} \left[ G \left( a + (b - a) \frac{k+1}{n} \right) - G \left( a + (b - a) \frac{k}{n} \right) \right] \\
= \frac{b - a}{n} \sum_{k=0}^{n-1} G'(\xi_k),
\]

for some \( \xi_k \) satisfying

\[
a + (b - a) \frac{k}{n} < \xi_k < a + (b - a) \frac{k+1}{n},
\]

as a consequence of the Mean Value Theorem. Given \( G' \in \mathcal{R}([a, b]) \), Darboux’s theorem (Theorem 0.4) implies that as \( n \to \infty \) one gets \( G(b) - G(a) = \int_a^b G'(t) \, dt \).

Note that the beautiful symmetry in Theorems 0.6–0.7 is not preserved in Propositions 0.9–0.10. The hypothesis of Proposition 0.10 requires \( G \) to be differentiable at each \( x \in [a, b] \), but the conclusion of Proposition 0.9 does not yield differentiability at all points. For this reason, we regard Propositions 0.9–0.10 as less “fundamental” than Theorems 0.6–0.7. There are more satisfactory extensions of the fundamental theorem of calculus, involving the Lebesgue integral, and a more subtle notion of the “derivative” of a non-smooth function. For this, we can point the reader to Chapters 10-11 of the text [T2], Measure Theory and Integration.

So far, we have dealt with integration of real valued functions. If \( f : I \to \mathbb{C} \), we set \( f = f_1 + if_2 \) with \( f_j : I \to \mathbb{R} \) and say \( f \in \mathcal{R}(I) \) if and only if \( f_1 \) and \( f_2 \) are in \( \mathcal{R}(I) \). Then

\[
\int_I f \, dx = \int_I f_1 \, dx + i \int_I f_2 \, dx.
\]
There are straightforward extensions of Propositions 0.5–0.10 to complex valued functions. Similar comments apply to functions \( f : I \to \mathbb{R}^n \).

If a function \( G \) is differentiable on \((a, b)\), and \( G' \) is continuous on \((a, b)\), we say \( G \) is a \( C^1 \) function, and write \( G \in C^1((a, b)) \). Inductively, we say \( G \in C^k((a, b)) \) provided \( G' \in C^{k-1}((a, b)) \).

An easy consequence of the definition (0.31) of the derivative is that, for any real constants \( a, b, c \),

\[
    f(x) = ax^2 + bx + c \implies \frac{d}{dx} f = 2ax + b.
\]

Now, it is a simple enough step to replace \( a, b, c \) by \( y, z, w \), in these formulas. Having done that, we can regard \( y, z, \) and \( w \) as variables, along with \( x \):

\[
    F(x, y, z, w) = yx^2 + zx + w.
\]

We can then hold \( y, z \) and \( w \) fixed (e.g., set \( y = a, z = b, w = c \)), and then differentiate with respect to \( x \); we get

\[
    \frac{\partial F}{\partial x} = 2yx + z,
\]

the partial derivative of \( F \) with respect to \( x \). Generally, if \( F \) is a function of \( n \) variables, \( x_1, \ldots, x_n \), we set

\[
    \frac{\partial F}{\partial x_j}(x_1, \ldots, x_n)
    = \lim_{h \to 0} \frac{1}{h} \left[ F(x_1, \ldots, x_{j-1}, x_j + h, x_{j+1}, \ldots, x_n) - F(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n) \right],
\]

where the limit exists. Section one carries on with a further investigation of the derivative of a function of several variables.

**Complementary results on Riemann integrability**

Here we provide a condition, more general then Proposition 0.2, which guarantees Riemann integrability.

**Proposition 0.11.** Let \( f : I \to \mathbb{R} \) be a bounded function, with \( I = [a, b] \). Suppose that the set \( S \) of points of discontinuity of \( f \) has the property

\[
    \text{cont}^+(S) = 0.
\]

Then \( f \in \mathcal{R}(I) \).

**Proof.** Say \(|f(x)| \leq M \). Take \( \varepsilon > 0 \). As in (0.21), take intervals \( J_1, \ldots, J_N \) such that \( S \subset \bigcup J_1 \cup \cdots \cup J_N \) and \( \sum_{k=1}^N \ell(J_k) < \varepsilon \). In fact, fatten each \( J_k \) such that \( S \) is contained
in the interior of this collection of intervals. Consider a partition $P_0$ of $I$, whose intervals
include $J_1, \ldots, J_N$, amongst others, which we label $I_1, \ldots, I_K$. Now $f$ is continuous on
each interval $I_\nu$, so, subdividing each $I_\nu$ as necessary, hence refining $P_0$ to a partition
$P_1$, we arrange that $\sup f - \inf f < \varepsilon$ on each such subdivided interval. Denote these
subdivided intervals $I'_1, \ldots, I'_L$. It readily follows that

$$0 \leq I_{P_1}(f) - I_{P_0}(f) < \sum_{k=1}^N 2M\ell(J_k) + \sum_{k=1}^L \varepsilon\ell(I'_k)$$

$$< 2\varepsilon M + \varepsilon\ell(I).$$

Since $\varepsilon$ can be taken arbitrarily small, this establishes that $f \in \mathcal{R}(I)$.

REMARK. An even better result is that such $f$ is Riemann integrable if and only if

$$(0.38A) \quad m^*(S) = 0,$$

where $m^*(S)$ is defined by (0.22). The implication $m^*(S) = 0 \Rightarrow f \in \mathcal{R}(I)$ is established in
Chapter 4 of [T4], and its higher dimensional generalization is established in Proposition
4.31 of this text. For the reverse implication $f \in \mathcal{R}(I) \Rightarrow m^*(S) = 0$, one can see standard
books on measure theory, such as [Fol] and [T2].

We give an example of a function to which Proposition 0.11 applies, and then an example
for which Proposition 0.11 fails to apply, though the function is Riemann integrable.

EXAMPLE 1. Let $I = [0,1]$. Define $f : I \to \mathbb{R}$ by

$$f(0) = 0,$$

$$f(x) = (-1)^j \text{ for } x \in (2^{-(j+1)}, 2^{-j}], \ j \geq 0.$$ 

Then $|f| \leq 1$ and the set of points of discontinuity of $f$ is

$$S = \{0\} \cup \{2^{-j} : j \geq 1\}.$$ 

It is easy to see that $\text{cont}^+ S = 0$. Hence $f \in \mathcal{R}(I)$.

See Exercises 19–20 below for a more elaborate example to which Proposition 0.11 applies.

EXAMPLE 2. Again $I = [0,1]$. Define $f : I \to \mathbb{R}$ by

$$f(x) = 0 \text{ if } x \notin \mathbb{Q},$$

$$\frac{1}{n} \text{ if } x = \frac{m}{n}, \text{ in lowest terms.}$$
Then $|f| \leq 1$ and the set of points of discontinuity of $f$ is

$$S = I \cap \mathbb{Q}.$$ 

As we have seen below (0.23), $\text{cont}^+ S = 1$, so Proposition 0.11 does not apply. Nevertheless, it is fairly easy to see directly that

$$\overline{I}(f) = \underline{I}(f) = 0, \quad \text{so } f \in \mathcal{R}(I).$$

In fact, given $\varepsilon > 0$, $f \geq \varepsilon$ only on a finite set, hence

$$\overline{I}(f) \leq \varepsilon, \quad \forall \varepsilon > 0.$$

As indicated below (0.23), (0.38A) does apply to this function.

By contrast, the function $\vartheta$ in (0.16) is discontinuous at each point of $I$.

We mention an alternative characterization of $\overline{I}(f)$ and $\underline{I}(f)$, which can be useful. Given $I = [a, b]$, we say $g : I \to \mathbb{R}$ is \textit{piecewise constant} on $I$ (and write $g \in \text{PK}(I)$) provided there exists a partition $\mathcal{P} = \{J_k\}$ of $I$ such that $g$ is constant on the interior of each interval $J_k$. Clearly $\text{PK}(I) \subset \mathcal{R}(I)$. It is easy to see that, if $f : I \to \mathbb{R}$ is bounded,

$$\overline{I}(f) = \inf \left\{ \int_I f_1 \, dx : f_1 \in \text{PK}(I), \ f_1 \geq f \right\},$$

(0.39)

$$\underline{I}(f) = \sup \left\{ \int_I f_0 \, dx : f_0 \in \text{PK}(I), \ f_0 \leq f \right\}.$$ 

Hence, given $f : I \to \mathbb{R}$ bounded,

$$f \in \mathcal{R}(I) \iff \text{for each } \varepsilon > 0, \ \exists f_0, f_1 \in \text{PK}(I) \text{ such that}$$

(0.40)

$$f_0 \leq f \leq f_1 \quad \text{and } \int_I (f_1 - f_0) \, dx < \varepsilon.$$ 

This can be used to prove

(0.41) \quad f, g \in \mathcal{R}(I) \implies fg \in \mathcal{R}(I),

via the fact that

(0.42) \quad f_j, g_j \in \text{PK}(I) \implies f_j g_j \in \text{PK}(I).

In fact, we have the following, which can be used to prove (0.41).
Proposition 0.12. Let $f \in \mathcal{R}(I)$, and assume $|f| \leq M$. Let $\varphi : [-M, M] \to \mathbb{R}$ be continuous. Then $\varphi \circ f \in \mathcal{R}(I)$.

Proof. We proceed in steps.

**Step 1.** We can obtain $\varphi$ as a uniform limit on $[-M, M]$ of a sequence $\varphi_\nu$ of continuous, piecewise linear functions. Then $\varphi_\nu \circ f \to \varphi \circ f$ uniformly on $I$. A uniform limit $g$ of functions $g_\nu \in \mathcal{R}(I)$ is in $\mathcal{R}(I)$ (see Exercise 12). So it suffices to prove Proposition 0.12 when $\varphi$ is continuous and piecewise linear.

**Step 2.** Given $\varphi : [-M, M] \to \mathbb{R}$ continuous and piecewise linear, it is an exercise to write $\varphi = \varphi_1 - \varphi_2$, with $\varphi_j : [-M, M] \to \mathbb{R}$ monotone, continuous, and piecewise linear. Now $\varphi_1 \circ f, \varphi_2 \circ f \in \mathcal{R}(I)$ imply $\varphi \circ f \in \mathcal{R}(I)$.

**Step 3.** We now demonstrate Proposition 0.12 when $\varphi : [-M, M] \to \mathbb{R}$ is monotone and Lipschitz. By Step 2, this will suffice. So we assume

$$-M \leq x_1 < x_2 \leq M \implies \varphi(x_1) \leq \varphi(x_2) \quad \text{and} \quad \varphi(x_2) - \varphi(x_1) \leq L(x_2 - x_1),$$

for some $L < \infty$. Given $\varepsilon > 0$, pick $f_0, f_1 \in \text{PK}(I)$, as in (0.40). Then

$$\varphi \circ f_0, \varphi \circ f_1 \in \text{PK}(I), \quad \varphi \circ f_0 \leq \varphi \circ f \leq \varphi \circ f_1,$$

and

$$\int_I (\varphi \circ f_1 - \varphi \circ f_0) \, dx \leq L \int_I (f_1 - f_0) \, dx \leq L \varepsilon.$$

This proves $\varphi \circ f \in \mathcal{R}(I)$.

**Exercises**

1. Let $c > 0$ and let $f : [ac, bc] \to \mathbb{R}$ be Riemann integrable. Working directly with the definition of integral, show that

$$(0.43) \quad \int_a^b f(cx) \, dx = \frac{1}{c} \int_{ac}^{bc} f(x) \, dx.$$

More generally, show that

$$(0.44) \quad \int_{a-d/c}^{b-d/c} f(cx + d) \, dx = \frac{1}{c} \int_{ac}^{bc} f(x) \, dx.$$

2. Let $f : I \times S \to \mathbb{R}$ be continuous, where $I = [a, b]$ and $S \subset \mathbb{R}^n$. Take $\varphi(y) = \int_I f(x, y) \, dx$. 

Show that \( \varphi \) is continuous on \( S \).

**Hint.** If \( f_j : I \to \mathbb{R} \) are continuous and \( |f_1(x) - f_2(x)| \leq \delta \) on \( I \), then

\[
\left| \int_I f_1 \, dx - \int_I f_2 \, dx \right| \leq \ell(I) \delta.
\]

**Hint.** Suppose \( y_j \in S, y_j \to y \in S \). Let \( \tilde{S} = \{y_j\} \cup \{y\} \). This is compact. Thus

\[
\varphi : I \times \tilde{S} \to \mathbb{R}
\]

is uniformly continuous. Hence

\[
|\varphi(x, y_j) - \varphi(x, y)| \leq \omega(|y_j - y|), \quad \forall x \in I,
\]

where \( \omega(\delta) \to 0 \) as \( \delta \to 0 \).

3. With \( f \) as in Exercise 2, suppose \( g_j : S \to \mathbb{R} \) are continuous and \( a \leq g_0(y) < g_1(y) \leq b \). Take \( \varphi(y) = \int_{g_0(y)}^{g_1(y)} f(x, y) \, dx \). Show that \( \varphi \) is continuous on \( S \).

**Hint.** Make a change of variables, linear in \( x \), to reduce this to Exercise 2.

4. Suppose \( f : (a, b) \to (c, d) \) and \( g : (c, d) \to \mathbb{R} \) are differentiable. Show that \( h(x) = g(f(x)) \) is differentiable and

\[
h'(x) = g'(f(x))f'(x).
\]

This is the **chain rule**.

**Hint.** Peek at the proof of the chain rule in §1.

5. If \( f_1 \) and \( f_2 \) are differentiable on \( (a, b) \), show that \( f_1(x)f_2(x) \) is differentiable and

\[
\frac{d}{dx}(f_1(x)f_2(x)) = f'_1(x)f_2(x) + f_1(x)f'_2(x).
\]

If \( f_2(x) \neq 0, \forall x \in (a, b) \), show that \( f_1(x)/f_2(x) \) is differentiable, and

\[
\frac{d}{dx}\left(\frac{f_1(x)}{f_2(x)}\right) = \frac{f_1'(x)f_2(x) - f_1(x)f_2'(x)}{f_2(x)^2}.
\]

6. Let \( \varphi : [a, b] \to [A, B] \) be \( C^1 \) on a neighborhood \( J \) of \( [a, b] \), with \( \varphi'(x) > 0 \) for all \( x \in [a, b] \). Assume \( \varphi(a) = A, \varphi(b) = B \). Show that the identity

\[
\int_A^B f(y) \, dy = \int_a^b f(\varphi(t))\varphi'(t) \, dt,
\]

for any \( f \in C(I), I = [A, B] \), follows from the chain rule and the Fundamental Theorem of Calculus.

**Hint.** Replace \( b \) by \( x \), \( B \) by \( \varphi(x) \), and differentiate.
7. Show that (0.46) holds for each $f \in \mathcal{P}(I)$. Using (0.39)–(0.40), show that $f \in \mathcal{R}(I) \Rightarrow f \circ \varphi \in \mathcal{R}([a,b])$ and (0.46) holds. (This result contains that of Exercise 1.)

8. Show that, if $f$ and $g$ are $C^1$ on a neighborhood of $[a,b]$, then

$$
(0.47) \quad \int_a^b f(s)g'(s) \, ds = - \int_a^b f'(s)g(s) \, ds + \left[ f(b)g(b) - f(a)g(a) \right].
$$

This transformation of integrals is called "integration by parts."

9. Let $f : (-a,a) \to \mathbb{R}$ be a $C^{j+1}$ function. Show that, for $x \in (-a,a)$,

$$
(0.48) \quad f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \cdots + \frac{f^{(j)}(0)}{j!}x^j + R_j(x),
$$

where

$$
(0.49) \quad R_j(x) = \int_0^x \frac{(x-s)^j}{j!} f^{(j+1)}(s) \, ds
$$

This is Taylor’s formula with remainder.

**Hint.** Use induction. If (0.48)–(0.49) holds for $0 \leq j \leq k$, show that it holds for $j = k + 1$, by showing that

$$
(0.50) \quad \int_0^x \frac{(x-s)^k}{k!} f^{(k+1)}(s) \, ds = \frac{f^{(k+1)}(0)}{(k+1)!} x^{k+1} + \int_0^x \frac{(x-s)^{k+1}}{(k+1)!} f^{(k+2)}(s) \, ds.
$$

To establish this, use the integration by parts formula (0.47), with $f(s)$ replaced by $f^{(k+1)}(s)$, and with appropriate $g(s)$. See §1 for another approach. Note that another presentation of (0.49) is

$$
(0.51) \quad R_j(x) = \frac{x^{j+1}}{(j+1)!} \int_0^1 f^{(j+1)} \left( (1-t^{j/(j+1)})x \right) \, dt.
$$

10. Assume $f : (-a,a) \to \mathbb{R}$ is a $C^j$ function. Show that, for $x \in (-a,a)$, (0.48) holds, with

$$
(0.52) \quad R_j(x) = \frac{1}{(j-1)!} \int_0^x (x-s)^{j-1} \left[ f^{(j)}(s) - f^{(j)}(0) \right] \, ds.
$$

**Hint.** Apply (0.49) with $j$ replaced by $j - 1$. Add and subtract $f^{(j)}(0)$ to the factor $f^{(j)}(s)$ in the resulting integrand.
11. Given \( I = [a, b] \), show that

\[
(0.53) \quad f, g \in \mathcal{R}(I) \implies fg \in \mathcal{R}(I),
\]

as advertised in (0.41).

12. Assume \( f_k \in \mathcal{R}(I) \) and \( f_k \to f \) uniformly on \( I \). Prove that \( f \in \mathcal{R}(I) \) and

\[
(0.54) \quad \int_I f_k \, dx \to \int_I f \, dx.
\]

13. Given \( I = [a, b] \), \( I_\varepsilon = [a + \varepsilon, b - \varepsilon] \), assume \( f_k \in \mathcal{R}(I) \), \( |f_k| \leq M \) on \( I \) for all \( k \), and

\[
(0.55) \quad f_k \to f \quad \text{uniformly on} \quad I_\varepsilon,
\]

for all \( \varepsilon \in (0, (b - a)/2) \). Prove that \( f \in \mathcal{R}(I) \) and (0.54) holds.

14. Use the fundamental theorem of calculus to compute

\[
(0.56) \quad \int_a^b x^r \, dx, \quad r \in \mathbb{Q} \setminus \{-1\},
\]

where \( 0 \leq a < b < \infty \) if \( r \geq 0 \) and \( 0 < a < b < \infty \) if \( r < 0 \).

15. Use the change of variable result of Exercise 6 to compute

\[
\int_0^1 x\sqrt{1 + x^2} \, dx.
\]

16. We say \( f \in \mathcal{R}(\mathbb{R}) \) provided \( f|_{[k, k+1]} \in \mathcal{R}([k, k+1]) \) for each \( k \in \mathbb{Z} \), and

\[
(0.57) \quad \sum_{k=-\infty}^{\infty} \int_k^{k+1} |f(x)| \, dx < \infty.
\]

If \( f \in \mathcal{R}(\mathbb{R}) \), we set

\[
(0.58) \quad \int_{-\infty}^{\infty} f(x) \, dx = \lim_{k \to \infty} \int_{-k}^{k} f(x) \, dx.
\]

Formulate and demonstrate basic properties of the integral over \( \mathbb{R} \) of elements of \( \mathcal{R}(\mathbb{R}) \).

17. This exercise discusses the integral test for absolute convergence of an infinite series,
which goes as follows. Let \( f \) be a positive, monotonically decreasing, continuous function on \([0, \infty)\), and suppose \( |a_k| = f(k) \). Then

\[
\sum_{k=0}^{\infty} |a_k| < \infty \iff \int_0^{\infty} f(x) \, dx < \infty.
\]

Prove this.

**Hint.** Use

\[
\sum_{k=1}^{N} |a_k| \leq \int_0^{N} f(x) \, dx \leq \sum_{k=0}^{N-1} |a_k|.
\]

18. Use the integral test to show that, if \( p > 0 \),

\[
\sum_{k=1}^{\infty} \frac{1}{k^p} < \infty \iff p > 1.
\]

**Hint.** Use Exercise 14 to evaluate \( I_N(p) = \int_1^{N} x^{-p} \, dx \), for \( p \neq -1 \), and let \( N \to \infty \). See if you can show \( \int_1^{\infty} x^{-1} \, dx = \infty \) without knowing about \( \log N \). **Subhint.** Show that \( \int_1^{2} x^{-1} \, dx = \int_{N}^{2N} x^{-1} \, dx \).

In Exercises 19–20, \( C \subset \mathbb{R} \) is the Cantor set, defined as follows. Take a closed, bounded interval \([a, b] = C_0\). Let \( C_1 \) be obtained from \( C_0 \) by deleting the open middle third interval, of length \( (b - a)/3 \). At the \( j \)th stage, \( C_j \) is a disjoint union of \( 2^j \) closed intervals, each of length \( 3^{-j}(b - a) \). Then \( C_{j+1} \) is obtained from \( C_j \) by deleting the open middle third of each of these \( 2^j \) intervals. We have \( C_0 \supset C_1 \supset \cdots \supset C_j \supset \cdots \), each a closed subset of \([a, b]\). The Cantor set is \( C = \cap_{j \geq 0} C_j \).

19. Show that \( \text{cont}^+ C_j = (2/3)^j(b - a) \), and conclude that

\[
\text{cont}^+ C = 0.
\]

20. Define \( f : [a, b] \to \mathbb{R} \) as follows. We call an interval of length \( 3^{-j}(b - a) \), omitted in passing from \( C_{j-1} \) to \( C_j \), a “\( j \)-interval.” Set

\[
f(x) = 0, \quad \text{if } x \in C,
\]

\[
(-1)^j, \quad \text{if } x \text{ belongs to a } j\text{-interval}.
\]

Show that the set of discontinuities of \( f \) is \( C \). Hence Proposition 0.11 implies \( f \in \mathcal{R}([a, b]) \).

21. Generalize Exercise 8 as follows. Assume \( f \) and \( g \) are differentiable on a neighborhood of \([a, b]\) and \( f', g' \in \mathcal{R}([a, b]) \). Then show that (0.47) holds.

**Hint.** Use the results of Exercise 11 to show that \((fg)' \in \mathcal{R}([a, b])\).
22. Let $f : I \to \mathbb{R}$ be bounded, $I = [a, b]$. Show that

$$
\overline{I}(f) = \inf \left\{ \int_I f_1 \, dx : f_1 \in C(I), \ f_1 \geq f \right\},
$$

$$
\underline{I}(f) = \sup \left\{ \int_I f_0 \, dx : f_0 \in C(I), \ f_0 \leq f \right\}.
$$

Compare (0.39). Then show that

$$
f \in \mathcal{R}(I) \iff \text{for each } \varepsilon > 0, \ \exists f_0, f_1 \in C(I) \text{ such that}\n
(0.61) \quad f_0 \leq f \leq f_1 \text{ and } \int_I (f_1 - f_0) \, dx < \varepsilon.
$$

Compare (0.40).
1. The derivative

Let $O$ be an open subset of $\mathbb{R}^n$, and $F: O \to \mathbb{R}^m$ a continuous function. We say $F$ is differentiable at a point $x \in O$, with derivative $L$, if $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation such that, for $y \in \mathbb{R}^n$, small,

\begin{equation}
F(x + y) = F(x) + Ly + R(x, y)
\end{equation}

with

\begin{equation}
\frac{\|R(x, y)\|}{\|y\|} \to 0 \quad \text{as} \quad y \to 0.
\end{equation}

We denote the derivative at $x$ by $DF(x) = L$. With respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, $DF(x)$ is simply the matrix of partial derivatives,

\begin{equation}
DF(x) = \left( \frac{\partial F_j}{\partial x_k} \right) = \begin{pmatrix} \partial F_1/\partial x_1 & \cdots & \partial F_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial F_m/\partial x_1 & \cdots & \partial F_m/\partial x_n \end{pmatrix},
\end{equation}

so that, if $v = (v_1, \ldots, v_n)^t$, (regarded as a column vector) then

\begin{equation}
DF(x)v = \begin{pmatrix} \sum_k (\partial F_1/\partial x_k)v_k \\ \vdots \\ \sum_k (\partial F_m/\partial x_k)v_k \end{pmatrix}.
\end{equation}

Recall the definition of the partial derivative $\partial f_j/\partial x_k$ from §0. It will be shown below that $F$ is differentiable whenever all the partial derivatives exist and are continuous on $O$. In such a case we say $F$ is a $C^1$ function on $O$. More generally, $F$ is said to be $C^k$ if all its partial derivatives of order $\leq k$ exist and are continuous. If $F$ is $C^k$ for all $k$, we say $F$ is $C^\infty$.

In (1.2), we can use the Euclidean norm on $\mathbb{R}^n$ and $\mathbb{R}^m$. This norm is defined by

\begin{equation}
\|x\| = (x_1^2 + \cdots + x_n^2)^{1/2}
\end{equation}

for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Any other norm would do equally well.

An application of the Fundamental Theorem of Calculus, to functions of each variable $x_j$ separately, yields the following. If we assume $F: O \to \mathbb{R}^m$ is differentiable in each variable separately, and that each $\partial F/\partial x_j$ is continuous on $O$, then

\begin{equation}
F(x + y) = F(x) + \sum_{j=1}^n [F(x + z_j) - F(x + z_{j-1})] = F(x) + \sum_{j=1}^n A_j(x, y)y_j,
\end{equation}

\begin{equation}
A_j(x, y) = \int_0^1 \frac{\partial F}{\partial x_j}(x + z_{j-1} + ty_j e_j) \, dt,
\end{equation}
where $z_0 = 0$, $z_j = (y_1, \ldots, y_j, 0, \ldots, 0)$, and $\{e_j\}$ is the standard basis of $\mathbb{R}^n$. Consequently,

$$F(x + y) = F(x) + \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(x) y_j + R(x, y),$$

(1.7)

$$R(x, y) = \sum_{j=1}^{n} y_j \int_{0}^{1} \left\{ \frac{\partial F}{\partial x_j}(x + z_{j-1} + ty_j e_j) - \frac{\partial F}{\partial x_j}(x) \right\} dt.$$

Now (1.7) implies $F$ is differentiable on $O$, as we stated below (1.4). Thus we have established the following.

**Proposition 1.1.** If $O$ is an open subset of $\mathbb{R}^n$ and $F : O \to \mathbb{R}^m$ is of class $C^1$, then $F$ is differentiable at each point $x \in O$.

As is shown in many calculus texts, one can use the Mean Value Theorem instead of the Fundamental Theorem of Calculus, and obtain a slightly sharper result.

Let us give some examples of derivatives. First, take $n = 2$, $m = 1$, and set

$$F(x) = (\sin x_1)(\sin x_2).$$

Then

$$DF(x) = ((\cos x_1)(\sin x_2), (\sin x_1)(\cos x_2)).$$

Next, take $n = m = 2$ and

$$F(x) = \begin{pmatrix} x_1 x_2 \\ x_1^2 - x_2^2 \end{pmatrix}.$$

Then

$$DF(x) = \begin{pmatrix} x_2 & x_1 \\ 2x_1 & -2x_2 \end{pmatrix}.$$

We can replace $\mathbb{R}^n$ and $\mathbb{R}^m$ by more general finite-dimensional real vector spaces, isomorphic to Euclidean space. For example, the space $M(n, \mathbb{R})$ of real $n \times n$ matrices is isomorphic to $\mathbb{R}^{n^2}$. Consider the function

$$S : M(n, \mathbb{R}) \to M(n, \mathbb{R}), \quad S(X) = X^2.$$

We have

$$(X + Y)^2 = X^2 + XY + YX + Y^2$$

$$= X^2 + DS(X)Y + R(X, Y),$$

(1.13)
with $R(X, Y) = Y^2$, and hence
\begin{equation}
(1.14) \quad DS(X)Y = XY + YX.
\end{equation}

For our next example, we take
\begin{equation}
(1.15) \quad \mathcal{O} = G\ell(n, \mathbb{R}) = \{X \in M(n, \mathbb{R}) : \text{det } X \neq 0\},
\end{equation}
which, as shown below, is open in $M(n, \mathbb{R})$. We consider
\begin{equation}
(1.16) \quad \Phi : G\ell(n, \mathbb{R}) \longrightarrow M(n, \mathbb{R}), \quad \Phi(X) = X^{-1},
\end{equation}
and compute $D\Phi(I)$. We use the following. If, for $A \in M(n, \mathbb{R})$,
\begin{equation}
(1.17) \quad \|A\| = \sup\{\|Av\| : v \in \mathbb{R}^n, \|v\| \leq 1\},
\end{equation}
then
\begin{equation}
(1.18) \quad A, B \in M(n, \mathbb{R}) \Rightarrow \|A + B\| \leq \|A\| + \|B\| \quad \text{and} \quad \|AB\| \leq \|A\| \cdot \|B\|,
\end{equation}
hence $Y \in M(n, \mathbb{R}) \Rightarrow \|Y^k\| \leq \|Y\|^k$.

Also
\begin{equation}
(1.19) \quad S_k = I - Y + Y^2 - \cdots + (-1)^kY^k
\end{equation}
\begin{equation}
\Rightarrow YS_k = S_kY = Y - Y^2 + Y^3 - \cdots + (-1)^kY^{k+1}
\end{equation}
\begin{equation}
\Rightarrow (I + Y)S_k = S_k(I + Y) = I + (-1)^kY^{k+1},
\end{equation}
hence
\begin{equation}
(1.20) \quad \|Y\| < 1 \implies (I + Y)^{-1} = \sum_{k=0}^{\infty} (-1)^kY^k = I - Y + Y^2 - \cdots,
\end{equation}
so
\begin{equation}
(1.21) \quad D\Phi(I)Y = -Y.
\end{equation}

Related calculations show that $G\ell(n, \mathbb{R})$ is open in $M(n, \mathbb{R})$. In fact, given $X \in G\ell(n, \mathbb{R})$, $Y \in M(n, \mathbb{R})$,
\begin{equation}
(1.22) \quad X + Y = X(I + X^{-1}Y),
\end{equation}
which by (1.20) is invertible as long as
\begin{equation}
(1.23) \quad \|X^{-1}Y\| < 1.
\end{equation}
One can proceed from here to compute $D\Phi(X)$. See the exercises.

We return to general considerations, and derive the chain rule for the derivative. Let $F : \mathcal{O} \to \mathbb{R}^m$ be differentiable at $x \in \mathcal{O}$, as above, let $U$ be a neighborhood of $z = F(x)$ in $\mathbb{R}^m$, and let $G : U \to \mathbb{R}^k$ be differentiable at $z$. Consider $H = G \circ F$. We have

$$H(x+y) = G(F(x+y))$$

(1.24)

$$= G(F(x)+DF(x)y+R(x,y))$$

$$= G(z) + DG(z)(DF(x)y+R(x,y)) + R_1(x,y)$$

$$= G(z) + DG(z)DF(x)y+R_2(x,y)$$

with

$$\|R_2(x,y)\| \|y\| \to 0 \text{ as } y \to 0.$$
for some $\alpha_j \in (0, 1)$, depending on $x$ and $h$. Iterating this, if $\partial_j (\partial_k f)$ exists on $\mathcal{O}$, then, for $x \in \mathcal{O}$, $h > 0$ sufficiently small,

$$
\Delta_{k, h} \Delta_{j, h} f(x) = \partial_k (\Delta_{j, h} f)(x + \alpha_k h e_k)
= \Delta_{j, h} (\partial_k f)(x + \alpha_k h e_k)
= \partial_j \partial_k f(x + \alpha_k h e_k + \alpha_j h e_j),
$$

with $\alpha_j, \alpha_k \in (0, 1)$. Here we have used the elementary result

$$
\partial_k \Delta_{j, h} f = \Delta_{j, h} (\partial_k f).
$$

We deduce the following.

**Proposition 1.3.** If $\partial_k f$ and $\partial_j \partial_k f$ exist on $\mathcal{O}$ and $\partial_j \partial_k f$ is continuous at $x_0 \in \mathcal{O}$, then,

$$
\partial_j \partial_k f(x_0) = \lim_{h \to 0} \Delta_{k, h} \Delta_{j, h} f(x_0).
$$

Clearly

$$
\Delta_{k, h} \Delta_{j, h} f = \Delta_{j, h} \Delta_{k, h} f,
$$

so we have the following, which easily implies Proposition 1.2.

**Corollary 1.4.** In the setting of Proposition 1.3, if also $\partial_j f$ and $\partial_k \partial_j f$ exist on $\mathcal{O}$ and $\partial_k \partial_j f$ is continuous at $x_0$, then

$$
\partial_j \partial_k f(x_0) = \partial_k \partial_j f(x_0).
$$

We now describe two convenient notations to express higher order derivatives of a $C^k$ function $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^n$ is open. In one, let $J$ be a $k$-tuple of integers between 1 and $n$; $J = (j_1, \ldots, j_k)$. We set

$$
f^{(J)}(x) = \partial_{j_k} \cdots \partial_{j_1} f(x), \quad \partial_j = \frac{\partial}{\partial x_j}.
$$

We set $|J| = k$, the total order of differentiation. As we have seen in Proposition 1.2, $\partial_k \partial_j f = \partial_j \partial_k f$ provided $f \in C^2(\Omega)$. It follows that, if $f \in C^k(\Omega)$, then $\partial_{j_k} \cdots \partial_{j_1} f = \partial_{j_k} \cdots \partial_{j_1} f$ whenever $\{\ell_1, \ldots, \ell_k\}$ is a permutation of $\{j_1, \ldots, j_k\}$. Thus, another convenient notation to use is the following. Let $\alpha$ be an $n$-tuple of non-negative integers, $\alpha = (\alpha_1, \ldots, \alpha_n)$. Then we set

$$
f^{(\alpha)}(x) = \partial_{1}^{\alpha_1} \cdots \partial_{n}^{\alpha_n} f(x), \quad |\alpha| = \alpha_1 + \cdots + \alpha_n.
$$

Note that, if $|J| = |\alpha| = k$ and $f \in C^k(\Omega)$,

$$
f^{(J)}(x) = f^{(\alpha)}(x), \text{ with } \alpha_i = \# \{\ell : j_\ell = i\}.
$$
Correspondingly, there are two expressions for monomials in \( x = (x_1, \ldots, x_n) \):

\[
x^J = x_{j_1} \cdots x_{j_k}, \quad x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n},
\]

and \( x^J = x^\alpha \) provided \( J \) and \( \alpha \) are related as in (1.37). Both these notations are called “multi-index” notations.

We now derive Taylor's formula with remainder for a smooth function \( F : \Omega \to \mathbb{R} \), making use of these multi-index notations. We will apply the one variable formula (0.48)--(0.49), i.e.,

\[
\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{1}{2}\varphi''(0)t^2 + \cdots + \frac{1}{k!}\varphi^{(k)}(0)t^k + r_k(t),
\]

with

\[
r_k(t) = \frac{1}{k!}\int_0^t (t-s)^k \varphi^{(k+1)}(s) \, ds,
\]

given \( \varphi \in C^{k+1}(I) \), \( I = (-a, a) \). (See Exercise 9 of §0, Exercise 7 of this section, and also Appendix D for further discussion.) Let us assume \( 0 \in \Omega \), and that the line segment from 0 to \( x \) is contained in \( \Omega \). We set \( \varphi(t) = F(tx) \), and apply (1.39)–(1.40) with \( t = 1 \). Applying the chain rule, we have

\[
\varphi'(t) = \sum_{j=1}^n \partial_j F(tx)x_j = \sum_{|J|=1} F^{(J)}(tx)x^J.
\]

Differentiating again, we have

\[
\varphi''(t) = \sum_{|J|=2} F^{(J+K)}(tx)x^Jx^K = \sum_{|J|=2} F^{(J)}(tx)x^J,
\]

where, if \(|J| = k, |K| = \ell\), we take \( J + K = (j_1, \ldots, j_k, k_1, \ldots, k_\ell) \). Inductively, we have

\[
\varphi^{(k)}(t) = \sum_{|J|=k} F^{(J)}(tx)x^J.
\]

Hence, from (1.39) with \( t = 1 \),

\[
F(x) = F(0) + \sum_{|J|=1} F^{(J)}(0)x^J + \cdots + \frac{1}{k!} \sum_{|J|=k} F^{(J)}(0)x^J + R_k(x),
\]

or, more briefly,

\[
F(x) = \sum_{|J| \leq k} \frac{1}{|J|!} F^{(J)}(0)x^J + R_k(x),
\]

\[
(1.44)
\]
where

\[ R_k(x) = \frac{1}{k!} \sum_{|J|=k+1} \left( \int_0^1 (1 - s)^k F^{(J)}(sx) \, ds \right) x^J. \]

This gives Taylor’s formula with remainder for \( F \in C^{k+1}(\Omega) \), in the \( J \)-multi-index notation.

We also want to write the formula in the \( \alpha \)-multi-index notation. We have

\[ \sum_{|J|=k} F^{(J)}(tx)x^J = \sum_{|\alpha|=k} \nu(\alpha) F^{(\alpha)}(tx)x^{\alpha}, \]

where

\[ \nu(\alpha) = \#\{J : \alpha = \alpha(J)\}, \]

and we define the relation \( \alpha = \alpha(J) \) to hold provided the condition (1.37) holds, or equivalently provided \( x^J = x^\alpha \). Thus \( \nu(\alpha) \) is uniquely defined by

\[ \sum_{|\alpha|=k} \nu(\alpha)x^{\alpha} = \sum_{|J|=k} x^J = (x_1 + \cdots + x_n)^k. \]

One sees that, if \( |\alpha| = k \), then \( \nu(\alpha) \) is equal to the product of the number of combinations of \( k \) objects, taken \( \alpha_1 \) at a time, times the number of combinations of \( k - \alpha_1 \) objects, taken \( \alpha_2 \) at a time, \( \cdots \) times the number of combinations of \( k - (\alpha_1 + \cdots + \alpha_{n-1}) \) objects, taken \( \alpha_n \) at a time. Thus

\[ \nu(\alpha) = \frac{k}{\alpha_1!} \frac{(k - \alpha_1)}{\alpha_2!} \cdots \frac{(k - \alpha_1 - \cdots - \alpha_{n-1})}{\alpha_n!} = \frac{k!}{\alpha_1!\alpha_2!\cdots\alpha_n!}. \]

In other words, for \( |\alpha| = k \),

\[ \nu(\alpha) = \frac{k!}{\alpha!}, \text{ where } \alpha! = \alpha_1! \cdots \alpha_n! \]

Thus the Taylor formula (1.44) can be rewritten

\[ F(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} F^{(\alpha)}(0)x^{\alpha} + R_k(x), \]

where

\[ R_k(x) = \sum_{|\alpha|=k+1} \frac{k + 1}{\alpha!} \left( \int_0^1 (1 - s)^k F^{(\alpha)}(sx) \, ds \right) x^{\alpha}. \]
The formula (1.51)\textendash(1.52) holds for $F \in C^{k+1}$. It is significant that (1.51), with a variant of (1.52), holds for $F \in C^k$. In fact, for such $F$, we can apply (1.52) with $k$ replaced by $k-1$, to get

\begin{equation}
F(x) = \sum_{|\alpha| \leq k-1} \frac{1}{\alpha!} F^{(\alpha)}(0) x^\alpha + R_{k-1}(x),
\end{equation}

with

\begin{equation}
R_{k-1}(x) = \sum_{|\alpha| = k} \frac{k}{\alpha!} \left( \int_0^1 (1-s)^{k-1} F^{(\alpha)}(sx) \, ds \right) x^\alpha.
\end{equation}

We can add and subtract $F^{(\alpha)}(0)$ to $F^{(\alpha)}(sx)$ in the integrand above, and obtain the following.

**Proposition 1.5.** If $F \in C^k$ on a ball $B_r(0)$, the formula (1.51) holds for $x \in B_r(0)$, with

\begin{equation}
R_k(x) = \sum_{|\alpha| = k} \frac{k}{\alpha!} \left( \int_0^1 (1-s)^{k-1} [F^{(\alpha)}(sx) - F^{(\alpha)}(0)] \, ds \right) x^\alpha.
\end{equation}

**Remark.** Note that (1.55) yields the estimate

\begin{equation}
|R_k(x)| \leq \sum_{|\alpha| = k} \frac{|x^\alpha|}{\alpha!} \sup_{0 \leq s \leq 1} |F^{(\alpha)}(sx) - F^{(\alpha)}(0)|.
\end{equation}

The term corresponding to $|J| = 2$ in (1.44), or $|\alpha| = 2$ in (1.51), is of particular interest. It is

\begin{equation}
\frac{1}{2} \sum_{|J|=2} F^{(J)}(0) x^J = \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2 F}{\partial x_k \partial x_j}(0) x_j x_k.
\end{equation}

We define the **Hessian** of a $C^2$ function $F : \mathcal{O} \to \mathbb{R}$ as an $n \times n$ matrix:

\begin{equation}
D^2 F(y) = \left( \frac{\partial^2 F}{\partial x_k \partial x_j}(y) \right).
\end{equation}

Then the power series expansion of second order about 0 for $F$ takes the form

\begin{equation}
F(x) = F(0) + DF(0)x + \frac{1}{2} x \cdot D^2 F(0) x + R_2(x),
\end{equation}

with
where, by (1.56),
\begin{equation}
|R_2(x)| \leq C_n |x|^2 \sup_{0 \leq s \leq 1, |\alpha|=2} |F^{(\alpha)}(sx) - F^{(\alpha)}(0)|.
\end{equation}

In all these formulas we can translate coordinates and expand about \( y \in \mathcal{O} \). For example, (1.59) extends to
\begin{equation}
F(x) = F(y) + DF(y)(x - y) + \frac{1}{2} (x - y) \cdot D^2F(y)(x - y) + R_2(x, y),
\end{equation}
with
\begin{equation}
|R_2(x, y)| \leq C_n |x - y|^2 \sup_{0 \leq s \leq 1, |\alpha|=2} |F^{(\alpha)}(y + s(x - y)) - F^{(\alpha)}(y)|.
\end{equation}

**Example.** If we take \( F(x) \) as in (1.8), so \( DF(x) \) is as in (1.9), then
\[
D^2F(x) = \begin{pmatrix}
-\sin x_1 & \sin x_2 \\
\cos x_1 & \cos x_2 \\
\sin x_1 & -\sin x_2
\end{pmatrix}.
\]

The results (1.61)–(1.62) are useful for extremal problems, i.e., determining where a sufficiently smooth function \( F : \mathcal{O} \to \mathbb{R} \) has local maxima and local minima. Clearly if \( F \in C^1(\mathcal{O}) \) and \( F \) has a local maximum or minimum at \( x_0 \in \mathcal{O} \), then \( DF(x_0) = 0 \). In such a case, we say \( x_0 \) is a **critical point** of \( F \). To check what kind of critical point \( x_0 \) is, we look at the \( n \times n \) matrix \( A = D^2F(x_0) \), assuming \( F \in C^2(\mathcal{O}) \). By Proposition 1.2, \( A \) is a symmetric matrix. A basic result in linear algebra is that if \( A \) is a real, symmetric \( n \times n \) matrix, then \( \mathbb{R}^n \) has an orthonormal basis of eigenvectors, \( \{v_1, \ldots, v_n\} \), satisfying \( Av_j = \lambda_j v_j \); the real numbers \( \lambda_j \) are the eigenvalues of \( A \). We say \( A \) is positive definite if all \( \lambda_j > 0 \), and we say \( A \) is negative definite if all \( \lambda_j < 0 \). We say \( A \) is strongly indefinite if some \( \lambda_j > 0 \) and another \( \lambda_k < 0 \). Equivalently, given a real, symmetric matrix \( A \),
\begin{equation}
A \text{ positive definite } \iff v \cdot Av \geq C|v|^2,
\end{equation}
\begin{equation}
A \text{ negative definite } \iff v \cdot Av \leq -C|v|^2,
\end{equation}
for some \( C > 0 \), all \( v \in \mathbb{R}^n \), and
\begin{equation}
A \text{ strongly indefinite } \iff \exists v, w \in \mathbb{R}^n, \text{ nonzero, such that } v \cdot Av \geq C|v|^2, \quad w \cdot Aw \leq -C|w|^2,
\end{equation}
for some \( C > 0 \). In light of (1.44)–(1.45), we have the following result.

**Proposition 1.6.** Assume \( F \in C^2(\mathcal{O}) \) is real valued, \( \mathcal{O} \) open in \( \mathbb{R}^n \). Let \( x_0 \in \mathcal{O} \) be a critical point for \( F \). Then
\begin{enumerate}[(i)]
\item \( D^2F(x_0) \) positive definite \( \Rightarrow \) \( F \) has a local minimum at \( x_0 \),
\item \( D^2F(x_0) \) negative definite \( \Rightarrow \) \( F \) has a local maximum at \( x_0 \),
\item \( D^2F(x_0) \) strongly indefinite \( \Rightarrow \) \( F \) has neither a local maximum nor a local minimum at \( x_0 \).
\end{enumerate}

In case (iii), we say \( x_0 \) is a **saddle point** for \( F \).

The following is a test for positive definiteness.
Proposition 1.7. Let $A = (a_{ij})$ be a real, symmetric, $n \times n$ matrix. For $1 \leq \ell \leq n$, form the $\ell \times \ell$ matrix $A_{\ell} = (a_{ij})_{1 \leq i,j \leq \ell}$. Then

\[(1.65) \quad A \text{ positive definite } \iff \det A_{\ell} > 0, \quad \forall \ell \in \{1, \ldots, n\}.\]

Regarding the implication $\Rightarrow$, note that if $A$ is positive definite, then $\det A = \det A_n$ is the product of its eigenvalues, all $> 0$, hence is $> 0$. Also in this case, it follows from the hypothesis on the left side of (1.65) that each $A_{\ell}$ must be positive definite, hence have positive determinant, so we have $\Rightarrow$.

The implication $\Leftarrow$ is easy enough for $2 \times 2$ matrices. If $A$ is symmetric and $\det A > 0$, then either both its eigenvalues are positive (so $A$ is positive definite) or both are negative (so $A$ is negative definite). In the latter case, $A_1 = (a_{11})$ must be negative, so we have $\Leftarrow$ in this case.

We prove $\Leftarrow$ for $n \geq 3$, using induction. The inductive hypothesis implies that if $\det A_{\ell} > 0$ for each $\ell \leq n$, then $A_{n-1}$ is positive definite. The next lemma then guarantees that $A = A_n$ has at least $n-1$ positive eigenvalues. The hypothesis that $\det A > 0$ does not allow that the remaining eigenvalue be $\leq 0$, so all the eigenvalues of $A$ must be positive. Thus Proposition 1.7 is proven, once we have the following.

Lemma 1.8. In the setting of Proposition 1.7, if $A_{n-1}$ is positive definite, then $A = A_n$ has at least $n-1$ positive eigenvalues.

Proof. Since $A$ is symmetric, $\mathbb{R}^n$ has an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of $A$; $Av_j = \lambda_j v_j$. If the conclusion of the lemma is false, at least two of the eigenvalues, say $\lambda_1, \lambda_2$, are $\leq 0$. Let $W = \text{Span}(v_1, v_2)$, so

\[w \in W \implies w \cdot Aw \leq 0.\]

Since $W$ has dimension 2, $\mathbb{R}^{n-1} \subset \mathbb{R}^n$ satisfies $\mathbb{R}^{n-1} \cap W \neq 0$, so there exists a nonzero $w \in \mathbb{R}^{n-1} \cap W$, and then

\[w \cdot A_{n-1}w = w \cdot Aw \leq 0,
\]

contradicting the hypothesis that $A_{n-1}$ is positive definite.

Remark. Given (1.65), we see by taking $A \mapsto -A$ that if $A$ is a real, symmetric $n \times n$ matrix,

\[(1.66) \quad A \text{ negative definite } \iff (-1)^\ell \det A_{\ell} > 0, \quad \forall \ell \in \{1, \ldots, n\}.
\]

We return to higher order power series formulas with remainder and complement Proposition 1.5. Let us go back to (1.39)–(1.40) and note that the integral in (1.40) is $1/(k+1)$ times a weighted average of $\varphi^{(k+1)}(s)$ over $s \in [0, t]$. Hence we can write

\[r_k(t) = \frac{1}{(k+1)!} \varphi^{(k+1)}(\theta t), \quad \text{for some } \theta \in [0, 1],\]
if \( \varphi \) is of class \( C^{k+1} \). This is the Lagrange form of the remainder; see Appendix D for more on this, and for a comparison with the Cauchy form of the remainder. If \( \varphi \) is of class \( C^k \), we can replace \( k+1 \) by \( k \) in (1.39) and write

\[
(1.67) \quad \varphi(t) = \varphi(0) + \varphi'(0)t + \cdots + \frac{1}{(k-1)!}\varphi^{(k-1)}(0)t^{k-1} + \frac{1}{k!}\varphi^{(k)}(\theta t)t^k,
\]

for some \( \theta \in [0, 1] \). Plugging (1.67) into (1.43) for \( \varphi(t) = F(tx) \) gives

\[
(1.68) \quad F(x) = \sum_{|J| \leq k-1} \frac{1}{|J|!} F^{(J)}(0)x^J + \frac{1}{k!} \sum_{|J|=k} F^{(J)}(\theta x)x^J,
\]

for some \( \theta \in [0, 1] \) (depending on \( x \) and on \( k \), but not on \( J \)), when \( F \) is of class \( C^k \) on a neighborhood \( B_r(0) \) of \( 0 \in \mathbb{R}^n \). Similarly, using the \( \alpha \)-multi-index notation, we have, as an alternative to (1.53)–(1.54),

\[
(1.69) \quad F(x) = \sum_{|\alpha| \leq k-1} \frac{1}{\alpha!} F^{(\alpha)}(0)x^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} F^{(\alpha)}(\theta x)x^\alpha,
\]

for some \( \theta \in [0, 1] \) (depending on \( x \) and on \( |\alpha| \), but not on \( \alpha \)), if \( F \in C^k(B_r(0)) \). Note also that

\[
(1.70) \quad \frac{1}{2} \sum_{|J|=2} F^{(J)}(\theta x)x^J = \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2 F}{\partial x_j \partial x_k}(\theta x)x_j x_k
\]

\[
= \frac{1}{2} x \cdot D^2 F(\theta x),
\]

with \( D^2 F(y) \) as in (1.58), so if \( F \in C^2(B_r(0)) \), we have, as an alternative to (1.59),

\[
(1.71) \quad F(x) = F(0) + DF(0)x + \frac{1}{2} x \cdot D^2 F(\theta x)x,
\]

for some \( \theta \in [0, 1] \).

We next complement the multi-index notations for higher derivatives of a function \( F \) by a multi-linear notation, defined as follows. If \( k \in \mathbb{N} \), \( F \in C^k(U) \), and \( y \in U \subset \mathbb{R}^n \), set

\[
(1.72) \quad D^k F(y)(u_1, \ldots, u_k) = \left. \partial_{t_1} \cdots \partial_{t_k} F(y + t_1 u_1 + \cdots + t_k u_k) \right|_{t_1 = \cdots = t_k = 0},
\]

for \( u_1, \ldots, u_k \in \mathbb{R}^n \). For \( k = 1 \), this formula is equivalent to the definition of \( DF \) given at the beginning of this section. For \( k = 2 \), we have

\[
(1.73) \quad D^2 F(y)(u, v) = u \cdot D^2 F(y)v,
\]
with $D^2F(y)$ on the right as in (1.58). Generally, (1.72) defines $D^kF(y)$ as a symmetric, $k$-linear form in $u_1, \ldots, u_k \in \mathbb{R}^n$.

We can relate (1.72) to the $J$-multi-index notation as follows. We start with
\[(1.74) \quad \partial_{t_1}\cdots\partial_{t_k}F(y + \Sigma t_j u_j) = \sum_{|J_1| = \cdots = |J_k| = 1} F^{(J_1+\cdots+J_k)}(y + \Sigma t_j u_j) u_1^{J_1}\cdots u_k^{J_k},\]
and inductively obtain
\[(1.75) \quad \partial_{t_1}\cdots\partial_{t_k}F(y + \Sigma t_j u_j) = \sum_{|J_1| = \cdots = |J_k| = 1} F^{(J_1+\cdots+J_k)}(y) u_1^{J_1}\cdots u_k^{J_k}.\]

In particular, if $u_1 = \cdots = u_k = u$, (1.77) yields
\[(1.76) \quad D^kF(y)(u, \ldots, u) = \sum_{|J| = k} F^{(J)}(y) u^J.\]

Hence (1.68) yields
\[(1.78) \quad F(x) = F(0) + DF(0)x + \cdots + \frac{1}{(k-1)!} D^{k-1}F(0)(x, \ldots, x) + \frac{1}{k!} D^kF(\theta x)(x, \ldots, x),\]
for some $\theta \in [0, 1]$, if $F \in C^k(B_r(0))$. In fact, rather than appealing to (1.68), we can note that
\[\varphi(t) = F(tx) \implies \varphi^{(k)}(t) = \partial_{t_1}\cdots\partial_{t_k} \varphi(t + t_1 + \cdots + t_k) \bigg|_{t_1 = \cdots = t_k = 0} = D^kF(tx)(x, \ldots, x),\]
and obtain (1.78) directly from (1.67). We can also use the notation
\[D^jF(y) x^{\otimes j} = D^jF(y)(x, \ldots, x),\]
with $j$ copies of $x$ within the last set of parentheses, and rewrite (1.78) as
\[(1.79) \quad F(x) = F(0) + DF(0)x + \cdots + \frac{1}{(k-1)!} D^{k-1}F(0)x^{\otimes(k-1)} + \frac{1}{k!} D^kF(\theta x)x^{\otimes k}.\]

Note how (1.78) and (1.79) generalize (1.71).

**Convergent power series and their derivatives**

Here we consider functions given by convergent power series, of the form
\[(1.80) \quad F(x) = \sum_{\alpha \geq 0} b_\alpha x^\alpha.\]
We allow $b_\alpha \in \mathbb{C}$, and take $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, with $x^\alpha$ given by (1.38). Here is our first result.
Proposition 1.9. Assume there exist \( y \in \mathbb{R}^n \) and \( C_0 < \infty \) such that

\[
|y_k| = a_k > 0, \forall k, \quad |b_\alpha y^\alpha| \leq C_0, \forall \alpha.
\]

Then, for each \( \delta \in (0,1) \), the series (1.80) converges absolutely and uniformly on each set

\[
R_\delta = \{ x \in \mathbb{R}^n : |x_k| \leq (1 - \delta)a_k, \forall k \}.
\]

The sum \( F(x) \) is continuous on \( \tilde{R} = \{ x \in \mathbb{R}^n : |x_k| < a_k, \forall k \} \).

Proof. We have

\[
x \in R_\delta \implies |b_\alpha x^\alpha| \leq C_0(1 - \delta)^{|\alpha|}, \forall \alpha,
\]

hence

\[
\sum_{\alpha \geq 0} |b_\alpha x^\alpha| \leq C_0 \sum_{\alpha \geq 0} (1 - \delta)^{|\alpha|} < \infty.
\]

Thus the power series (1.80) is absolutely convergent whenever \( x \in R_\delta \). We also have, for each \( N \in \mathbb{N} \),

\[
F(x) = \sum_{|\alpha| \leq N} b_\alpha x^\alpha + R_N(x),
\]

and, for \( x \in R_\delta \),

\[
|R_N(x)| \leq \sum_{|\alpha| > N} |b_\alpha x^\alpha|
\]

\[
\leq C_0 \sum_{|\alpha| > N} (1 - \delta)^{|\alpha|}
\]

\[
= \varepsilon_N \to 0 \text{ as } N \to \infty.
\]

This shows that \( R_N(x) \to 0 \) uniformly for \( x \in R_\delta \), and completes the proof of Proposition 1.9.

We next discuss differentiability.

Proposition 1.10. In the setting of Proposition 1.9, \( F \) is differentiable on \( \tilde{R} \) and, for each \( j \in \{1, \ldots, n\} \),

\[
\frac{\partial F}{\partial x_j}(x) = \sum_{\alpha \geq \varepsilon_j} \alpha_j b_\alpha x^{\alpha - \varepsilon_j}, \quad \forall x \in \tilde{R}.
\]

Here, we set \( \varepsilon_j = (0, \ldots, 1, \ldots, 0) \), with the 1 in the \( j \)th slot. It is convenient to begin the proof of Proposition 1.10 with the following.
Lemma 1.11. In the setting of Proposition 1.9, for each \( j \in \{1, \ldots, n\} \),
\[
G_j(x) = \sum_{\alpha \geq \varepsilon_j} \alpha_j b_\alpha x^{\alpha - \varepsilon_j}
\]
is absolutely convergent for \( x \in \tilde{R} \), uniformly on \( R_\delta \) for each \( \delta \in (0,1) \), therefore defining \( G_j \) as a continuous function on \( \tilde{R} \).

Proof. Take \( a = (a_1, \ldots, a_n) \), with \( a_j \) as in (1.81). Given \( x \in R_\delta \), we have
\[
\sum_{\alpha \geq \varepsilon_j} \alpha_j |b_\alpha x^{\alpha - \varepsilon_j}| \leq \sum_{\alpha \geq \varepsilon_j} \alpha_j (1 - \delta)^{|\alpha| - 1} |b_\alpha a^{\alpha - \varepsilon_j}|
\]
(1.89)
\[
\leq \frac{C_0}{a_j (1 - \delta)} \sum_{\alpha \geq 0} \alpha_j (1 - \delta)^{|\alpha|},
\]
and this is
\[
\leq M_\delta < \infty, \quad \forall \delta \in (0,1).
\]
This gives the asserted convergence on \( R_\delta \) and hence defines the function \( G_j \), continuous on \( \tilde{R} \).

To prove Proposition 1.10, we need to show that
\[
\frac{\partial F}{\partial x_j} = G_j \quad \text{on} \quad \tilde{R},
\]
(1.91)
for each \( j \). Let us use the notation
\[
\tilde{x}_j = (x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_n) = x - x_j e_j,
\]
(1.92)
where \( e_j \) is the \( j \)th standard basis vector of \( \mathbb{R}^n \). Now, given \( x \in R_\delta, \delta \in (0,1) \), the uniform convergence of (1.88) on \( R_\delta \) implies
\[
\int_0^{x_j} G_j(\tilde{x}_j + te_j) \, dt = \sum_{\alpha \geq \varepsilon_j} \alpha_j b_\alpha \int_0^{x_j} (\tilde{x}_j + te_j)^{\alpha - \varepsilon_j} \, dt
\]
(1.93)
\[
= \sum_{\alpha \geq \varepsilon_j} \alpha_j b_\alpha \alpha_j^{-1} x^{\alpha}
\]
\[
= \sum_{\alpha \geq \varepsilon_j} b_\alpha x^{\alpha}
\]
\[
= F(x) - F(\tilde{x}_j).
\]
Applying \( \partial/\partial x_j \) to the left side of (1.93) and using the fundamental theorem of calculus then yields (1.91) as desired. This gives the identity (1.87). Since each \( G_j \) is continuous on \( \tilde{R} \), this implies \( F \) is differentiable on \( \tilde{R} \).

We can iterate Proposition 1.10, obtaining \( \partial_k \partial_j F(x) = \partial_k G_j(x) \) as a convergent power series on \( \tilde{R} \), etc. In particular, we have the following.
Corollary 1.12. In the setting of Proposition 1.9, we have $F \in C^\infty(\mathbb{R})$.

Exercises

1. Consider the following function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$:

   $$f(x, y) = (\cos x)(\cos y).$$

   Find all its critical points, and determine which of these are local maxima, local minima, and saddle points.

2. Let $M(n, \mathbb{R})$ denote the space of real $n \times n$ matrices. Assume $F, G : M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R})$ are of class $C^1$. Show that $H(X) = F(X)G(X)$ defines a $C^1$ map $H : M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R})$, and

   $$DH(X)Y = DF(X)YG(X) + F(X)DG(X)Y.$$

3. Let $Gl(n, \mathbb{R}) \subset M(n, \mathbb{R})$ denote the set of invertible matrices. Show that

   $$\Phi : Gl(n, \mathbb{R}) \rightarrow M(n, \mathbb{R}), \quad \Phi(X) = X^{-1}$$

   is of class $C^1$ and that

   $$D\Phi(X)Y = -X^{-1}YX^{-1}.$$

4. Identify $\mathbb{R}^2$ and $\mathbb{C}$ via $z = x + iy$. Then multiplication by $i$ on $\mathbb{C}$ corresponds to applying

   $$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

   Let $O \subset \mathbb{R}^2$ be open, $f : O \rightarrow \mathbb{R}^2$ be $C^1$. Say $f = (u, v)$. Regard $Df(x, y)$ as a $2 \times 2$ real matrix. One says $f$ is holomorphic, or complex-analytic, provided the Cauchy-Riemann equations hold:

   $$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

   Show that this is equivalent to the condition

   $$Df(x, y)J = JDf(x, y).$$

   Generalize to $O$ open in $\mathbb{C}^n$, $f : O \rightarrow \mathbb{C}^n$.

5. Let $f$ be $C^1$ on a region in $\mathbb{R}^2$ containing $[a, b] \times \{y\}$. Show that, as $h \rightarrow 0$,

   $$\frac{1}{h}[f(x, y + h) - f(x, y)] \rightarrow \frac{\partial f}{\partial y}(x, y), \text{ uniformly on } [a, b] \times \{y\}. $$
Hint. Show that the left side is equal to
\[ \frac{1}{h} \int_{0}^{h} \frac{\partial f}{\partial y}(x, y + s) \, ds, \]
and use the uniform continuity of \( \frac{\partial f}{\partial y} \) on \([a, b] \times [y - \delta, y + \delta]\); cf. Proposition A.15.

6. In the setting of Exercise 5, show that
\[ \frac{d}{dy} \int_{a}^{b} f(x, y) \, dx = \int_{a}^{b} \frac{\partial f}{\partial y}(x, y) \, dx. \]

7. Considering the power series
\[ f(x) = f(y) + f'(y)(x - y) + \cdots + \frac{f^{(j)}(y)}{j!} (x - y)^{j} + R_{j}(x, y), \]
show that
\[ \frac{\partial R_{j}}{\partial y} = -\frac{1}{j!} f^{(j+1)}(y)(x - y)^{j}, \quad R_{j}(x, x) = 0. \]
Use this to re-derive (0.49), and hence (1.39)–(1.40).

We define “big oh” and “little oh” notation:
\[ f(x) = O(x) \quad (\text{as } x \to 0) \iff \left| \frac{f(x)}{x} \right| \leq C \quad \text{as } x \to 0, \]
\[ f(x) = o(x) \quad (\text{as } x \to 0) \iff \frac{f(x)}{x} \to 0 \quad \text{as } x \to 0. \]

8. Let \( \mathcal{O} \subset \mathbb{R}^{n} \) be open and \( y \in \mathcal{O} \). Show that
\[ f \in C^{k+1}({\mathcal{O}}) \Rightarrow f(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} f^{(\alpha)}(y)(x - y)^{\alpha} + O(|x - y|^{k+1}), \]
\[ f \in C^{k}({\mathcal{O}}) \Rightarrow f(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} f^{(\alpha)}(y)(x - y)^{\alpha} + o(|x - y|^{k}). \]

9. Assume \( G : U \to \mathcal{O}, \; F : \mathcal{O} \to \Omega \). Show that
\[ (1.94) \quad F, G \in C^{1} \iff F \circ G \in C^{1}. \]
More generally, show that, for \( k \in \mathbb{N} \),
\[ (1.95) \quad F, G \in C^{k} \iff F \circ G \in C^{k}. \]
Hint. Write \( H = F \circ G \), with \( h_\ell(x) = f_\ell(g_1(x), \ldots, g_n(x)) \), and use (1.25) to get

\[
\partial_j h_\ell(x) = \sum_{k=1}^n \partial_k f_\ell(g_1, \ldots, g_n) \partial_j g_k.
\]  

Show that this yields (1.94). To proceed, deduce from (1.96) that

\[
\partial_{j_1} \partial_{j_2} h_\ell(x) = \sum_{k_1, k_2=1}^n \partial_{k_1} \partial_{k_2} f_\ell(g_1, \ldots, g_n) (\partial_{j_1} g_{k_1})(\partial_{j_2} g_{k_2})
\]

\[+ \sum_{k=1}^n \partial_k f_\ell(g_1, \ldots, g_n) \partial_{j_1} \partial_{j_2} g_k.
\]

Use this to get (1.95) for \( k = 2 \). Proceeding inductively, show that there exist constants \( C(\mu, J^\#, k^\#) = C(\mu, J_1, \ldots, J_\mu, k_1, \ldots, k_\mu) \) such that if \( F, G \in C^k \) and \( |J| \leq k \),

\[
h_\ell^{(J)}(x) = \sum C(\mu, J^\#, k^\#) g_{k_1}^{(J_1)} \ldots g_{k_\mu}^{(J_\mu)} f_\ell^{(k_1, \ldots, k_\mu)}(g_1, \ldots, g_n),
\]

where the sum is over \( \mu \leq |J|, J_1 + \cdots + J_\mu \sim J, |J_\nu| \geq 1 \),

and \( J_1 + \cdots + J_\mu \sim J \) means \( J \) is a rearrangement of \( J_1 + \cdots + J_\mu \). Show that (1.95) follows from this.

10. Show that the map \( \Phi : Gl(n, \mathbb{R}) \to Gl(n, \mathbb{R}) \) given by \( \Phi(X) = X^{-1} \) is \( C^k \) for each \( k \), i.e., \( \Phi \in C^\infty \).

**Hint.** Start with the material of Exercise 3. Write \( D\Phi(X)Y = -X^{-1}XYX^{-1} \) as

\[
\partial_{x_{\ell m}} \Phi(X) = \frac{\partial}{\partial x_{\ell m}} \Phi(X) = D\Phi(X)E_{\ell m} = -\Phi(X)E_{\ell m} \Phi(X),
\]

where \( X = (x_{\ell m}) \) and \( E_{\ell m} \) has just one nonzero entry, at position \((\ell, m)\). Iterate this to get

\[
\partial_{\ell_2 m_2} \partial_{\ell_1 m_1} \Phi(X) = -(\partial_{\ell_2 m_2} \Phi(X))E_{\ell_1 m_1} \Phi(X) - \Phi(X)E_{\ell_1 m_1} (\partial_{\ell_2 m_2} \Phi(X)),
\]

and continue.

**Auxiliary exercises on determinants**

If \( M(n, \mathbb{F}) \) denotes the space of \( n \times n \) matrices with coefficients in \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{C} \), we want to show that there is a map

\[
det : M(n, \mathbb{F}) \to \mathbb{F}
\]
which is uniquely specified as a function $\vartheta : M(n, F) \to F$ satisfying:

(a) $\vartheta$ is linear in each column $a_j$ of $A$.
(b) $\vartheta(\bar{A}) = -\vartheta(A)$ if $\bar{A}$ is obtained from $A$ by interchanging two columns.
(c) $\vartheta(I) = 1$.

A detailed presentation of results in these exercises can be found in §5 of [T7].

1. Let $A = (a_1, \ldots, a_n)$, where $a_j$ are column vectors; $a_j = (a_{1j}, \ldots, a_{nj})^t$. Show that, if (a) holds, we have the expansion

$$\det A = \sum_j a_{j1} \det (e_j, a_2, \ldots, a_n) = \cdots$$

$$= \sum_{j_1, \ldots, j_n} a_{j_11} \cdots a_{j_n n} \det (e_{j_1}, e_{j_2}, \ldots, e_{j_n}),$$

where $\{e_1, \ldots, e_n\}$ is the standard basis of $F^n$.

2. Show that, if (b) and (c) also hold, then

$$\det A = \sum_{\sigma \in S_n} (\text{sgn } \sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n},$$

where $S_n$ is the set of permutations of $\{1, \ldots, n\}$, and

$$\text{sgn } \sigma = \det (e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \pm 1.$$

To define $\text{sgn } \sigma$, the “sign” of a permutation $\sigma$, we note that every permutation $\sigma$ can be written as a product of transpositions: $\sigma = \tau_1 \cdots \tau_\nu$, where a transposition of $\{1, \ldots, n\}$ interchanges two elements and leaves the rest fixed. We say $\text{sgn } \sigma = 1$ if $\nu$ is even and $\text{sgn } \sigma = -1$ if $\nu$ is odd. It is necessary to show that $\text{sgn } \sigma$ is independent of the choice of such a product representation. (Referring to (1.102) begs the question until we know that $\det$ is well defined.)

3. Let $\sigma \in S_n$ act on a function of $n$ variables by

$$(\sigma f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Let $P$ be the polynomial

$$P(x_1, \ldots, x_n) = \prod_{1 \leq j < k \leq n} (x_j - x_k).$$

Show that

$$(\sigma P)(x) = (\text{sgn } \sigma) P(x),$$
and that this implies that \( \text{sgn} \sigma \) is well defined.

4. Deduce that there is a unique determinant satisfying (a)–(c), and that it is given by (1.101).

5. Show that (1.101) implies

\[
\det A = \det A^t.
\]

Conclude that one can replace columns by rows in the characterization (a)–(c) of determinants.

**Hint.** \( a_{\sigma(j) \ell} = a_{\ell \tau(\ell)} \) with \( \ell = \sigma(j), \tau = \sigma^{-1} \). Also, \( \text{sgn} \sigma = \text{sgn} \tau \).

6. Show that, if (a)–(c) hold (for rows), it follows that

(d) \( \vartheta(\tilde{A}) = \vartheta(A) \) if \( \tilde{A} \) is obtained from \( A \) by adding \( c \rho_k \) to \( \rho_k \), for some \( c \in \mathbb{F} \), where \( \rho_1, \ldots, \rho_n \) are the rows of \( A \).

Re-prove the uniqueness of \( \vartheta \) satisfying (a)–(d) (for rows) by applying row operations to \( A \) until either some row vanishes or \( A \) is converted to \( I \).

7. Show that

\[
\text{det}(BA) = (\text{det} B)(\text{det} A).
\]

**Hint.** For fixed \( A \in M(n, \mathbb{F}) \), compare \( \vartheta_1(A) = \text{det}(BA) \) and \( \vartheta_2(A) = (\text{det} B)(\text{det} A) \).

For uniqueness, show that if (c) is modified to

(c') \( \vartheta(I) = r \),

then arguments in Exercises 1–4 yield \( \vartheta(A) = r \det A \).

8. Given \( A \in M(n, \mathbb{F}) \), show that if \( A \) is invertible, then \( \det A \not= 0 \).

**Hint.** \( \det(AB^{-1}) = (\det A)(\det B^{-1}) \).

If \( A \) is invertible, we say \( A \in \text{Gl}(n, \mathbb{F}) \). If \( \mathbb{F} = \mathbb{R} \), we say

\( A \in \text{Gl}_+(n, \mathbb{R}) \) provided \( \det A > 0 \).

9. Given \( A \in M(n, \mathbb{F}) \), show that if \( \det A \not= 0 \), then \( A \) is invertible.

**Hint.** If \( A \) is not invertible, there is a linear dependence relation among its rows. Use Exercise 6 to show that, in such a case, there exists \( \tilde{A} \in M(n, \mathbb{F}) \), obtained by a sequence of transformations of the form (d), having a row \((0, \ldots, 0)\), so \( \det \tilde{A} = 0 \).

**Conclusion.** Given \( A \in M(n, \mathbb{F}) \), \( A \) is invertible \( \iff \det A \not= 0 \). This result was used in (1.16).

10. Show that

\[
\det \begin{pmatrix}
1 & a_{12} & \cdots & a_{1n} \\
0 & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
0 & a_{n2} & \cdots & a_{nn}
\end{pmatrix} = \det \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
0 & a_{n2} & \cdots & a_{nn}
\end{pmatrix} = \det A_{11}
\]
where $A_{11} = (a_{jk})_{2 \leq j, k \leq n}.$

*Hint.* Do the first identity by the analogue of (d), for columns. Then exploit uniqueness for det on $M(n-1, \mathbb{F}).$

11. Deduce that $\det(e_j, a_2, \ldots, a_n) = (-1)^{j-1} \det A_{1j}$ where $A_{kj}$ is formed by deleting the $k^{th}$ column and the $j^{th}$ row from $A.$

12. Deduce from the first sum in (1.100) that

$$\det A = \sum_{j=1}^{n} (-1)^{j-1} a_{j1} \det A_{1j}. \quad (1.109)$$

More generally, for any $k \in \{1, \ldots, n\},$

$$\det A = \sum_{j=1}^{n} (-1)^{j-k} a_{jk} \det A_{kj}. \quad (1.110)$$

This is called an expansion of det $A$ by minors, down the $k^{th}$ column. By definition, the *cofactor matrix* of $A$ is given by

$$\text{Cof}(A)_{jk} = (-1)^{j-k} \det A_{kj}.$$ 

13. Show that

$$\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{nn} \end{pmatrix} = a_{11}a_{22}\cdots a_{nn}. \quad (1.111)$$

*Hint.* Use (1.108) and induction. *Alternative.* Apply (1.101).

Exercises 14–16 deal with properties of the determinant, as a differentiable function on spaces of matrices.

14. Let $M(n, \mathbb{F})$ be the space of $n \times n$ matrices with coefficients in $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, det : $M(n, \mathbb{F}) \to \mathbb{F}$ the determinant. Show that, if $I$ is the identity matrix,

$$D \det(I)B = \text{Tr} B,$$

i.e.,

$$\frac{d}{dt} \det(I + tB)|_{t=0} = \text{Tr} B.$$
15. If $A(t) = (a_{jk}(t))$ is a curve in $M(n, F)$, use the expansion of $(d/dt)\det A(t)$ as a sum of $n$ determinants, in which the rows of $A(t)$ are successively differentiated, to show that

$$\frac{d}{dt} \det A(t) = \text{Tr} \left( \text{Cof}(A(t))^t \cdot A'(t) \right),$$

and deduce that, for $A, B \in M(n, F)$,

$$D \det(A)B = \text{Tr}(\text{Cof}(A)^t \cdot B).$$

16. Suppose $A \in M(n, F)$ is invertible. Using

$$\det(A + tB) = (\det A) \det(I + tA^{-1}B),$$

show that

$$D \det(A)B = (\det A) \text{Tr}(A^{-1}B).$$

Comparing the result of Exercise 15, deduce Cramer’s formula:

$$(\det A)A^{-1} = \text{Cof}(A)^t.$$

**Auxiliary exercises on the cross product**

1. If $u, v \in \mathbb{R}^3$, show that the formula

$$w \cdot (u \times v) = \det \begin{pmatrix} w_1 & u_1 & v_1 \\ w_2 & u_2 & v_2 \\ w_3 & u_3 & v_3 \end{pmatrix}$$

for $u \times v = \Pi(u, v)$ defines uniquely a bilinear map $\Pi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$. Show that it satisfies

$$i \times j = k, \quad j \times k = i, \quad k \times i = j,$$

where $\{i, j, k\}$ is the standard basis of $\mathbb{R}^3$.

2. We say $T \in SO(3)$ provided that $T$ is a real $3 \times 3$ matrix satisfying $T^tT = I$ and $\det T > 0$, (hence $\det T = 1$). Show that

$$T \in SO(3) \implies Tu \times Tv = T(u \times v).$$

*Hint.* Multiply the $3 \times 3$ matrix in Exercise 1 on the left by $T$. 

3. Show that, if $\theta$ is the angle between $u$ and $v$ in $\mathbb{R}^3$, then

$$|u \times v| = |u| |v| \sin \theta.$$  

*Hint.* Check this for $u = i$, $v = ai + bj$, and use Exercise 2 to show this suffices.

4. Show that, for all $u, v, w, x \in \mathbb{R}^3$, 

$$ (u \times v) \cdot (w \times x) = \det \begin{pmatrix} u \cdot w & v \cdot w \\ u \cdot x & v \cdot x \end{pmatrix}.$$  

Taking $w = u, x = v$, show that this implies (1.114).  

*Hint.* Using Exercise 2, show that it suffices to check this for 

$$w = i, x = ai + bj,$$

in which case $w \times x = bk$. Then the left side of (1.115) is 

$$ (u \times v) \cdot bk = \det \begin{pmatrix} 0 & u \cdot i & v \cdot i \\ 0 & u \cdot j & v \cdot j \\ b & u \cdot k & v \cdot k \end{pmatrix}.$$  

Show that this equals the right side of (1.115).

5. Show that $\kappa : \mathbb{R}^3 \to \text{Skew}(3)$, the set of antisymmetric real $3 \times 3$ matrices, given by 

$$ \kappa(y_1, y_2, y_3) = \begin{pmatrix} 0 & -y_3 & y_2 \\ y_3 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{pmatrix} $$

satisfies 

$$Kx = y \times x, \quad K = \kappa(y).$$

Show that, with $[A, B] = AB - BA$, 

$$\kappa(x \times y) = [\kappa(x), \kappa(y)],$$

$$\text{Tr}(\kappa(x)\kappa(y)^t) = 2x \cdot y.$$
2. Inverse function and implicit function theorem

The Inverse Function Theorem gives a condition under which a function can be locally inverted. This theorem and its corollary the Implicit Function Theorem are fundamental results in multivariable calculus. First we state the Inverse Function Theorem. Here, we assume $k \geq 1$.

**Theorem 2.1.** Let $F$ be a $C^k$ map from an open neighborhood $\Omega$ of $p_0 \in \mathbb{R}^n$ to $\mathbb{R}^n$, with $q_0 = F(p_0)$. Suppose the derivative $DF(p_0)$ is invertible. Then there is a neighborhood $U$ of $p_0$ and a neighborhood $V$ of $q_0$ such that $F : U \to V$ is one-to-one and onto, and $F^{-1} : V \to U$ is a $C^k$ map. (One says $F : U \to V$ is a diffeomorphism.)

First we show that $F$ is one-to-one on a neighborhood of $p_0$, under these hypotheses. In fact, we establish the following result, of interest in its own right.

**Proposition 2.2.** Assume $\Omega \subset \mathbb{R}^n$ is open and convex, and let $f : \Omega \to \mathbb{R}^n$ be $C^1$. Assume that the symmetric part of $Df(u)$ is positive-definite, for each $u \in \Omega$. Then $f$ is one-to-one on $\Omega$.

**Proof.** Take distinct points $u_1, u_2 \in \Omega$, and set $u_2 - u_1 = w$. Consider $\varphi : [0, 1] \to \mathbb{R}$, given by

$$\varphi(t) = w \cdot f(u_1 + tw).$$

Then $\varphi'(t) = w \cdot Df(u_1 + tw)w > 0$ for $t \in [0, 1]$, so $\varphi(0) \neq \varphi(1)$. But $\varphi(0) = w \cdot f(u_1)$ and $\varphi(1) = w \cdot f(u_2)$, so $f(u_1) \neq f(u_2)$.

To continue the proof of Theorem 2.1, let us set

$$(2.1) \quad f(u) = A(F(p_0 + u) - q_0), \quad A = DF(p_0)^{-1}.$$ 

Then $f(0) = 0$ and $Df(0) = I$, the identity matrix. We will show that $f$ maps a neighborhood of 0 one-to-one and onto some neighborhood of 0. We can write

$$f(u) = u + R(u), \quad R(0) = 0, \quad DR(0) = 0,$$

and $R$ is $C^1$. Pick $b > 0$ such that

$$(2.3) \quad \|u\| \leq 2b \implies \|DR(u)\| \leq \frac{1}{2}.$$ 

Then $Df = I + DR$ has positive definite symmetric part on

$$B_{2b}(0) = \{u \in \mathbb{R}^n : \|u\| < 2b\},$$

so by Proposition 2.2,

$$f : B_{2b}(0) \to \mathbb{R}^n$$

is one-to-one.
We will show that the range \( f(B_{2b}(0)) \) contains \( B_b(0) \), that is to say, we can solve
\[
(2.4) \quad f(u) = v,
\]
given \( v \in B_b(0) \), for some (unique) \( u \in B_{2b}(0) \). This is equivalent to \( u + R(u) = v \).

To get the solution, we set
\[
(2.5) \quad T_v(u) = v - R(u).
\]
Then solving (2.4) is equivalent to solving
\[
(2.6) \quad T_v(u) = u.
\]

We look for a fixed point
\[
(2.7) \quad u = K(v) = f^{-1}(v).
\]

Also, we want to show that \( DK(0) = I \), i.e., that
\[
(2.8) \quad K(v) = v + r(v), \quad r(v) = o(\|v\|).
\]

The “little oh” notation is defined in Exercise 8 of §1. If we succeed in doing this, it follows that, for \( y \) close to \( q_0 \), \( G(y) = F^{-1}(y) \) is defined. Also, taking
\[
x = p_0 + u, \quad y = F(x), \quad v = f(u) = A(y - q_0),
\]
as in (2.1), we have, via (2.8),
\[
G(y) = p_0 + u = p_0 + K(v)
= p_0 + K(A(y - q_0))
= p_0 + A(y - q_0) + o(\|y - q_0\|).
\]
Hence \( G \) is differentiable at \( q_0 \) and
\[
(2.9) \quad DG(q_0) = A = DF(p_0)^{-1}.
\]

A parallel argument, with \( p_0 \) replaced by a nearby \( x \) and \( y = F(x) \), gives
\[
(2.10) \quad DG(y) = DF(G(y))^{-1}.
\]

Thus our task is to solve (2.6). To do this, we use the following general result, known as the Contraction Mapping Theorem.
Theorem 2.3. Let $X$ be a complete metric space, and let $T : X \to X$ satisfy

\begin{equation}
\text{dist}(Tx, Ty) \leq r \text{ dist}(x, y),
\end{equation}

for some $r < 1$. (We say $T$ is a contraction.) Then $T$ has a unique fixed point $x$. For any $y_0 \in X$, $T^k y_0 \to x$ as $k \to \infty$.

Proof. Pick $y_0 \in X$ and let $y_k = T^k y_0$. Then $\text{dist}(y_k, y_{k+1}) \leq r^k \text{ dist}(y_0, y_1)$, so

\begin{equation}
\text{dist}(y_k, y_{k+m}) \leq \text{dist}(y_k, y_{k+1}) + \cdots + \text{dist}(y_{k+m-1}, y_{k+m}) \\
\leq (r^k + \cdots + r^{k+m-1}) \text{ dist}(y_0, y_1) \\
\leq r^k (1 - r)^{-1} \text{ dist}(y_0, y_1).
\end{equation}

It follows that $(y_k)$ is a Cauchy sequence, so it converges; $y_k \to x$. Since $Ty_k = y_{k+1}$ and $T$ is continuous, it follows that $Tx = x$, i.e., $x$ is a fixed point. Uniqueness of the fixed point is clear from the estimate $\text{dist}(Tx, Tx') \leq r \text{ dist}(x, x')$, which implies dist$(x, x') = 0$ if $x$ and $x'$ are fixed points. This proves Theorem 2.3.

Returning to the task of solving (2.6), having $b$ as in (2.3), we claim that

\begin{equation}
\|v\| \leq b \implies T_v : X_v \to X_v,
\end{equation}

where

\begin{equation}
X_v = \{u \in \overline{B_{2b}(0)} : \|u - v\| \leq A_v\}, \\
A_v = \sup_{\|w\| \leq 2\|v\|} \|R(w)\|.
\end{equation}

See Fig. 2.1. Note from (2.2)–(2.3) that

\[ \|w\| \leq 2b \implies \|R(w)\| \leq \frac{1}{2}\|w\|, \text{ and } \|R(w)\| = o(\|w\|). \]

Hence

\begin{equation}
\|v\| \leq b \implies A_v \leq \|v\|, \text{ and } A_v = o(\|v\|).
\end{equation}

Thus $\|u - v\| \leq A_v \Rightarrow u \in X_v$. Also

\begin{equation}
u \in X_v \implies \|u\| \leq 2\|v\| \\
\implies \|R(u)\| \leq A_v \\
\implies \|T_v(u) - v\| \leq A_v,
\end{equation}

so (2.13) holds.
As for the contraction property, given \( u_j \in X_b, \|v\| \leq b \),

\[
\|T_v(u_1) - T_v(u_2)\| = \|R(u_2) - R(u_1)\| \\
\leq \frac{1}{2}\|u_1 - u_2\|,
\]

(2.17)

the last inequality by (2.3), so the map (2.13) is a contraction. Hence, by Theorem 2.3, there is a unique fixed point, \( u = K(v) \in X_v \). Also, since \( u \in X_v \),

\[
\|K(v) - v\| \leq A_v = o(\|v\|).
\]

(2.18)

Thus we have (2.8). This establishes the existence of the inverse function \( G = F^{-1} : V \to U \), and we have the formula (2.10) for the derivative \( DG \). Since \( G \) is differentiable on \( V \), it is certainly continuous, so (2.10) implies \( DG \) is continuous, given \( F \in C^1(U) \).

To finish the proof of the Inverse Function Theorem and show that \( G \) is \( C^k \) if \( F \) is \( C^k \), for \( k \geq 2 \), one uses an inductive argument. See Exercise 6 at the end of this section for an approach to this last argument.

Thus if \( DF \) is invertible on the domain of \( F \), \( F \) is a local diffeomorphism. Stronger hypotheses are needed to guarantee that \( F \) is a global diffeomorphism onto its range. Proposition 2.2 provides one tool for doing this. Here is a slight strengthening.

**Corollary 2.4.** Assume \( \Omega \subset \mathbb{R}^n \) is open and convex, and that \( F : \Omega \to \mathbb{R}^n \) is \( C^1 \). Assume there exist \( n \times n \) matrices \( A \) and \( B \) such that the symmetric part of \( ADF(u)B \) is positive definite for each \( u \in \Omega \). Then \( F \) maps \( \Omega \) diffeomorphically onto its image, an open set in \( \mathbb{R}^n \).

**Proof.** Exercise.

We make a comment about solving the equation \( F(x) = y \), under the hypotheses of Theorem 2.1, when \( y \) is close to \( q_0 \). The fact that finding the fixed point for \( T_v \) in (2.13) is accomplished by taking the limit of \( T^k_v(v) \) implies that, when \( y \) is sufficiently close to \( q_0 \), the sequence \((x_k)\), defined by

\[
x_0 = p_0, \quad x_{k+1} = x_k + DF(p_0)^{-1}(y - F(x_k)),
\]

(2.19)

converges to the solution \( x \). An analysis of the rate at which \( x_k \to x \), and \( F(x_k) \to y \), can be made by applying \( F \) to (2.19), yielding

\[
F(x_{k+1}) = F(x_k + DF(p_0)^{-1}(y - F(x_k))) \\
= F(x_k) + DF(x_k)DF(p_0)^{-1}(y - F(x_k)) + R(x_k, DF(p_0)^{-1}(y - F(x_k))),
\]

and hence

\[
y - F(x_{k+1}) = (I - DF(x_k)DF(p_0)^{-1})(y - F(x_k)) + \tilde{R}(x_k, y - F(x_k)),
\]

(2.20)
with $\| \tilde{R}(x_k, y - F(x_k)) \| = o(\| y - F(x_k) \|)$.

It turns out that replacing $p_0$ by $x_k$ in (2.19) yields a faster approximation. This method, known as Newton’s method, is described in the exercises.

We consider some examples of maps to which Theorem 2.1 applies. First, we look at

$$F : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2, \quad F(r, \theta) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix} = \begin{pmatrix} x(r, \theta) \\ y(r, \theta) \end{pmatrix}.\tag{2.21}$$

Then

$$DF(r, \theta) = \begin{pmatrix} \partial_x x & \partial_x y \\ \partial_y x & \partial_y y \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix},\tag{2.22}$$

so

$$\det DF(r, \theta) = r \cos^2 \theta + r \sin^2 \theta = r.\tag{2.23}$$

Hence $DF(r, \theta)$ is invertible for all $(r, \theta) \in (0, \infty) \times \mathbb{R}$. Theorem 2.1 implies that each $(r_0, \theta_0) \in (0, \infty) \times \mathbb{R}$ has a neighborhood $U$ and $(x_0, y_0) = (r_0 \cos \theta_0, r_0 \sin \theta_0)$ has a neighborhood $V$ such that $F$ is a smooth diffeomorphism of $U$ onto $V$. In this simple situation, it can be verified directly that

$$F : (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}\tag{2.24}$$

is a smooth diffeomorphism.

Note that $DF(1, 0) = I$ in (2.22). Let us check the domain of applicability of Proposition 2.2. The symmetric part of $DF(r, \theta)$ in (2.22) is

$$S(r, \theta) = \begin{pmatrix} \cos \theta & \frac{1}{2} (1 - r) \sin \theta \\ \frac{1}{2} (1 - r) \sin \theta & r \cos \theta \end{pmatrix}.\tag{2.25}$$

By Proposition 1.7, this is positive definite if and only if

$$\cos \theta > 0,\tag{2.26}$$

and

$$\det S(r, \theta) = r \cos^2 \theta - \frac{1}{4} (1 - r)^2 \sin^2 \theta > 0.\tag{2.27}$$

Now (2.26) holds for $\theta \in (-\pi/2, \pi/2)$, but not on all of $(-\pi, \pi)$. Furthermore, (2.27) holds for $(r, \theta)$ in a neighborhood of $(r_0, \theta_0) = (1, 0)$, but it does not hold on all of $(0, \infty) \times (-\pi/2, \pi/2)$. We see that Proposition 2.2 does not capture the full force of (2.24).

We move on to another example. As in §1, we can extend Theorem 2.1, replacing $\mathbb{R}^n$ by a finite dimensional real vector space, isometric to a Euclidean space, such as $M(n, \mathbb{R}) \approx \mathbb{R}^{n^2}$. As an example, consider

$$\text{Exp} : M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R}), \quad \text{Exp}(X) = e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.\tag{2.28}$$
Smoothness of $\text{Exp}$ follows from Corollary 1.12. Since
\[(2.29) \quad \text{Exp}(Y) = I + Y + \frac{1}{2}Y^2 + \cdots,\]
we have
\[(2.30) \quad D\text{Exp}(0)Y = Y, \quad \forall Y \in M(n, \mathbb{R}),\]
so $D\text{Exp}(0)$ is invertible. Then Theorem 2.1 implies that there exist a neighborhood $U$ of $0 \in M(n, \mathbb{R})$ and a neighborhood $V$ of $I \in M(n, \mathbb{R})$ such that $\text{Exp} : U \to V$ is a smooth diffeomorphism.

To motivate the next result, we consider the following example. Take $a > 0$ and consider the equation
\[(2.31) \quad x^2 + y^2 = a^2, \quad F(x, y) = x^2 + y^2.\]
Note that
\[(2.32) \quad DF(x, y) = (2x \ 2y), \quad D_xF(x, y) = 2x, \quad D_yF(x, y) = 2y.\]
The equation (2.31) defines $y$ “implicitly” as a smooth function of $x$ if $|x| < a$. Explicitly,
\[(2.33) \quad |x| < a \implies y = \sqrt{a^2 - x^2},\]
Similarly, (2.31) defines $x$ implicitly as a smooth function of $y$ if $|y| < a$; explicitly
\[(2.34) \quad |y| < a \implies x = \sqrt{a^2 - y^2}.\]
Now, given $x_0 \in \mathbb{R}$, $a > 0$, there exists $y_0 \in \mathbb{R}$ such that $F(x_0, y_0) = a^2$ if and only if $|x_0| \leq a$. Furthermore,
\[(2.35) \quad \text{given} \quad F(x_0, y_0) = a^2, \quad D_yF(x_0, y_0) \neq 0 \iff |x_0| < a.\]
Similarly, given $y_0 \in \mathbb{R}$, there exists $x_0$ such that $F(x_0, y_0) = a^2$ if and only if $|y_0| \leq a$, and
\[(2.36) \quad \text{given} \quad F(x_0, y_0) = a^2, \quad D_xF(x_0, y_0) \neq 0 \iff |x_0| < a.\]
Note also that, whenever $(x, y) \in \mathbb{R}^2$ and $F(x, y) = a^2 > 0$,
\[(2.37) \quad DF(x, y) \neq 0,\]
so either $D_xF(x, y) \neq 0$ or $D_yF(x, y) \neq 0$, and, as seen above whenever $(x_0, y_0) \in \mathbb{R}^2$ and $F(x_0, y_0) = a^2 > 0$, we can solve $F(x, y) = a^2$ for either $y$ as a smooth function of $x$ for $x$ near $x_0$ or for $x$ as a smooth function of $y$ for $y$ near $y_0$.

We move from these observations to the next result, the Implicit Function Theorem.
**Theorem 2.5.** Suppose $U$ is a neighborhood of $x_0 \in \mathbb{R}^m$, $V$ a neighborhood of $y_0 \in \mathbb{R}^\ell$, and we have a $C^k$ map

\[ F : U \times V \to \mathbb{R}^\ell, \quad F(x_0, y_0) = u_0. \]

Assume $D_y F(x_0, y_0)$ is invertible. Then the equation $F(x, y) = u_0$ defines $y = g(x, u_0)$ for $x$ near $x_0$ (satisfying $g(x_0, u_0) = y_0$) with $g$ a $C^k$ map.

To prove this, consider $H : U \times V \to \mathbb{R}^m \times \mathbb{R}^\ell$ defined by

\[ H(x, y) = (x, F(x, y)). \]

(Actually, regard $(x, y)$ and $(x, F(x, y))$ as column vectors.) We have

\[ DH = \begin{pmatrix} I & 0 \\ D_x F & D_y F \end{pmatrix}. \]

Thus $DH(x_0, y_0)$ is invertible, so $J = H^{-1}$ exists, on a neighborhood of $(x_0, u_0)$, and is $C^k$, by the Inverse Function Theorem. It is clear that $J(x, u_0)$ has the form

\[ J(x, u_0) = (x, g(x, u_0)), \]

and $g$ is the desired map.

Here is an example where Theorem 2.5 applies. Set

\[ F : \mathbb{R}^4 \to \mathbb{R}^2, \quad F(u, v, x, y) = \begin{pmatrix} x(u^2 + v^2) \\ xu + yv \end{pmatrix}. \]

We have

\[ F(2, 0, 1, 1) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}. \]

Note that

\[ D_{u,v} F(u, v, x, y) = \begin{pmatrix} 2xu & 2xv \\ x & y \end{pmatrix}, \]

hence

\[ D_{u,v} F(2, 0, 1, 1) = \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix} \]

is invertible, so Theorem 2.5 (with $(u, v)$ in place of $y$ and $(x, y)$ in place of $x$) implies that the equation

\[ F(u, v, x, y) = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \]
defines smooth functions

\[(2.47) \quad u = u(x,y), \quad v = v(x,y),\]

for \((x,y)\) near \((x_0,y_0) = (1,1)\), satisfying \((2.46)\), with \((u(1,1), v(1,1)) = (2,0)\).

Let us next focus on the case \(\ell = 1\) of Theorem 2.5, so

\[(2.48) \quad z = (x,y) \in \mathbb{R}^n, \quad x \in \mathbb{R}^{n-1}, \quad y \in \mathbb{R}, \quad F(z) \in \mathbb{R}.\]

Then \(D_y F = \partial_y F\). If \(F(x_0,y_0) = u_0\), Theorem 2.5 says that if

\[(2.49) \quad \partial_y F(x_0,y_0) \neq 0,\]

then one can solve

\[(2.50) \quad F(x,y) = u_0 \text{ for } y = g(x,u_0),\]

for \(x\) near \(x_0\) (satisfying \(g(x_0,u_0) = y_0\)), with \(g\) a \(C^k\) function. This phenomenon was illustrated in \((2.31)-(2.35)\). To generalize the observations involving \((2.36)-(2.37)\), we note the following. Set \((x,y) = z = (z_1,\ldots,z_n), \ z_0 = (x_0,y_0)\). The condition \((2.49)\) is that \(\partial_{z_n} F(z_0) \neq 0\). Now a simple permutation of variables allows us to assume

\[(2.51) \quad \partial_{z_j} F(z_0) \neq 0, \quad F(z_0) = u_0,\]

and deduce that one can solve

\[(2.52) \quad F(z) = u_0, \text{ for } z_j = g(z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_n).\]

Let us record this result, changing notation and replacing \(z\) by \(x\).

**Proposition 2.6.** Let \(\Omega\) be a neighborhood of \(x_0 \in \mathbb{R}^n\). Assume we have a \(C^k\) function

\[(2.53) \quad F : \Omega \to \mathbb{R}, \quad F(x_0) = u_0,\]

and assume

\[(2.54) \quad DF(x_0) \neq 0, \quad \text{i.e.,} \quad (\partial_1 F(x_0),\ldots,\partial_n F(x_0)) \neq 0.\]

Then there exists \(j \in \{1,\ldots,n\}\) such that one can solve \(F(x) = u_0\) for

\[(2.55) \quad x_j = g(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n),\]

with \((x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n) = x_0\), for a \(C^k\) function \(g\).

**Remark.** For \(F : \Omega \to \mathbb{R}\), it is common to denote \(DF(x)\) by \(\nabla F(x)\),

\[(2.56) \quad \nabla F(x) = (\partial_1 F(x),\ldots,\partial_n F(x)).\]
Here is an example to which Proposition 2.6 applies. Using the notation \((x, y) = (x_1, x_2)\), set
\[
F : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad F(x, y) = x^2 + y^2 - x.
\]
Then
\[
\nabla F(x, y) = (2x - 1, 2y),
\]
which vanishes if and only if \(x = 1/2, \ y = 0\). Hence Proposition 2.6 applies if and only if \((x_0, y_0) \neq (1/2, 0)\).

Let us give an example involving a real valued function on \(M(n, \mathbb{R})\), namely
\[
det : M(n, \mathbb{R}) \rightarrow \mathbb{R}.
\]
As indicated in Exercise 11 of §1 (the first exercise set), if \(\det X \neq 0\),
\[
D \det(X) Y = (\det X) \text{Tr}(X^{-1}Y),
\]
so
\[
det X \neq 0 \implies D \det(X) \neq 0.
\]
We deduce that, if
\[
X_0 \in M(n, \mathbb{R}), \quad \det X_0 = a \neq 0,
\]
then, writing
\[
X = (x_{jk})_{1 \leq j, k \leq n},
\]
there exist \(\mu, \nu \in \{1, \ldots, n\}\) such that the equation
\[
det X = a
\]
has a smooth solution of the form
\[
x_{\mu \nu} = g(x_{\alpha \beta} : (\alpha, \beta) \neq (\mu, \nu)),
\]
such that, if the argument of \(g\) consists of the matrix entries of \(X_0\) other than the \(\mu, \nu\) entry, then the left side of (2.65) is the \(\mu, \nu\) entry of \(X_0\).

Let us return to the setting of Theorem 2.5, with \(\ell\) not necessarily equal to 1. In notation parallel to that of (2.51), we assume \(F\) is a \(C^k\) map,
\[
F : \Omega \rightarrow \mathbb{R}^\ell, \quad F(z_0) = u_0,
\]
where \( \Omega \) is a neighborhood of \( z_0 \) in \( \mathbb{R}^n \). We assume

\[
(2.67) \quad DF(z_0) : \mathbb{R}^n \rightarrow \mathbb{R}^\ell \text{ is surjective.}
\]

Then, upon reordering the variables \( z = (z_1, \ldots, z_n) \), we can write \( z = (x, y) \), \( x = (x_1, \ldots, x_{n-\ell}) \), \( y = (y_1, \ldots, y_\ell) \), such that \( D_y F(z_0) \) is invertible, and Theorem 2.5 applies. Thus (for this reordering of variables), we have a \( C^k \) solution to

\[
(2.68) \quad F(x, y) = u_0, \quad y = g(x, u_0),
\]

satisfying \( y_0 = g(x_0, u_0) \), \( z_0 = (x_0, y_0) \).

To give one example to which this result applies, we take another look at \( F : \mathbb{R}^4 \rightarrow \mathbb{R}^2 \) in (2.42). We have

\[
(2.69) \quad DF(u, v, x, y) = \begin{pmatrix} 2xu & 2xv & u^2 + v^2 & 0 \\ x & y & u & v \end{pmatrix}.
\]

The reader is invited to determine for which \((u, v, x, y) \in \mathbb{R}^4\) the matrix on the right side of (2.69) has rank 2.

Here is another example, involving a map defined on \( M(n, \mathbb{R}) \). Set

\[
(2.70) \quad F : M(n, \mathbb{R}) \rightarrow \mathbb{R}^2, \quad F(X) = \begin{pmatrix} \det X \\ \text{Tr} X \end{pmatrix}.
\]

Parallel to (2.60), if \( \det X \neq 0 \), \( Y \in M(n, \mathbb{R}) \),

\[
(2.71) \quad DF(X)Y = \begin{pmatrix} (\det X) \text{Tr}(X^{-1}Y) \\ \text{Tr} Y \end{pmatrix}.
\]

Hence, given \( \det X \neq 0 \), \( DF(X) : M(n, \mathbb{R}) \rightarrow \mathbb{R}^2 \) is surjective if and only if

\[
(2.72) \quad L : M(n, \mathbb{R}) \rightarrow \mathbb{R}^2, \quad LY = \begin{pmatrix} \text{Tr}(X^{-1}Y) \\ \text{Tr} Y \end{pmatrix}
\]

is surjective. This is seen to be the case if and only if \( X \) is not a scalar multiple of the identity \( I \in M(n, \mathbb{R}) \).

Exercises

1. Suppose \( F : U \rightarrow \mathbb{R}^n \) is a \( C^2 \) map, \( p \in U \), open in \( \mathbb{R}^n \), and \( DF(p) \) is invertible. With \( q = F(p) \), define a map \( N \) on a neighborhood of \( p \) by

\[
(2.73) \quad N(x) = x + DF(x)^{-1}(q - F(x)).
\]
Show that there exists $\varepsilon > 0$ and $C < \infty$ such that, for $0 \leq r < \varepsilon$,

$$\|x - p\| \leq r \implies \|N(x) - p\| \leq C \cdot r^2.$$ 

Conclude that, if $\|x_1 - p\| \leq r$ with $r < \min(\varepsilon, 1/2C)$, then $x_{j+1} = N(x_j)$ defines a sequence converging very rapidly to $p$. This is the basis of Newton's method, for solving $F(p) = q$ for $p$.

**Hint.** Apply $F$ to both sides of (2.73).

2. Applying Newton’s method to $f(x) = 1/x$, show that you get a fast approximation to division using only addition and multiplication.

**Hint.** Carry out the calculation of $N(x)$ in this case and notice a “miracle.”

3. Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ via $z = x + iy$, as in Exercise 4 of §1. Let $U \subset \mathbb{R}^{2n}$ be open, $F : U \to \mathbb{R}^{2n}$ be $C^1$. Assume $p \in U$, $DF(p)$ invertible. If $F^{-1} : V \to U$ is given as in Theorem 2.1, show that $F^{-1}$ is holomorphic provided $F$ is.

4. Let $O \subset \mathbb{R}^n$ be open. We say a function $f \in C^\infty(O)$ is real analytic provided that, for each $x_0 \in O$, we have a convergent power series expansion

$$f(x) = \sum_{\alpha \geq 0} \frac{1}{\alpha!} f^{(\alpha)}(x_0)(x - x_0)^\alpha,$$

valid in a neighborhood of $x_0$. Show that we can let $x$ be complex in (2.16), and obtain an extension of $f$ to a neighborhood of $O$ in $\mathbb{C}^n$. Show that the extended function is holomorphic, i.e., satisfies the Cauchy-Riemann equations.

**Hint.** Use Proposition 1.10.

**Remark.** It can be shown that, conversely, any holomorphic function has a power series expansion. See §10. For the next exercise, assume this as known.

5. Let $O \subset \mathbb{R}^n$ be open, $p \in O$, $f : O \to \mathbb{R}^n$ be real analytic, with $Df(p)$ invertible. Take $f^{-1} : V \to U$ as in Theorem 2.1. Show $f^{-1}$ is real analytic.

**Hint.** Consider a holomorphic extension $F : \Omega \to \mathbb{C}^n$ of $f$ and apply Exercise 3.

6. Use (2.10) to show that if a $C^1$ diffeomorphism has a $C^1$ inverse $G$, and if actually $F$ is $C^k$, then also $G$ is $C^k$.

**Hint.** Use induction on $k$. Write (2.10) as

$$G(x) = \Phi \circ F \circ G(x),$$

with $\Phi(X) = X^{-1}$, as on Exercises 3 and 13 of §1, $G(x) = DG(x), F(x) = DF(x)$. Apply Exercise 12 of §1 to show that, in general

$$G, F, \Phi \in C^k \implies G \in C^k.$$
Deduce that if one is given $F \in C^k$ and one knows that $G \in C^{k-1}$, then this result applies to give $G = DG \in C^{k-1}$, hence $G \in C^k$.

7. Show that there is a neighborhood $O$ of $(1,0) \in \mathbb{R}^2$ and there are functions $u, v, w \in C^1(O)$ $(u = u(x,y)$, etc.) satisfying the equations

$$u^3 + v^3 - xw^3 = 0,$$
$$u^2 + yw^2 + v = 1,$$
$$xu + yvw = 1,$$

for $(x,y) \in O$, and satisfying

$$u(1,0) = 1, \quad v(1,0) = 0, \quad w(1,0) = 1.$$

**Hint.** Define $F : \mathbb{R}^5 \to \mathbb{R}^3$ by

$$F(u,v,w,x,y) = \begin{pmatrix} u^3 + v^3 - xw^3 \\ u^2 + yw^2 + v \\ xu + yvw \end{pmatrix},$$

Then $F(1,0,1,0,0) = (0,1,1)^t$. Evaluate the $3 \times 3$ matrix $D_{u,v,w}F(1,0,1,1,0)$. Compare (2.42)–(2.47).

8. Consider $F : M(n, \mathbb{R}) \to M(n, \mathbb{R})$, given by $F(X) = X^2$. Show that $F$ is a diffeomorphism of a neighborhood of the identity matrix $I$ onto a neighborhood of $I$. Show that $F$ is not a diffeomorphism of a neighborhood of

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

onto a neighborhood of $I$ (in case $n = 2$).

9. Prove Corollary 2.4.

10. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be a $C^1$ map. Assume

$$\frac{\partial f}{\partial x}(0) \times \frac{\partial f}{\partial y}(0) = (0,0,1).$$

Show that there exist neighborhoods $O$ and $\Omega$ of $0 \in \mathbb{R}^2$ and a $C^1$ map $u : \Omega \to \mathbb{R}$ such that the image of $O$ under $f$ in $\mathbb{R}^3$ is the graph of $u$ over $\Omega$.

**Hint.** Let $\Pi : \mathbb{R}^3 \to \mathbb{R}^2$ be $\Pi(x,y,z) = (x,y)$, and consider

$$\varphi(x,y) = \Pi(f(x,y)), \quad \varphi : \mathbb{R}^2 \to \mathbb{R}^2.$$
Show that \( D\varphi(0) : \mathbb{R}^2 \to \mathbb{R}^2 \) is invertible, and apply the inverse function theorem. Then let \( u \) be the \( z \)-component of \( f \circ \varphi^{-1} \).

11. Generalize Exercise 10 to the setting where \( f : \mathbb{R}^m \to \mathbb{R}^n \) (\( m < n \)) is \( C^1 \) and
\[
Df(0) : \mathbb{R}^m \to \mathbb{R}^n
\]
is injective.

Remark. For related results, see the opening paragraphs of §5.

12. Let \( \Omega \subset \mathbb{R}^n \) be open and contain \( p_0 \). Assume \( F : \overline{\Omega} \to \mathbb{R}^n \) is continuous and \( F(p_0) = q_0 \). Assume \( F \) is \( C^1 \) on \( \Omega \) and \( DF(x) \) is invertible for all \( x \in \Omega \). Finally, assume there exists \( R > 0 \) such that
\[
(2.75) \quad x \in \partial \Omega \implies \| F(x) - q_0 \| \geq R.
\]
Show that
\[
(2.76) \quad F(\Omega) \supset B_{R/2}(q_0).
\]

Hint. Given \( y_0 \in B_{R/2}(q_0) \), use compactness to show that there exists \( x_0 \in \overline{\Omega} \) such that
\[
\| F(x_0) - y_0 \| = \inf_{x \in \overline{\Omega}} \| F(x) - y_0 \|.
\]
Use the hypothesis (2.75) to show that \( x_0 \in \Omega \). If \( F(x_0) \neq y_0 \), use
\[
F(x_0 + tz) = F(x_0) + tDF(x_0)z + o(\|tz\|),
\]
to produce \( z \in \mathbb{R}^n \) (say \( DF(x_0)z = y_0 - F(x_0) \)) such that \( F(x_0 + tz) \) is closer to \( y_0 \) than \( F(x_0) \) is, for small \( t > 0 \). Contradiction.

13. Do Exercise 12 with the conclusion (2.76) strengthened to
\[
(2.77) \quad F(\Omega) \supset B_R(q_0).
\]

Hint. It suffices to show that \( F(\Omega) \supset B_S(q_0) \) for each \( S < R \). Given such \( S \), produce a diffeomorphism \( \varphi : \mathbb{R}^n \to \mathbb{R}^n \) such that Exercise 12 applies to \( \varphi \circ F \), and yields the desired conclusion.
3. Systems of differential equations

In this section we study $n \times n$ systems of ODE,

$$\frac{dy}{dt} = F(t, y), \quad y(t_0) = y_0.$$  \hfill (3.1)

To begin, we prove the following fundamental existence and uniqueness result.

**Theorem 3.1.** Let $y_0 \in \Omega$, an open subset of $\mathbb{R}^n$, $I \subset \mathbb{R}$ an interval containing $t_0$. Suppose $F$ is continuous on $I \times \Omega$ and satisfies the following Lipschitz estimate in $y$:

$$\|F(t, y_1) - F(t, y_2)\| \leq L\|y_1 - y_2\|$$  \hfill (3.2)

for $t \in I$, $y_j \in \Omega$. Then the equation (3.1) has a unique solution on some $t$-interval containing $t_0$.

To begin the proof, we note that the equation (3.1) is equivalent to the integral equation

$$y(t) = y_0 + \int_{t_0}^{t} F(s, y(s)) \, ds.$$  \hfill (3.3)

Existence will be established via the Picard iteration method, which is the following. Guess $y_0(t)$, e.g., $y_0(t) = y_0$. Then set

$$y_k(t) = y_0 + \int_{t_0}^{t} F(s, y_{k-1}(s)) \, ds.$$  \hfill (3.4)

We aim to show that, as $k \to \infty$, $y_k(t)$ converges to a (unique) solution of (3.3), at least for $t$ close enough to $t_0$.

To do this, we use the Contraction Mapping Principle, established in §2. We look for a fixed point of $\Phi$, defined by

$$(\Phi y)(t) = y_0 + \int_{t_0}^{t} F(s, y(s)) \, ds.$$  \hfill (3.5)

Let

$$X = \{ u \in C(J, \mathbb{R}^n) : u(t_0) = y_0, \sup_{t \in J} \|u(t) - y_0\| \leq R \}.$$  \hfill (3.6)

Here $J = [t_0 - T, t_0 + T]$, where $T$ will be chosen, sufficiently small, below. The quantity $R$ is picked so that

$$\overline{B_R}(y_0) = \{ y : \|y - y_0\| \leq R \}$$
is contained in $\Omega$, and we also suppose $J \subset I$. Then there exists $M$ such that
\[
(3.7) \quad \sup_{s \in J, \|y - y_0\| \leq R} \|F(s, y)\| \leq M.
\]
Then, provided
\[
(3.8) \quad T \leq \frac{R}{M},
\]
we have
\[
(3.9) \quad \Phi : X \to X.
\]
Now, using the Lipschitz hypothesis (3.2), we have, for $t \in J$,
\[
(3.10) \quad \|(\Phi y)(t) - (\Phi z)(t)\| \leq \int_{t_0}^{t} L\|y(s) - z(s)\| \, ds \\
\quad \leq TL \sup_{s \in J} \|y(s) - z(s)\|
\]
assuming $y$ and $z$ belong to $X$. It follows that $\Phi$ is a contraction on $X$ provided one has
\[
(3.11) \quad T < \frac{1}{L}
\]
in addition to the hypotheses above. This proves Theorem 3.1.

Note that the bound $M$ and the Lipschitz hypothesis on $F$ were needed only on $\overline{B_R(y_0)}$. Thus we can extend Theorem 3.1 to the following setting:
\[
(3.12) \quad \text{For each compact } K \subset \Omega, \text{ there exists } M_K < \infty \text{ such that} \\
\quad \|F(t, x)\| \leq M_K, \; \forall x \in K, \; t \in I,
\]
and
\[
(3.13) \quad \text{For each } K \text{ as above, there exists } L_K < \infty \text{ such that} \\
\quad \|F(t, x) - F(t, y)\| \leq L_K\|x - y\|, \; \forall x, y \in K, \; t \in I.
\]
Note that, if $K \subset \Omega$ is compact, there exists $R_K > 0$ such that
\[
\tilde{K} = \{ x \in \mathbb{R}^n : \text{dist}(x, K) \leq R_K \} \subset \Omega,
\]
and $\tilde{K}$ is compact. It follows that for each $y_0 \in K$, the solution to (3.1) exists on the interval
\[
(3.14) \quad \{ t \in I : |t - t_0| \leq \min(R_K/M_{\tilde{K}}, 1/2L_{\tilde{K}}) \}.
\]

Now that we have local solutions to (3.1), it is of interest to investigate when global solutions exist. Here is an example where breakdown occurs:
\[
(3.15) \quad \frac{dy}{dt} = y^2, \quad y(0) = 1.
\]
The solution blows up in finite time. See Exercise 1. It is useful to know that “blowing up” is the only way a solution can fail to exist globally. We have the following result.
Proposition 3.2. Let $F$ be as in Theorem 3.1, but with the boundedness and Lipschitz hypotheses replaced by (3.12)–(3.13). Assume $[a, b]$ is contained in the open interval $I$, and assume $y(t)$ solves (3.1) for $t \in (a, b)$. Assume there exists a compact $K \subset \Omega$ such that $y(t) \in K$ for all $t \in (a, b)$. Then there exist $a_1 < a$ and $b_1 > b$ such that $y(t)$ solves (3.1) for $t \in (a_1, b_1)$.

Proof. We deduce from (3.14) that there exists $\delta > 0$ such that for each $y_1 \in K$, $t_1 \in [a, b]$, the solution to

$$\frac{dy}{dt} = F(t, y), \quad y(t_1) = y_1 \tag{3.16}$$

exists on the interval $[t_1 - \delta, t_1 + \delta]$. Now, under the current hypotheses, take $t_1 \in (b - \delta/2, b)$ and $y_1 = y(t_1)$, with $y(t)$ solving (3.1). Then solving (3.16) continues $y(t)$ past $t = b$. Similarly one can continue $y(t)$ past $t = a$.

Here is an example of a global existence result that can be deduced from Proposition 3.2. Consider the $2 \times 2$ system for $y = (x, v)$:

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -x^3. \tag{3.17}$$

Here we take $\Omega = \mathbb{R}^2$, $F(t, y) = F(t, x, v) = (u, -x^3)$. If (3.17) holds for $t \in (a, b)$, we have

$$\frac{d}{dt} \left( \frac{v^2}{2} + \frac{x^4}{4} \right) = v \frac{dv}{dt} + x^3 \frac{dx}{dt} = 0, \tag{3.18}$$

so each $y(t) = (x(t), v(t))$ solving (3.17) lies on a level curve $x^4/4 + v^2/2 = C$, hence is confined to a compact subset of $\mathbb{R}^2$, yielding global existence of solutions to (3.17).

For more examples of global existence, see Exercises 2–4 below, and also section 3b below, treating linear systems.

The discussion above dealt with first order systems. Often one wants to deal with a higher-order ODE. There is a standard method of reducing an $n$th-order ODE

$$y^{(n)}(t) = f(t, y, y', \ldots, y^{(n-1)}) \tag{3.19}$$

to a first-order system. One sets $u = (u_0, \ldots, u_{n-1})$ with

$$u_0 = y, \quad u_j = y^{(j)}, \tag{3.20}$$

and then

$$\frac{du}{dt} = (u_1, \ldots, u_{n-1}, f(t, u_0, \ldots, u_{n-1})) = g(t, u). \tag{3.21}$$
If $y$ takes values in $\mathbb{R}^k$, then $u$ takes values in $\mathbb{R}^{kn}$.

If the system (3.1) is non-autonomous, i.e., if $F$ explicitly depends on $t$, it can be converted to an autonomous system (one with no explicit $t$-dependence) as follows. Set $z = (t, y)$. We then have

$$
\frac{dz}{dt} = \left(1, \frac{dy}{dt}\right) = (1, F(z)) = G(z).
$$

Sometimes this process destroys important features of the original system (3.1). For example, if (3.1) is linear, (3.22) might be nonlinear. Nevertheless, the trick of converting (3.1) to (3.22) has some uses.

### 3b. Linear systems

Here we consider linear systems, of the form

$$
\frac{dx}{dt} = A(t)x, \quad x(0) = x_0,
$$

given $A(t)$ continuous in $t \in I$ (an interval about 0), with values in $M(n, \mathbb{R})$. We will apply Proposition 3.2 to establish global existence of solutions. It suffices to establish the following.

**Proposition 3.3.** If $\| A(t) \| \leq K$ for $t \in I$, then the solution to (3.23) satisfies

$$
\| x(t) \| \leq e^{K|t|} \| x_0 \|.
$$

**Proof.** It suffices to prove (3.24) for $t \geq 0$. Then $y(t) = e^{-Kt}x(t)$ satisfies

$$
\frac{dy}{dt} = C(t)y, \quad y(0) = x_0, \quad C(t) = A(t) - KI.
$$

We claim that, for $t \geq 0$,

$$
\| y(t) \| \leq \| y(0) \|,
$$

which then implies (3.24), for $t \geq 0$. In fact,

$$
\frac{d}{dt} \| y(t) \|^2 = y'(t) \cdot y(t) + y(t) \cdot y'(t)
$$

$$
= 2y(t) \cdot (A(t) - K)y(t).
$$

Now

$$
y(t) \cdot A(t)y(t) \leq \| y(t) \| \cdot \| A(t)y(t) \| \leq \| A(t) \| \cdot \| y(t) \|^2,
$$
so the hypothesis \( \|A(t)\| \leq K \) implies

\[
\frac{d}{dt} \|y(t)\|^2 \leq 0.
\]

yielding (3.26).

Thanks to Proposition 3.3, we have, for \( s, t \in I \), the solution operator for (3.23),

\[
(3.29) \quad S(t, s) \in M(n, \mathbb{R}), \quad S(t, s)x(s) = x(t).
\]

We have

\[
(3.30) \quad \frac{\partial}{\partial t} S(t, s) = A(t)S(t, s), \quad S(s, s) = I.
\]

Note that \( S(t, s)S(s, r) = S(t, r) \). In particular, \( S(t, s) = S(s, t)^{-1} \).

We can use the solution operator \( S(t, s) \) to solve the inhomogeneous system

\[
(3.31) \quad \frac{dx}{dt} = A(t)x + f(t), \quad x(t_0) = x_0.
\]

Namely, we can take

\[
(3.32) \quad x(t) = S(t, t_0)x_0 + \int_{t_0}^{t} S(t, s)f(s) \, ds.
\]

This is known as Duhamel’s formula. Verifying that this solves (3.30) is an exercise. We will make good use of this in the next subsection.

3c. Dependence of solutions on initial data and other parameters

We study how the solution to a system of differential equations

\[
(3.33) \quad \frac{dx}{dt} = F(x), \quad x(0) = y
\]

depends on the initial condition \( y \). As shown in (3.22), there is no loss of generality in considering the autonomous system (3.33). We will assume \( F : \Omega \to \mathbb{R}^n \) is smooth, \( \Omega \subset \mathbb{R}^n \) open and convex, and denote the solution to (3.33) by \( x = x(t, y) \). We want to examine smoothness in \( y \). Let \( DF(x) \) denote the \( n \times n \) matrix valued function of partial derivatives of \( F \).

To start, we assume \( F \) is of class \( C^1 \), i.e., \( DF \) is continuous on \( \Omega \), and we want to show \( x(t, y) \) is differentiable in \( y \). Let us recall what this means. Take \( y \in \Omega \) and pick
$R > 0$ such that $\overline{B_R(y)}$ is contained in $\Omega$. We seek an $n \times n$ matrix $W(t, y)$ such that, for $w_0 \in \mathbb{R}^n$, $\|w_0\| \leq R$,

\begin{equation}
(3.34) \quad x(t, y + w_0) = x(t, y) + W(t, y)w_0 + r(t, y, w_0),
\end{equation}

where

\begin{equation}
(3.35) \quad r(t, y, w_0) = o(\|w_0\|),
\end{equation}

which means

\begin{equation}
(3.36) \quad \lim_{w_0 \to 0} \frac{r(t, y, w_0)}{\|w_0\|} = 0.
\end{equation}

When this holds, $x(t, y)$ is differentiable in $y$, and

\begin{equation}
(3.37) \quad D_y x(t, y) = W(t, y).
\end{equation}

In other words,

\begin{equation}
(3.38) \quad x(t, y + w_0) = x(t, y) + D_y x(t, y)w_0 + o(\|w_0\|).
\end{equation}

In the course of proving this differentiability, we also want to produce an equation for $W(t, y) = D_y x(t, y)$. This can be done as follows. Suppose $x(t, y)$ were differentiable in $y$. (We do not yet know that it is, but that is okay.) Then $F(x(t, y))$ is differentiable in $y$, so we can apply $D_y$ to (3.32). Using the chain rule, we get the following equation,

\begin{equation}
(3.39) \quad \frac{dW}{dt} = DF(x)W, \quad W(0, y) = I,
\end{equation}

called the linearization of (3.33). Here, $I$ is the $n \times n$ identity matrix. Equivalently, given $w_0 \in \mathbb{R}^n$,

\begin{equation}
(3.40) \quad w(t, y) = W(t, y)w_0
\end{equation}

is expected to solve

\begin{equation}
(3.41) \quad \frac{dw}{dt} = DF(x)w, \quad w(0) = w_0.
\end{equation}

Now, we do not yet know that $x(t, y)$ is differentiable, but we do know from results of §3b that (3.39) and (3.41) are uniquely solvable. It remains to show that, with such a choice of $W(t, y)$, (3.34)–(3.35) hold.

To rephrase the task, set

\begin{equation}
(3.42) \quad x(t) = x(t, y), \quad x_1(t) = x(t, y + w_0), \quad z(t) = x_1(t) - x(t),
\end{equation}
and let \( w(t) \) solve (3.41). The task of verifying (3.34)–(3.35) is equivalent to the task of verifying

\[
\|z(t) - w(t)\| = o(\|w_0\|). \tag{3.43}
\]

To show this, we will obtain for \( z(t) \) an equation similar to (3.41). To begin, (3.42) implies

\[
\frac{dz}{dt} = F(x_1) - F(x), \quad z(0) = w_0. \tag{3.44}
\]

Now the fundamental theorem of calculus gives

\[
F(x_1) - F(x) = G(x_1, x)(x_1 - x), \tag{3.45}
\]

with

\[
G(x_1, x) = \int_0^1 DF(\tau x_1 + (1 - \tau)x) \ d\tau. \tag{3.46}
\]

If \( F \) is \( C^1 \), then \( G \) is continuous. Then (3.44)–(3.45) yield

\[
\frac{dz}{dt} = G(x_1, x)z, \quad z(0) = w_0. \tag{3.47}
\]

Given that

\[
\|DF(u)\| \leq L, \quad \forall u \in \Omega, \tag{3.48}
\]

which we have by continuity of \( DF \), after possibly shrinking \( \Omega \) slightly, we deduce from Proposition 3.3 that

\[
\|z(t)\| \leq e^{\|t\|L}\|w_0\|, \tag{3.49}
\]

that is,

\[
\|x(t, y) - x(t, y + w_0)\| \leq e^{\|t\|L}\|w_0\|. \tag{3.50}
\]

This establishes that \( x(t, y) \) is Lipschitz in \( y \).

To proceed, since \( G \) is continuous and \( G(x, x) = DF(x) \), we can rewrite (3.47) as

\[
\frac{dz}{dt} = G(x + z, x)z + R(x, z), \quad z(0) = w_0, \tag{3.51}
\]

where

\[
F \in C^1(\Omega) \implies \|R(x, z)\| = o(\|z\|) = o(\|w_0\|). \tag{3.52}
\]
Now comparing (3.51) with (3.41), we have

\[(3.53) \quad \frac{d}{dt}(z - w) = DF(x)(z - w) + R(x, z), \quad (z - w)(0) = 0.\]

Then Duhamel’s formula gives

\[(3.54) \quad z(t) - w(t) = \int_0^t S(t, s)R(x(s), z(s)) \, ds,\]

where \(S(t, s)\) is the solution operator for \(d/dt - B(t)\), with \(B(t) = G(x_1(t), x(t))\), which as in (3.49), satisfies

\[(3.55) \quad \|S(t, s)\| \leq e^{(t-s)\|L\|.}\]

We hence have (3.43), i.e.,

\[(3.56) \quad \|z(t) - w(t)\| = o(\|w_0\|).\]

This is precisely what is required to show that \(x(t, y)\) is differentiable with respect to \(y\), with derivative \(W = D_y x(t, y)\) satisfying (3.39). Hence we have:

**Proposition 3.4.** If \(F \in C^1(\Omega)\) and if solutions to (3.33) exist for \(t \in (-T_0, T_1)\), then, for each such \(t\), \(x(t, y)\) is \(C^1\) in \(y\), with derivative \(D_y x(t, y)\) satisfying (3.39).

We have shown that \(x(t, y)\) is both Lipschitz and differentiable in \(y\). The continuity of \(W(t, y)\) in \(y\) follows easily by comparing the differential equations of the form (3.39) for \(W(t, y)\) and \(W(t, y + w_0)\), in the spirit of the analysis of \(z(t)\) done above.

If \(F\) possesses further smoothness, we can establish higher differentiability of \(x(t, y)\) in \(y\) by the following trick. Couple (3.33) and (3.39), to get a system of differential equations for \((x, W)\):

\[(3.57) \quad \frac{dx}{dt} = F(x), \quad \frac{dW}{dt} = DF(x)W,\]

with initial conditions

\[(3.58) \quad x(0) = y, \quad W(0) = I.\]

We can reiterate the preceding argument, getting results on \(D_y (x, W)\), hence on \(D_y^2 x(t, y)\), and continue, proving:
Proposition 3.5. If $F \in C^k(\Omega)$, then $x(t, y)$ is $C^k$ in $y$.

Similarly, we can consider dependence of the solution to

\begin{equation}
\frac{dx}{dt} = F(\tau, x), \quad x(0) = y
\end{equation}

on a parameter $\tau$, assuming $F$ smooth jointly in $(\tau, x)$. This result can be deduced from the previous one by the following trick. Consider the system

\begin{equation}
\frac{dx}{dt} = F(z, y), \quad \frac{dz}{dt} = 0, \quad x(0) = y, \quad z(0) = \tau.
\end{equation}

Then we get smoothness of $x(t, \tau, y)$ jointly in $(\tau, y)$. As a special case, let $F(\tau, x) = \tau F(x)$. In this case $x(t_0, \tau, y) = x(\tau t_0, y)$, so we can improve the conclusion in Proposition 3.5 to the following:

\begin{equation}
F \in C^k(\Omega) \implies x \in C^k \text{ jointly in } (t, y).
\end{equation}

3d. Vector fields and flows

Let $U \subset \mathbb{R}^n$ be open. A vector field on $U$ is a smooth map

\begin{equation}
X : U \to \mathbb{R}^n.
\end{equation}

Consider the corresponding ODE

\begin{equation}
\frac{dy}{dt} = X(y), \quad y(0) = x,
\end{equation}

with $x \in U$. A curve $y(t)$ solving (3.63) is called an integral curve of the vector field $X$. It is also called an orbit. For fixed $t$, write

\begin{equation}
y = y(t, x) = \mathcal{F}_X^t(x).
\end{equation}

The locally defined $\mathcal{F}_X^t$, mapping (a subdomain of) $U$ to $U$, is called the flow generated by the vector field $X$. As a consequence of the results of §3c, in (3.64), $y$ is a smooth function of $(t, x)$.

The vector field $X$ defines a differential operator on scalar functions, as follows:

\begin{equation}
\mathcal{L}_X f(x) = \lim_{h \to 0} h^{-1} \left[ f(\mathcal{F}_X^h x) - f(x) \right] = \frac{d}{dt} f(\mathcal{F}_X^t x) \bigg|_{t=0}.
\end{equation}

We also use the common notation

\begin{equation}
\mathcal{L}_X f(x) = X f,
\end{equation}
that is, we apply $X$ to $f$ as a first order differential operator.

Note that, if we apply the chain rule to (3.65) and use (3.63), we have

$$(3.67) \quad \mathcal{L}_X f(x) = X(x) \cdot \nabla f(x) = \sum a_j(x) \frac{\partial f}{\partial x_j},$$

if $X = \sum a_j(x)e_j$, with $\{e_j\}$ the standard basis of $\mathbb{R}^n$. In particular, using the notation (3.66), we have

$$(3.68) \quad a_j(x) = Xx_j.$$  

In the notation (3.66),

$$(3.69) \quad X = \sum a_j(x) \frac{\partial}{\partial x_j}.$$  

We note that $X$ is a derivation, that is, a map on $C^\infty(U)$, linear over $\mathbb{R}$, satisfying

$$(3.70) \quad X(fg) = (Xf)g + f(Xg).$$

Conversely, any derivation on $C^\infty(U)$ defines a vector field, i.e., has the form (3.69), as we now show.

**Proposition 3.6.** If $X$ is a derivation on $C^\infty(U)$, then $X$ has the form (3.69).

**Proof.** Set $a_j(x) = Xx_j$, $X^\# = \sum a_j(x)\partial/\partial x_j$, and $Y = X - X^\#$. Then $Y$ is a derivation satisfying $Yx_j = 0$ for each $j$. We aim to show that $Yf = 0$ for all $f$. Note that whenever $Y$ is a derivation

$$1 \cdot 1 = 1 \Rightarrow Y \cdot 1 = 2Y \cdot 1 \Rightarrow Y \cdot 1 = 0.$$  

Thus $Y$ annihilates constants. Thus in this case $Y$ annihilates all polynomials of degree $\leq 1$.

Now we show that $Yf(p) = 0$ for all $p \in U$. Without loss of generality, we can suppose $p = 0$. Then, with $b_j(x) = \int_0^1 (\partial_j f)(tx) \, dt$, we can write

$$f(x) = f(0) + \sum b_j(x)x_j.$$  

It immediately follows that $Yf$ vanishes at 0, so the proposition is proved.

A fundamental fact about vector fields is that they can be “straightened out” near points where they do not vanish. To see this, let $X$ be a smooth vector field on $U$, and suppose $X(p) \neq 0$. Then near $p$ there is a hyperplane $H$ that is not tangent to $X$ near $p$, say on a portion we denote $M$. We can choose coordinates near $p$ so that $p$ is the origin and $M$ is given by $\{x_n = 0\}$. Thus we can identify a point $x' \in \mathbb{R}^{n-1}$ near the origin with $x' \in M$. We can define a map

$$(3.71) \quad F : M \times (-t_0, t_0) \rightarrow U.$$
by

\[ (3.72) \quad F(x', t) = F_X^t(x'). \]

This is \( C^\infty \) and has surjective derivative at \((0, 0)\), and so by the inverse function theorem is a local diffeomorphism. This defines a new coordinate system near \( p \), in which the flow generated by \( X \) has the form

\[ (3.73) \quad F_X^s(x', t) = (x', t + s). \]

If we denote the new coordinates by \((u_1, \ldots, u_n)\), we see that the following result is established.

**Theorem 3.7.** If \( X \) is a smooth vector field on \( U \) and \( X(p) \neq 0 \), then there exists a coordinate system \((u_1, \ldots, u_n)\), centered at \( p \) (so \( u_j(p) = 0 \)) with respect to which

\[ (3.74) \quad X = \frac{\partial}{\partial u_n}. \]

We consider further mapping properties of vector fields. If \( F : V \to W \) is a diffeomorphism between two open domains in \( \mathbb{R}^n \), and \( Y \) is a vector field on \( W \), we define a vector field \( F_\# Y \) on \( V \) so that

\[ (3.75) \quad F^t \cdot F_\# Y = F^{-1} \circ F_Y^t \circ F, \]

or equivalently, by the chain rule,

\[ (3.76) \quad F_\# Y(x) = (DF^{-1})(F(x))Y(F(x)). \]

In particular, if \( U \subset \mathbb{R}^n \) is open and \( X \) is a vector field on \( U \), defining a flow \( F^t \), then for a vector field \( Y \), \( F_\#^t Y \) is defined on most of \( U \), for \(|t| \) small, and we can define the Lie derivative:

\[ (3.77) \quad L_X Y = \lim_{h \to 0} h^{-1}(F_\#^h Y - Y) = \frac{d}{dt} F_\#^t Y|_{t=0}, \]

as a vector field on \( U \).

Another natural construction is the operator-theoretic bracket:

\[ (3.78) \quad [X, Y] = XY - YX, \]

where the vector fields \( X \) and \( Y \) are regarded as first order differential operators on \( C^\infty(U) \). One verifies that (3.78) defines a vector field on \( U \). In fact, if \( X = \sum a_j(x)\partial/\partial x_j \), \( Y = \sum b_j(x)\partial/\partial x_j \), then

\[ (3.79) \quad [X, Y] = \sum_{j,k} \left( a_k \frac{\partial b_j}{\partial x_k} - b_k \frac{\partial a_j}{\partial x_k} \right) \frac{\partial}{\partial x_j}. \]

The basic fact about the Lie bracket is the following.
Theorem 3.8. If $X$ and $Y$ are smooth vector fields, then

$$\mathcal{L}_X Y = [X, Y].$$

Proof. We examine $\mathcal{L}_X Y = (d/ds)\mathcal{F}_X^s Y|_{s=0}$, using (3.76), which implies that

$$Y_s(x) = \mathcal{F}_X^s Y(x) = D\mathcal{F}_X^{-s}(\mathcal{F}_X^s(x))Y(\mathcal{F}_X^s(x)).$$

Let us set $G^s = D\mathcal{F}_X^{-s}$. Note that $G^s : U \to \text{End}(\mathbb{R}^n)$. Hence, for $x \in U$, $DG^s(x)$ is an element of $\text{Hom}(\mathbb{R}^n, \text{End}(\mathbb{R}^n))$. Taking the $s$-derivative of (3.81), we have

$$\frac{d}{ds} Y_s(x) = -DX(\mathcal{F}_X^s(x))Y(\mathcal{F}_X^s(x)) + DG^s(\mathcal{F}_X^s(x))X(\mathcal{F}_X^s(x))Y(\mathcal{F}_X^s(x)) + D\mathcal{F}_X^{-s}(\mathcal{F}_X^s(x))DY(\mathcal{F}_X^s(x))X(\mathcal{F}_X^s(x)).$$

Note that $G^0(x) = I \in \text{End}(\mathbb{R}^n)$ for all $x \in U$, so $DG^0 = 0$. Thus

$$\mathcal{L}_X Y = \frac{d}{ds} Y_s(x)|_{s=0} = -DX(x)Y(x) + DY(x)X(x),$$

which agrees with the formula (3.79) for $[X, Y]$.

Corollary 3.9. If $X$ and $Y$ are smooth vector fields on $U$, then

$$\frac{d}{dt} \mathcal{F}_X^t Y = \mathcal{F}_X^t [X, Y]$$

for all $t$.

Proof. Since locally $\mathcal{F}_X^{t+s} = \mathcal{F}_X^s \mathcal{F}_X^t$, we have the same identity for $\mathcal{F}_X^{t+s}$. Hence

$$\frac{d}{dt} \mathcal{F}_X^t Y = \frac{d}{ds} \mathcal{F}_X^s \mathcal{F}_X^t Y|_{s=0} = \mathcal{F}_X^t \mathcal{L}_X Y,$$

which yields (3.84).

Exercises

1. Solve the initial value problem

$$\frac{dy}{dt} = y^2, \quad y(0) = a,$$

given $a \in \mathbb{R}$. On what $t$-interval is the solution defined?
2. Assume in (3.1) that $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is $C^1$ and satisfies $\|F(t,y)\| \leq M$ for all $(t,y) \in \mathbb{R} \times \mathbb{R}^n$. Use Proposition 3.2 to show that (3.1) has a unique solution for all $t \in \mathbb{R}$.

3. Let $M$ be a compact smooth surface in $\mathbb{R}^n$. Suppose $F : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map (vector field), such that, for each $x \in M$, $F(x)$ is tangent to $M$, i.e., the line $\gamma_x(t) = x + tF(x)$ is tangent to $M$ at $x$, at $t = 0$. Show that, if $x \in M$, then the initial value problem

$$\frac{dy}{dt} = F(y), \quad y(0) = x$$

has a solution for all $t \in \mathbb{R}$, and $y(t) \in M$ for all $t$.

*Hint.* Locally, straighten out $M$ to be a linear subspace of $\mathbb{R}^n$, to which $F$ is tangent. Use uniqueness. Material in §2 helps do this local straightening.

4. Show that the initial value problem

$$\frac{dx}{dt} = -x(x^2 + y^2), \quad \frac{dy}{dt} = -y(x^2 + y^2), \quad x(0) = x_0, \quad y(0) = y_0$$

has a solution for all $t \geq 0$, but not for all $t < 0$, unless $(x_0, y_0) = (0, 0)$.

5. Verify Duhamel’s formula (3.32), for the solution to (3.31).

**Exercises on exponential functions**

1. Let $a \in \mathbb{R}$. Show that the unique solution to $u'(t) = au(t), \ u(0) = 1$ is given by

$$(3.86) \quad u(t) = \sum_{j=0}^{\infty} \frac{a^j}{j!} t^j.$$  

We denote this function by $u(t) = e^{at}$, the exponential function. We also write $\exp(t) = e^t$.

*Hint.* Integrate the series term by term and use the Fundamental Theorem of Calculus.

*Alternative.* Setting $u_0(t) = 1$, and using the Picard iteration method (3.4) to define the sequence $u_k(t)$, show that $u_k(t) = \sum_{j=0}^{k} a^j t^j / j!$.

2. Show that, for all $s, t \in \mathbb{R}$,

$$(3.87) \quad e^{a(s+t)} = e^{as} e^{at}.$$  

*Hint.* Show that $u_1(t) = e^{a(s+t)}$ and $u_2(t) = e^{as} e^{at}$ solve the same initial value problem.

*Alternative.* Apply $d/dt$ to $e^{a(s+t)} e^{-at}$.

3. Show that $\exp : \mathbb{R} \to (0, \infty)$ is a diffeomorphism. We denote the inverse by

$$\log : (0, \infty) \to \mathbb{R}.$$
Show that \(v(x) = \log x\) solves the ODE \(dv/dx = 1/x\), \(v(1) = 0\), and deduce that
\[
\int_1^x \frac{1}{y} \, dy = \log x.
\]

4. Let \(a \in \mathbb{C}\). Show that the unique solution to \(f'(t) = af\), \(f(0) = 1\) is given by
\[
f(t) = \sum_{j=0}^{\infty} \frac{a^j}{j!} t^j.
\]
We denote this function by \(f(t) = e^{at}\). Show that, for all \(t \in \mathbb{R}\), \(a, b \in \mathbb{C}\),
\[
e^{(a+b)t} = e^{at}e^{bt}.
\]

Hint. See the alternative hint for Problem 2.

5. Write
\[
e^{it} = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} t^{2j} + i \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} t^{2j+1} = u(t) + iv(t).
\]
Show that
\[
u'(t) = -v(t), \quad v'(t) = u(t).
\]
We denote these functions by \(u(t) = \cos t\), \(v(t) = \sin t\). The identity
\[
e^{it} = \cos t + i \sin t
\]
is called Euler’s formula. For a presentation of the standard geometric definition of \(\sin t\) and \(\cos t\), and its equivalence with (3.92), see exercises in \(\S5\).

Auxiliary exercises on trigonometric functions

We use the definitions of \(\sin t\) and \(\cos t\) given in Exercise 5 of the last exercise set.

1. Use (3.90) to derive the identities
\[
\sin(x + y) = \sin x \cos y + \cos x \sin y
\]
\[
\cos(x + y) = \cos x \cos y - \sin x \sin y.
\]

2. Use (3.90)–(3.93) to show that
\[
\sin^2 t + \cos^2 t = 1, \quad \cos^2 t = \frac{1}{2}(1 + \cos 2t).
\]
Hint. Show that $e^{it} = e^{-it}$ and hence $|e^{it}|^2 = 1$, for $t \in \mathbb{R}$.

3. Show that

\begin{equation}
\gamma(t) = (\cos t, \sin t)
\end{equation}

is a map of $\mathbb{R}$ onto the unit circle $S^1 \subset \mathbb{R}^2$ with non-vanishing derivative, and, as $t$ increases, $\gamma(t)$ moves monotonically, counterclockwise. Use Exercise 5 above to show that

\begin{equation}
\gamma'(t) = (-\sin t, \cos t),
\end{equation}

and deduce that $\gamma(t)$ has unit speed.

We define $\pi$ to be the smallest number $t_1 \in (0, \infty)$ such that $\gamma(t_1) = (-1, 0)$, so

$$\cos \pi = -1, \quad \sin \pi = 0.$$ 

Show that $2\pi$ is the smallest number $t_2 \in (0, \infty)$ such that $\gamma(t_2) = (1, 0)$, so

$$\cos 2\pi = 1, \quad \sin 2\pi = 0.$$ 

Show that

$$\cos(t + 2\pi) = \cos t, \quad \cos(t + \pi) = -\cos t$$

$$\sin(t + 2\pi) = \sin t, \quad \sin(t + \pi) = -\sin t.$$ 

Show that $\gamma(\pi/2) = (0, 1)$, and that

$$\cos\left(t + \frac{\pi}{2}\right) = -\sin t, \quad \sin\left(t + \frac{\pi}{2}\right) = \cos t.$$ 

4. Show that $\sin : (-\pi/2, \pi/2) \to (-1, 1)$ is a diffeomorphism. We denote its inverse by

$$\arcsin : (-1, 1) \mapsto \left(-\frac{1}{2}\pi, \frac{1}{2}\pi\right).$$

Show that $u(t) = \arcsin t$ solves the ODE

$$\frac{du}{dt} = \frac{1}{\sqrt{1-t^2}}, \quad u(0) = 0.$$ 

Hint. Apply the chain rule to $\sin(u(t)) = t$.

Deduce that, for $t \in (-1, 1),

\begin{equation}
\arcsin t = \int_0^t \frac{dx}{\sqrt{1-x^2}}.
\end{equation}
5. Show that
\[ e^{\pi i/3} = \frac{1}{2} + \frac{\sqrt{3}}{2} i, \quad e^{\pi i/6} = \frac{\sqrt{3}}{2} + \frac{1}{2} i. \]

*Hint.* First compute
\[ \left( \frac{1}{2} + \frac{\sqrt{3}}{2} i \right)^3 \]
and use Exercise 3. Then compute \( e^{\pi i/2} e^{-\pi i/3} \).

For intuition behind these formulas, look at Fig. 3.1.

6. Show that \( \sin(\pi/6) = 1/2 \), and hence that
\[ \frac{\pi}{6} = \int_0^{1/2} \frac{dx}{\sqrt{1 - x^2}} = \sum_{n=0}^{\infty} a_n \left( \frac{1}{2} \right)^{2n+1}, \]
where
\[ a_0 = 1, \quad a_{n+1} = \frac{2n + 1}{2n + 2} a_n. \]
Show that
\[ \frac{\pi}{6} - \sum_{n=0}^{k} a_n \left( \frac{1}{2} \right)^{2n+1} < \frac{4^{-k}}{3(2k+3)}. \]
Using a calculator, sum the series over \( 0 \leq n \leq 20 \), and verify that
\[ \pi \approx 3.141592653589 \ldots \]

7. For \( x \neq (k + 1/2)\pi, \ k \in \mathbb{Z} \), set
\[ \tan x = \frac{\sin x}{\cos x}. \]
Show that \( 1 + \tan^2 x = 1/\cos^2 x \). Show that \( w(x) = \tan x \) satisfies the ODE
\[ \frac{dw}{dx} = 1 + w^2, \quad w(0) = 0. \]

8. Show that \( \tan : (-\pi/2, \pi/2) \to \mathbb{R} \) is a diffeomorphism. Denote the inverse by
\[ \arctan : \mathbb{R} \to \left( -\frac{1}{2} \pi, \frac{1}{2} \pi \right). \]
Show that
\[ \arctan y = \int_0^y \frac{dx}{1 + x^2}. \]
9. Use the identity
\[ \tan \frac{\pi}{6} = \frac{1}{\sqrt{3}} \]
together with (3.98) to produce an infinite series for \( \pi \) that is different from that of Exercise 6.

**Exercises on the matrix exponential**

1. Let \( A \) be an \( n \times n \) matrix. Parallel to Exercises 1 and 3 of the exercise set on exponential functions, show that

(3.99) \[ e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \]
solves

(3.100) \[ \frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA}A, \quad e^{tA} |_{t=0} = I. \]

2. Show that, for \( t \in \mathbb{R}, \ A \in M(n, \mathbb{C}), \)

(3.101) \[ e^{tA}e^{-tA} = I. \]

*Hint.* Compute \((d/dt)e^{tA}e^{-tA}\). Show it is 0.

3. In the setting of Exercise 2, show that for \( s, t \in \mathbb{R}, \)

(3.102) \[ e^{(s+t)A} = e^{sA}e^{tA}. \]

*Hint.* Compute \((d/dt) e^{(s+t)A}e^{-tA}\). Show it is 0.

4. Let \( A, B \in M(n, \mathbb{C}) \), and assume

(3.103) \[ AB = BA. \]

Show that

(3.104) \[ e^{t(A+B)} = e^{tA}e^{tB}. \]

*Hint.* Compute

(3.105) \[ \frac{d}{dt} e^{t(A+B)}e^{-tB}e^{-tA}. \]
Show it is 0. To get this, you will want to show that, if $A$ and $B$ commute, then
\[ e^{-tB}A = Ae^{-tB}. \]

5. We desire to compute $e^{tJ}$ when
\[ J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \]
Noting that $J^2 = -I$, $J^3 = -J$, $J^4 = I$, show that the power series for $e^{tJ}$ resembles (3.91) and deduce that
\[ e^{tJ} = (\cos t)I + (\sin t)J = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \]
Note the resemblance of (3.107) and (3.92).

6. Note that $X(t) = \exp(tA) = e^{tA}$ is given as the solution to a system of the form (3.59):
\[ \frac{dX}{dt} = F(A, X), \quad X(0) = I, \]
where
\[ F(A, X) = AX. \]
Deduce from the discussion of (3.59)–(3.60) that
\[ \exp : M(n, \mathbb{C}) \rightarrow M(n, \mathbb{C}) \]
is smooth, of class $C^k$, for each $k \in \mathbb{N}$. Note that such smoothness also follows from Corollary 1.12.
4. The Riemann integral in $n$ variables

We define the Riemann integral of a bounded function $f : R \to \mathbb{R}$, where $R \subset \mathbb{R}^n$ is a cell, i.e., a product of intervals $R = I_1 \times \cdots \times I_n$, where $I_{\nu} = [a_{\nu}, b_{\nu}]$ are intervals in $\mathbb{R}$. Recall that a partition of an interval $I = [a, b]$ is a finite collection of subintervals $\{J_k : 0 \leq k \leq N\}$, disjoint except for their endpoints, whose union is $I$. We can take $J_k = [x_k, x_{k+1}]$, where

$$a = x_0 < x_1 < \cdots < x_N < x_{N+1} = b.$$  

Now, if one has a partition of each $I_{\nu}$ into $J_{\nu,1} \cup \cdots \cup J_{\nu,N(\nu)}$, then a partition $P$ of $R$ consists of the cells

$$R_\alpha = J_{1\alpha_1} \times J_{2\alpha_2} \times \cdots \times J_{n\alpha_n},$$

where $0 \leq \alpha_\nu \leq N(\nu)$. For such a partition, define

$$\text{maxsize}(P) = \max_\alpha \text{diam } R_\alpha,$$

where $(\text{diam } R_\alpha)^2 = \ell(J_{1\alpha_1})^2 + \cdots + \ell(J_{n\alpha_n})^2$. Here, $\ell(J)$ denotes the length of an interval $J$. Each cell has $n$-dimensional volume

$$V(R_\alpha) = \ell(J_{1\alpha_1}) \cdots \ell(J_{n\alpha_n}).$$

Sometimes we use $V_n(R_\alpha)$ for emphasis on the dimension. We also use $A(R)$ for $V_2(R)$, and, of course, $\ell(R)$ for $V_1(R)$.

We set

$$\bar{I}_P(f) = \sum \sup_{R_\alpha} f(x) V(R_\alpha),$$

$$\underline{I}_P(f) = \sum \inf_{R_\alpha} f(x) V(R_\alpha).$$

Note that $\underline{I}_P(f) \leq \bar{I}_P(f)$. These quantities should approximate the Riemann integral of $f$, if the partition $P$ is sufficiently “fine.”

To be more precise, if $P$ and $Q$ are two partitions of $R$, we say $P$ refines $Q$, and write $P \succ Q$, if each partition of each interval factor $I_{\nu}$ of $R$ involved in the definition of $Q$ is further refined in order to produce the partitions of the factors $I_{\nu}$, used to define $P$, via (4.2). It is an exercise to show that any two partitions of $R$ have a common refinement. Note also that

$$P \succ Q \implies \bar{I}_P(f) \leq \bar{I}_Q(f), \text{ and } \underline{I}_P(f) \geq \underline{I}_Q(f).$$
Consequently, if \( P_j \) are any two partitions of \( R \) and \( Q \) is a common refinement, we have

\[
I_{P_1}(f) \leq I_Q(f) \leq \bar{I}_Q(f) \leq \bar{I}_{P_2}(f).
\]

(4.7)

Now, whenever \( f : R \to \mathbb{R} \) is bounded, the following quantities are well defined:

\[
\bar{I}(f) = \inf_{P \in \Pi(R)} \bar{I}_P(f), \quad \underline{I}(f) = \sup_{P \in \Pi(R)} \underline{I}_P(f),
\]

(4.8)

where \( \Pi(R) \) is the set of all partitions of \( R \), as defined above. Clearly, by (4.7), \( \underline{I}(f) \leq \bar{I}(f) \).

We then say that \( f \) is Riemann integrable (on \( R \)) provided \( \bar{I}(f) = \underline{I}(f) \), and in such a case, we set

\[
\int_R f(x) \, dV(x) = \bar{I}(f) = \underline{I}(f).
\]

(4.9)

We will denote the set of Riemann integrable functions on \( R \) by \( R(R) \). If \( \text{dim} \, R = 2 \), we will often use \( dA(x) \) instead of \( dV(x) \) in (4.9). For general \( n \), we might also use simply \( dx \).

We derive some basic properties of the Riemann integral. First, the proofs of Proposition 0.3 and Corollary 0.4 readily extend, to give:

**Proposition 4.1.** Let \( P_{\nu} \) be any sequence of partitions of \( R \) such that

\[
\maxsize(P_{\nu}) = \delta_{\nu} \to 0,
\]

and let \( \xi_{\nu\alpha} \) be any choice of one point in each cell \( R_{\nu\alpha} \) in the partition \( P_{\nu} \). Then, whenever \( f \in R(R) \),

\[
\int_R f(x) \, dV(x) = \lim_{\nu \to \infty} \sum_{\alpha} f(\xi_{\nu\alpha}) \, V(R_{\nu\alpha}).
\]

(4.10)

Also, we can extend the proof of Proposition 0.1, to obtain:

**Proposition 4.2.** If \( f_j \in R(R) \) and \( c_j \in \mathbb{R} \), then \( c_1 f_1 + c_2 f_2 \in R(R) \), and

\[
\int_R (c_1 f_1 + c_2 f_2) \, dV = c_1 \int_R f_1 \, dV + c_2 \int_R f_2 \, dV.
\]

(4.11)

Next, we establish an integrability result analogous to Proposition 0.2.
Proposition 4.3. If $f$ is continuous on $R$, then $f \in \mathcal{R}(R)$.

Proof. As in the proof of Proposition 0.2, we have that,

$$\text{maxsize}(\mathcal{P}) \leq \delta \implies \mathcal{I}_\mathcal{P}(f) - \mathcal{I}_\mathcal{P}(f) \leq \omega(\delta) \cdot V(R),$$

where $\omega(\delta)$ is a modulus of continuity for $f$ on $R$. This proves the proposition.

When the number of variables exceeds one, it becomes more important to identify some nice classes of discontinuous functions on $R$ that are Riemann integrable. A useful tool for this is the following notion of size of a set $S \subset R$, called content. Extending (0.18)-(0.19), we define “upper content” $\text{cont}^+$ and “lower content” $\text{cont}^-$ by

$$\text{cont}^+(S) = \mathcal{T}(\chi_S), \quad \text{cont}^-(S) = \mathcal{I}(\chi_S),$$

where $\chi_S$ is the characteristic function of $S$. We say $S$ has content, or “is contented,” if these quantities are equal, which happens if and only if $\chi_S \in \mathcal{R}(R)$, in which case the common value of $\text{cont}^+(S)$ and $\text{cont}^-(S)$ is

$$V(S) = \int_R \chi_S(x) \, dV(s).$$

We mention that, if $S = I_1 \times \cdots \times I_n$ is a cell, it is readily verified that the definitions in (4.5), (4.8), and (4.13) yield

$$\text{cont}^+(S) = \text{cont}^-(S) = \ell(I_1) \cdots \ell(I_n),$$

so the definition of $V(S)$ given by (4.14) is consistent with that given in (4.4).

It is easy to see that

$$\text{cont}^+(S) = \inf \left\{ \sum_{k=1}^N V(R_k) : S \subset R_1 \cup \cdots \cup R_N \right\},$$

where $R_k$ are cells contained in $R$. In the formal definition, the $R_\alpha$ in (4.15) should be part of a partition $\mathcal{P}$ of $R$, as defined above, but if $\{R_1, \ldots, R_N\}$ are any cells in $R$, they can be chopped up into smaller cells, some perhaps thrown away, to yield a finite cover of $S$ by cells in a partition of $R$, so one gets the same result.

It is an exercise to see that, for any set $S \subset R$,

$$\text{cont}^+(S) = \text{cont}^+(\overline{S}),$$

where $\overline{S}$ is the closure of $S$.

We note that, generally, for a bounded function $f$ on $R$,

$$\mathcal{I}(f) + \mathcal{I}(1-f) = V(R).$$
This follows directly from (4.5). In particular, given $S \subset R$,

$$\text{cont}^-(S) + \text{cont}^+(R \setminus S) = V(R).$$

Using this together with (4.16), with $S$ and $R \setminus S$ switched, we have

$$\text{cont}^-(S) = \text{cont}^-(\overset{\circ}S),$$

where $\overset{\circ}S$ is the interior of $S$. The difference $\overline{S} \setminus \overset{\circ}S$ is called the boundary of $S$, and denoted $bS$.

Note that

$$\text{cont}^-(S) = \sup \left\{ \sum_{k=1}^{N} V(R_k) : R_1 \cup \cdots \cup R_N \subset S \right\},$$

where here we take $\{R_1, \ldots, R_N\}$ to be cells within a partition $\mathcal{P}$ of $R$, and let $\mathcal{P}$ vary over all partitions of $R$. Now, given a partition $\mathcal{P}$ of $R$, classify each cell in $\mathcal{P}$ as either being contained in $R \setminus \overline{S}$, or intersecting $bS$, or contained in $\overset{\circ}S$. Letting $\mathcal{P}$ vary over all partitions of $R$, we see that

$$\text{cont}^+(\overline{S}) = \text{cont}^+(bS) + \text{cont}^-(\overset{\circ}S).$$

In particular, we have:

**Proposition 4.4.** If $S \subset R$, then $S$ is contented if and only if $\text{cont}^+(bS) = 0$.

If a set $\Sigma \subset R$ has the property that $\text{cont}^+(\Sigma) = 0$, we say that $\Sigma$ has content zero, or is a nil set. Clearly $\Sigma$ is nil if and only if $\overline{\Sigma}$ is nil. It follows easily from Proposition 4.2 that, if $\Sigma_j$ are nil, $1 \leq j \leq K$, then $\bigcup_{j=1}^{K} \Sigma_j$ is nil.

If $S_1, S_2 \subset R$ and $S = S_1 \cup S_2$, then $\overline{S} = \overline{S}_1 \cup \overline{S}_2$ and $\overset{\circ}S \supset \overset{\circ}S_1 \cup \overset{\circ}S_2$. Hence $bS \subset b(S_1) \cup b(S_2)$. It follows then from Proposition 4.4 that, if $S_1$ and $S_2$ are contented, so is $S_1 \cup S_2$. Clearly, if $S_j$ are contented, so are $S_j^c = R \setminus S_j$. It follows that, if $S_1$ and $S_2$ are contented, so is $S_1 \cap S_2 = (S_1^c \cup S_2^c)^c$. A family $\mathcal{F}$ of subsets of $R$ is called an algebra of subsets of $R$ provided the following conditions hold:

$$R \in \mathcal{F},$$

$$S_j \in \mathcal{F} \Rightarrow S_1 \cup S_2 \in \mathcal{F},$$

$$S \in \mathcal{F} \Rightarrow R \setminus S \in \mathcal{F}.$$

Algebras of sets are automatically closed under finite intersections also. We see that:

**Proposition 4.5.** The family of contented subsets of $R$ is an algebra of sets.

The following result specifies a useful class of Riemann integrable functions. For a sharper result, see Proposition 4.31.
Proposition 4.6. If \( f : \mathbb{R} \to \mathbb{R} \) is bounded and the set \( S \) of points of discontinuity of \( f \) is a nil set, then \( f \in \mathcal{R}(\mathbb{R}) \).

Proof. Suppose \( |f| \leq M \) on \( R \), and take \( \varepsilon > 0 \). Take a partition \( \mathcal{P} \) of \( R \), and write \( \mathcal{P} = \mathcal{P}' \cup \mathcal{P}'' \), where cells in \( \mathcal{P}' \) do not meet \( S \), and cells in \( \mathcal{P}'' \) do intersect \( S \). Since \( \text{cont}^+(S) = 0 \), we can pick \( \mathcal{P} \) so that the cells in \( \mathcal{P}'' \) have total volume \( \leq \varepsilon \). Now \( f \) is continuous on each cell in \( \mathcal{P}' \). Further refining the partition if necessary, we can assume that \( f \) varies by \( \leq \varepsilon \) on each cell in \( \mathcal{P}' \). Thus

\[
\bar{I}_\mathcal{P}(f) - \underline{I}_\mathcal{P}(f) \leq \left[ V(R) + 2M \right] \varepsilon.
\]

This proves the proposition.

To give an example, suppose \( K \subset \mathbb{R} \) is a closed set such that \( bK \) is nil. Let \( f : K \to \mathbb{R} \) be continuous. Define \( \tilde{f} : \mathbb{R} \to \mathbb{R} \) by

\[
\tilde{f}(x) = f(x) \quad \text{for } x \in K, \quad 0 \quad \text{for } x \in \mathbb{R} \setminus K.
\]

Then the set of points of discontinuity of \( \tilde{f} \) is contained in \( bK \). Hence \( \tilde{f} \in \mathcal{R}(\mathbb{R}) \). We set

\[
\int_K f \, dV = \int_{\mathbb{R}} \tilde{f} \, dV.
\]

In connection with this, we note the following fact, whose proof is an exercise. Suppose \( R \) and \( \hat{R} \) are cells, with \( R \subset \hat{R} \). Suppose that \( g \in \mathcal{R}(\mathbb{R}) \) and that \( \hat{g} \) is defined on \( R \), to be equal to \( g \) on \( R \) and to be 0 on \( \hat{R} \setminus R \). Then

\[
\hat{g} \in \mathcal{R}(\hat{R}), \quad \text{and} \quad \int_R g \, dV = \int_{\hat{R}} \hat{g} \, dV.
\]

This can be shown by an argument involving refining any given pair of partitions of \( R \) and \( \hat{R} \), respectively, to a pair of partitions \( \mathcal{P}_R \) and \( \mathcal{P}_{\hat{R}} \) with the property that each cell in \( \mathcal{P}_R \) is a cell in \( \mathcal{P}_{\hat{R}} \).

The following is a useful criterion for a set \( S \subset \mathbb{R}^n \) to have content zero.

Proposition 4.7. Let \( \Sigma \subset \mathbb{R}^{n-1} \) be a closed bounded set and let \( g : \Sigma \to \mathbb{R} \) be continuous. Then the graph of \( g \),

\[
\mathfrak{G} = \{ (x, g(x)) : x \in \Sigma \}
\]

is a nil subset of \( \mathbb{R}^n \).

Proof. Put \( \Sigma \) in a cell \( R_0 \subset \mathbb{R}^{n-1} \). Suppose \( |f| \leq M \) on \( \Sigma \). Take \( N \in \mathbb{Z}^+ \) and set \( \varepsilon = M/N \). Pick a partition \( \mathcal{P}_0 \) of \( R_0 \), sufficiently fine that \( g \) varies by at most \( \varepsilon \) on each cell \( \Sigma \cap R_\alpha \), for any cell \( R_\alpha \in \mathcal{P}_0 \). Partition the interval \( I = [-M, M] \) into \( 2N \) equal intervals \( J_\nu \), of length...
Then \( R_\alpha \times J_\nu = \{ Q_{\alpha \nu} \} \) forms a partition of \( R_0 \times I \). Now, over each cell \( R_\alpha \in \mathcal{P}_0 \), there lie at most 2 cells \( Q_{\alpha \nu} \) which intersect \( \mathcal{G} \), so \( \text{cont}^+(\mathcal{G}) \leq 2\varepsilon \cdot V(R_0) \). Letting \( N \to \infty \), we have the proposition.

Similarly, for any \( j \in \{ 1, \ldots, n \} \), the graph of \( x_j \) as a continuous function of the complementary variables is a nil set in \( \mathbb{R}^n \). So are finite unions of such graphs. Such sets arise as boundaries of many ordinary-looking regions in \( \mathbb{R}^n \).

Here is a further class of nil sets.

**Proposition 4.8.** Let \( O \subset \mathbb{R}^n \) be open and let \( S \subset O \) be a compact nil subset. Assume \( f : O \to \mathbb{R}^n \) is a Lipschitz map. Then \( f(S) \) is a nil subset of \( \mathbb{R}^n \).

**Proof.** The Lipschitz hypothesis on \( f \) is that there exists \( L < \infty \) such that, for \( p, q \in O \),

\[
|f(p) - f(q)| \leq L|p - q|.
\]

If we cover \( S \) with \( k \) cells (in a partition), of total volume \( \leq \alpha \), each cubical with edgesize \( \delta \), then \( f(S) \) is covered by \( k \) sets of diameter \( \leq L\sqrt{n}\delta \), hence it can be covered by \( k \) cubical cells of edgesize \( L\sqrt{n}\delta \), having total volume \( \leq (L\sqrt{n})^n \alpha \). From this we have the (not very sharp) general bound

\[
(4.26) \quad \text{cont}^+(f(S)) \leq (L\sqrt{n})^n \text{cont}^+(S),
\]

which proves the proposition.

In evaluating \( n \)-dimensional integrals, it is usually convenient to reduce them to iterated integrals. The following is a special case of a result known as Fubini’s Theorem.

**Theorem 4.9.** Let \( \Sigma \subset \mathbb{R}^{n-1} \) be a closed, bounded contented set and let \( g_j : \Sigma \to \mathbb{R} \) be continuous, with \( g_0(x) < g_1(x) \) on \( \Sigma \). Take

\[
(4.27) \quad \Omega = \{ (x, y) \in \mathbb{R}^n : x \in \Sigma, \ g_0(x) \leq y \leq g_1(x) \}.
\]

Then \( \Omega \) is a contented set in \( \mathbb{R}^n \). If \( f : \Omega \to \mathbb{R} \) is continuous, then

\[
(4.28) \quad \varphi(x) = \int_{g_0(x)}^{g_1(x)} f(x, y) \, dy
\]

is continuous on \( \Sigma \), and

\[
(4.29) \quad \int_{\Omega} f \, dV_n = \int_{\Sigma} \varphi \, dV_{n-1},
\]

i.e.,

\[
(4.30) \quad \int_{\Omega} f \, dV_n = \int_{\Sigma} \left( \int_{g_0(x)}^{g_1(x)} f(x, y) \, dy \right) \, dV_{n-1}(x).
\]
Proof. The continuity of (4.28) is an exercise in one-variable integration; see Exercises 2–3 of §0. Let \( \omega(\delta) \) be a modulus of continuity for \( g_0, g_1, \) and \( f, \) and also \( \varphi. \) We also can assume that \( \omega(\delta) \geq \delta. \)

Put \( \Sigma \) in a cell \( R_0 \) and let \( \mathcal{P}_0 \) be a partition of \( R_0. \) If \( A \leq g_0 \leq g_1 \leq B, \) partition the interval \([A, B],\) and from this and \( \mathcal{P}_0 \) construct a partition \( \mathcal{P} \) of \( R = R_0 \times [A, B]. \) We denote a cell in \( \mathcal{P}_0 \) by \( R_\alpha \) and a cell in \( \mathcal{P} \) by \( R_{\alpha \ell} = R_\alpha \times J_\ell. \) Pick points \( x_{\alpha \ell} \in R_{\alpha \ell}. \)

Write \( \mathcal{P}_0 = \mathcal{P}_0' \cup \mathcal{P}_0'' \cup \mathcal{P}_0''' \), consisting respectively of cells inside \( \hat{\Sigma}, \) meeting \( b\Sigma, \) and inside \( R_0 \setminus \Sigma. \) Similarly write \( \mathcal{P} = \mathcal{P}' \cup \mathcal{P}'' \cup \mathcal{P}''' \), consisting respectively of cells inside \( \hat{\Omega}, \) meeting \( b\Omega, \) and inside \( R \setminus \Omega, \) as illustrated in Fig. 4.1. For fixed \( \alpha, \) let

\[
z'(\alpha) = \{ \ell: R_{\alpha \ell} \in \mathcal{P}' \},
\]

and let \( z''(\alpha) \) and \( z'''(\alpha) \) be similarly defined. Note that

\[
z'(\alpha) \neq \emptyset \iff R_\alpha \in \mathcal{P}_0',
\]

provided we assume \( \max(\mathcal{P}) \leq \delta \) and \( 2\delta < \min[g_1(x) - g_0(x)], \) as we will from here on.

It follows from (0.9)–(0.10) that

\[
(4.31) \quad \left| \sum_{\ell \in z'(\alpha)} f(\xi_{\alpha \ell})\ell(J_\ell) - \int_{A(\alpha)} f(x, y) \, dy \right| \leq (B - A)\omega(\delta), \quad \forall \, x \in R_\alpha,
\]

where \( \bigcup_{\ell \in z'(\alpha)} J_\ell = [A(\alpha), B(\alpha)]. \) Note that \( A(\alpha) \) and \( B(\alpha) \) are within \( 2\omega(\delta) \) of \( g_0(x) \) and \( g_1(x), \) respectively, for all \( x \in R_\alpha, \) if \( R_\alpha \in \mathcal{P}_0'. \) Hence, if \( |f| \leq M, \)

\[
(4.32) \quad \left| \int_{A(\alpha)} f(x, y) \, dy - \varphi(x) \right| \leq 4M\omega(\delta), \quad \forall \, x \in R_\alpha.
\]

Thus, with \( C = B - A + 4M, \)

\[
(4.33) \quad \left| \sum_{\ell \in z'(\alpha)} f(\xi_{\alpha \ell})\ell(J_\ell) - \varphi(x) \right| \leq C\omega(\delta), \quad \forall \, x \in R_\alpha \in \mathcal{P}_0'.
\]

Multiplying by \( V_{n-1}(R_\alpha) \) and summing over \( R_\alpha \in \mathcal{P}_0', \) we have

\[
(4.34) \quad \left| \sum_{R_{\alpha \ell} \in \mathcal{P}'} f(\xi_{\alpha \ell})V_n(R_{\alpha \ell}) - \sum_{R_\alpha \in \mathcal{P}_0'} \varphi(x_\alpha)V_{n-1}(R_\alpha) \right| \leq CV(0)\omega(\delta),
\]

where \( x_\alpha \) is an arbitrary point in \( R_\alpha. \)

Now, if \( \mathcal{P}_0 \) is a sufficiently fine partition of \( R_0, \) it follows from the proof of Proposition 4.6 that the second sum in (4.34) is arbitrarily close to \( \int_\Sigma \varphi \, dV_{n-1}, \) since \( b\Sigma \) has content zero. Furthermore, an argument such as used to prove Proposition 4.7 shows that \( b\Omega \) has content zero, and one verifies that, for a sufficiently fine partition, the first sum in (4.34) is arbitrarily close to \( \int_{\Omega} f \, dV_n. \) This proves the desired identity (4.29).

We next take up the change of variables formula for multiple integrals, extending the one-variable formula, (0.42). We begin with a result on linear changes of variables. The set of invertible real \( n \times n \) matrices is denoted \( GL(n, \mathbb{R}). \) In (4.35) and subsequent formulas, \( \int f \, dV \) denotes \( \int_R f \, dV \) for some cell \( R \) on which \( f \) is supported. The integral is independent of the choice of such a cell; cf. (4.25).
Proposition 4.10. Let $f$ be a continuous function with compact support in $\mathbb{R}^n$. If $A \in \text{Gl}(n, \mathbb{R})$, then

$$ (4.35) \quad \int f(x) \, dV = |\det A| \int f(Ax) \, dV. $$

Proof. Let $\mathcal{G}$ be the set of elements $A \in \text{Gl}(n, \mathbb{R})$ for which (4.35) is true. Clearly $I \in \mathcal{G}$. Using $\det A^{-1} = (\det A)^{-1}$, and $\det AB = (\det A)(\det B)$, we can conclude that $\mathcal{G}$ is a subgroup of $\text{Gl}(n, \mathbb{R})$.

In more detail, for $A \in \text{Gl}(n, \mathbb{R})$, $f$ as above, let

$$ I_A(f) = \int f_A \, dV = I(f_A), \quad f_A(x) = f(Ax). $$

Then

$$ A \in \mathcal{G} \iff I_A(f) = |\det A|^{-1}I(f), $$

for all such $f$. We see that

$$ I_{AB}(f) = I(f_{AB}) = I_B(f_A), $$

so

$$ A, B \in \mathcal{G} \implies I_{AB}(f) = |\det B|^{-1}I(f_A) $$

$$ = |\det B|^{-1}|\det A|^{-1}I(f) = |\det AB|^{-1}I(f) $$

$$ \implies AB \in \mathcal{G}. $$

Applying a similar argument to $I_{AA^{-1}}(f) = I(f)$, also yields the implication $A \in \mathcal{G} \implies A^{-1} \in \mathcal{G}$.

To prove the proposition, it will therefore suffice to show that $\mathcal{G}$ contains all elements of the following 3 forms, since (as shown in the exercises on row reduction at the end of this section) the method of applying elementary row operations to reduce a matrix shows that any element of $\text{Gl}(n, \mathbb{R})$ is a product of a finite number of these elements. Here, $\{e_j : 1 \leq j \leq n\}$ denotes the standard basis of $\mathbb{R}^n$, and $\sigma$ a permutation of $\{1, \ldots, n\}$.

$$ (4.36) \begin{align*}
A_1 e_j &= e_{\sigma(j)}, \\
A_2 e_j &= e_j e_j, \quad e_j \neq 0 \\
A_3 e_2 &= e_2 + ce_1, \quad A_3 e_j &= e_j \text{ for } j \neq 2.
\end{align*} $$

The proofs of (4.35) in the first two cases are elementary consequences of the definition of the Riemann integral, and can be left as exercises.

We show that (4.35) holds for transformations of the form $A_3$ by using Theorem 4.9 (in a special case), to reduce it to the case $n = 1$. Given $f \in C(\mathbb{R})$, compactly supported, and $b \in \mathbb{R}$, we clearly have

$$ (4.37) \quad \int f(x) \, dx = \int f(x + b) \, dx. $$
Now, for the case \( A = A_3 \), with \( x = (x_1, x') \), we have

\[
\int f(x_1 + cx_2, x') \, dV_n(x) = \int \left( \int f(x_1 + cx_2, x') \, dx_1 \right) \, dV_{n-1}(x')
\]

(4.38)

\[
= \int \left( \int f(x_1, x') \, dx_1 \right) \, dV_{n-1}(x'),
\]

the second identity by (4.37). Thus we get (4.35) in case \( A = A_3 \); so the proposition is proved.

It is desirable to extend Proposition 4.10 to some discontinuous functions. Given a cell \( R \) and \( f : R \to \mathbb{R} \), bounded, we say

\[
f \in \mathcal{C}(R) \iff \text{the set of discontinuities of } f \text{ is nil.}
\]

(4.39)

Proposition 4.6 implies

\[
\mathcal{C}(R) \subset \mathcal{R}(R).
\]

From the closure of the class of nil sets under finite unions it is clear that \( \mathcal{C}(R) \) is closed under sums and products, i.e., that \( \mathcal{C}(R) \) is an algebra of functions on \( R \). We will denote by \( \mathcal{C}_c(\mathbb{R}^n) \) the set of bounded functions \( f : \mathbb{R}^n \to \mathbb{R} \) such that \( f \) has compact support and its set of discontinuities is nil. Any \( f \in \mathcal{C}_c(\mathbb{R}^n) \) is supported in some cell \( R \), and \( f|_R \in \mathcal{C}(R) \).

Here is another useful class of functions. Given a cell \( R \subset \mathbb{R}^n \) and \( f : R \to \mathbb{R} \) bounded, we say

\[
f \in \mathcal{PK}(R) \iff \exists \text{ a partition } P \text{ of } R \text{ such that } f \text{ is constant on the interior of each cell } R_\alpha \in P.
\]

(4.41)

The following will be a useful tool for extending Proposition 4.10. It is also of interest in its own right, and it will have other uses.

**Proposition 4.11.** Given a cell \( R \subset \mathbb{R}^n \) and \( f : R \to \mathbb{R} \) bounded,

\[
I(f) = \inf \left\{ \int_R g \, dV : g \in \mathcal{PK}(R), g \geq f \right\}
\]

\[
= \inf \left\{ \int_R g \, dV : g \in \mathcal{C}(R), g \geq f \right\}
\]

\[
= \inf \left\{ \int_R g \, dV : g \in \mathcal{C}(R), g \geq f \right\}.
\]

(4.42)

Similarly,

\[
I(f) = \sup \left\{ \int_R g \, dV : g \in \mathcal{PK}(R), g \leq f \right\}
\]

\[
= \sup \left\{ \int_R g \, dV : g \in \mathcal{C}(R), g \leq f \right\}
\]

\[
= \sup \left\{ \int_R g \, dV : g \in \mathcal{C}(R), g \leq f \right\}.
\]

(4.43)
Proof. Denote the three quantities on the right side of (4.42) by \( T_1(f), T_2(f), \) and \( T_3(f) \), respectively. The definition of \( T_1(f) \) is sufficiently close to that of \( T(f) \) in (4.8) that the identity \( T(f) = T_1(f) \) is apparent. Now \( T_2(f) \) is an inf over a larger class of functions \( g \) than that defining \( T_1(f) \), so \( T_2(f) \leq T_1(f) \). On the other hand, \( T(g) \geq T(f) \) for all \( g \) involved in defining \( T_2(f) \), so \( T_2(f) \geq T(f) \), hence \( T_2(f) = T(f) \).

Next, \( T_3(f) \) is an inf over a larger class of functions \( g \) than that defining \( T_2(f) \), so \( T_3(f) \geq T(f) \). On the other hand, given \( \varepsilon > 0 \) and \( \psi \in \text{PK}(R) \), one can readily find \( g \in C(R) \) such that \( g \geq \psi \) and \( \int_R (g - \psi) \, dV < \varepsilon \). This implies \( T_3(f) \leq T(f) + \varepsilon \) for all \( \varepsilon > 0 \) and finishes the proof of (4.42). The proof of (4.43) is similar.

We can now extend Proposition 4.10. Say \( f \in \mathcal{R}_c(\mathbb{R}^n) \) if \( f \) has compact support, say in some cell \( R \), and \( f \in \mathcal{R}(R) \). Also say \( f \in C_c(\mathbb{R}^n) \) if \( f \) is continuous on \( \mathbb{R}^n \), with compact support.

**Proposition 4.12.** Given \( A \in \text{Gl}(n, \mathbb{R}) \), the identity (4.35) holds for all \( f \in \mathcal{R}_c(\mathbb{R}^n) \).

*Proof.* We have from Proposition 4.11 that, for each \( \nu \in \mathbb{N} \), there exist \( g_\nu, h_\nu \in C_c(\mathbb{R}^n) \) such that \( h_\nu \leq f \leq g_\nu \) and, with \( B = \int f \, dV \),

\[
B - \frac{1}{\nu} \leq \int h_\nu \, dV \leq B \leq \int g_\nu \, dV \leq B + \frac{1}{\nu}.
\]

Now Proposition 4.10 applies to \( g_\nu \) and \( h_\nu \), so

\[
(4.44) \quad B - \frac{1}{\nu} \leq |\det A| \int h_\nu(Ax) \, dV \leq B \leq |\det A| \int g_\nu(Ax) \, dV \leq B + \frac{1}{\nu}. \tag{4.44}
\]

Furthermore, with \( f_A(x) = f(Ax) \), we have \( h_\nu(Ax) \leq f_A(x) \leq g_\nu(Ax) \), so (4.44) gives

\[
(4.45) \quad B - \frac{1}{\nu} \leq |\det A| T(f_A) \leq |\det A| T(f_A) \leq B + \frac{1}{\nu},
\]

for all \( \nu \), and letting \( \nu \to \infty \) we obtain (4.35).

**Corollary 4.13.** If \( \Sigma \subset \mathbb{R}^n \) is a compact, contented set and \( A \in \text{Gl}(n, \mathbb{R}) \), then \( A(\Sigma) = \{ Ax : x \in \Sigma \} \) is contented, and

\[
(4.46) \quad V(A(\Sigma)) = |\det A| V(\Sigma).
\]

We now extend Proposition 4.10 to nonlinear changes of variables.

**Proposition 4.14.** Let \( \mathcal{O} \) and \( \Omega \) be open in \( \mathbb{R}^n \), \( G : \mathcal{O} \to \Omega \) a \( C^1 \) diffeomorphism, and \( f \) a continuous function with compact support in \( \Omega \). Then

\[
(4.47) \quad \int_{\Omega} f(y) \, dV(y) = \int_{\mathcal{O}} f(G(x)) |\det DG(x)| \, dV(x).
\]
Proof. It suffices to prove the result under the additional assumption that \( f \geq 0 \), which we make from here on. Also, using a partition of unity (see Appendix B), we can write \( f \) as a finite sum of continuous functions with small supports, so it suffices to treat the case where \( f \) is supported in a cell \( \tilde{R} \subset \Omega \) and \( f \circ G \) is supported in a cell \( R \subset \mathcal{O} \). See Fig. 4.2. Let \( \mathcal{P} = \{ R_\alpha \} \) be a partition of \( R \). Note that for each \( R_\alpha \in \mathcal{P} \), \( bG(R_\alpha) = G(bR_\alpha) \), so \( G(R_\alpha) \) is contented, in view of Propositions 4.4 and 4.8.

Let \( \xi_\alpha \) be the center of \( R_\alpha \), and let \( \tilde{R}_\alpha = R_\alpha - \xi_\alpha \), a cell with center at the origin. Then

\[
G(\xi_\alpha) + DG(\xi_\alpha)(\tilde{R}_\alpha) = \eta_\alpha + H_\alpha
\]

is an \( n \)-dimensional parallelipiped, each point of which is very close to a point in \( G(R_\alpha) \), if \( R_\alpha \) is small enough. To be precise, for \( y \in \tilde{R}_\alpha \),

\[
G(\xi_\alpha + y) = \eta_\alpha + DG(\xi_\alpha)y + \Phi(\xi_\alpha, y) y,
\]

\[
\Phi(\xi_\alpha, y) = \int_0^1 [DG(\xi_\alpha + ty) - DG(\xi_\alpha)] \, dt.
\]

See Fig. 4.3. Consequently, given \( \varepsilon > 0 \), if \( \delta > 0 \) is small enough and \( \text{maxsize}(\mathcal{P}) \leq \delta \), then we have

\[
\eta_\alpha + (1 + \varepsilon)H_\alpha \supset G(R_\alpha),
\]

for all \( R_\alpha \in \mathcal{P} \). Now, by (4.46),

\[
V(H_\alpha) = |\det DG(\xi_\alpha)| V(R_\alpha).
\]

Hence

\[
V(G(R_\alpha)) \leq (1 + \varepsilon)^n |\det DG(\xi_\alpha)| V(R_\alpha).
\]

Now we have

\[
\int f \, dV = \sum_{\alpha} \int_{G(R_\alpha)} f \, dV
\]

\[
\leq \sum_{\alpha} \sup_{R_\alpha} f \circ G(x) V(G(R_\alpha))
\]

\[
\leq (1 + \varepsilon)^n \sum_{\alpha} \sup_{R_\alpha} f \circ G(x) |\det DG(\xi_\alpha)| V(R_\alpha).
\]

To see that the first line of (4.52) holds, note that \( f\chi_{G(R_\alpha)} \) is Riemann integrable, by Proposition 4.6; note also that \( \sum_\alpha f\chi_{G(R_\alpha)} = f \) except on a set of content zero. Then the additivity result in Proposition 4.2 applies. The first inequality in (4.52) is elementary; the second inequality uses (4.51) and \( f \geq 0 \). If we set

\[
h(x) = f \circ G(x) |\det DG(x)|,
\]
then we have

\[\sup_{\mathcal{R}_a} f \circ G(x) |\det DG(\xi_\alpha)| \leq \sup_{\mathcal{R}_a} h(x) + M\omega(\delta),\]

provided \(|f| \leq M\) and \(\omega(\delta)\) is a modulus of continuity for \(DG\). Taking arbitrarily fine partitions, we get

\[\int_{\Omega} f \, dV \leq \int_{\mathcal{O}} h \, dV.\]

If we apply this result, with \(G\) replaced by \(G^{-1}\), \(\mathcal{O}\) and \(\Omega\) switched, and \(f\) replaced by \(h\), given by (4.53), we have

\[\int_{\mathcal{O}} h \, dV \leq \int_{\Omega} h \circ G^{-1}(y) |\det DG^{-1}(y)| \, dV(y) = \int_{\Omega} f \, dV.\]

The inequalities (4.55) and (4.56) together yield the identity (4.47).

We now extend Proposition 4.14 to more general Riemann integrable functions. Recall that \(f \in \mathcal{R}_c(\mathbb{R}^n)\) if \(f\) has compact support, say in some cell \(R\), and \(f \in \mathcal{R}(R)\). If \(\Omega \subset \mathbb{R}^n\) is open and \(f \in \mathcal{R}_c(\mathbb{R}^n)\) has support in \(\Omega\), we say \(f \in \mathcal{C}_c(\Omega)\). We also say \(f \in \mathcal{C}_c(\mathbb{R}^n)\) if \(f\) has support in \(\mathbb{R}^n\), and we say \(f \in C_c(\Omega)\) if \(f\) is continuous with compact support in \(\Omega\). Theorem 4.15. Let \(\mathcal{O}\) and \(\Omega\) be open in \(\mathbb{R}^n\), \(G : \mathcal{O} \to \Omega\) a \(C^1\) diffeomorphism. If \(f \in \mathcal{C}_c(\Omega)\), then \(f \circ G \in \mathcal{C}_c(\mathcal{O})\), and (4.47) holds.

Proof. The proof is similar to that of Proposition 4.12. Given \(\nu \in \mathbb{N}\), we have from Proposition 4.11 that there exist \(g_\nu, h_\nu \in \mathcal{C}_c(\Omega)\) such that \(h_\nu \leq f \leq g_\nu\) and, with \(B = \int_{\Omega} f \, dV\),

\[B - \frac{1}{\nu} \leq \int_{\mathcal{O}} h_\nu \, dV \leq B \leq \int_{\mathcal{O}} g_\nu \, dV \leq B + \frac{1}{\nu}.\]

Then Proposition 4.14 applies to \(h_\nu\) and \(g_\nu\), so

\[B - \frac{1}{\nu} \leq \int_{\mathcal{O}} h_\nu(G(x))|\det DG(x)| \, dV(x) \leq \int_{\mathcal{O}} g_\nu(G(x))|\det DG(x)| \, dV(x) \leq B + \frac{1}{\nu}.\]

Now, with \(f_G(x) = f(G(x))\), we have \(h_\nu(G(x)) \leq f_G(x) \leq g_\nu(G(x))\), so

\[B - \frac{1}{\nu} \leq I(f_G|\det DG|) \leq I(f_G|\det DG|) \leq B + \frac{1}{\nu},\]

for all \(\nu\), and letting \(\nu \to \infty\), we obtain (4.47).

We have seen how Proposition 4.11 has been useful. The following result, to some degree a variant of Proposition 4.11, is also useful.
Lemma 4.16. Let $F : R \to \mathbb{R}$ be bounded, $B \in \mathbb{R}$. Suppose that, for each $\nu \in \mathbb{Z}^+$, there exist $\Psi, \Phi \in \mathcal{R}(R)$ such that

\[(4.57A) \quad \Psi \leq F \leq \Phi,\]

and

\[(4.57B) \quad B - \delta \leq \int_R \Psi(x) dV(x) \leq \int_R \Phi(x) dV(x) \leq B + \delta, \quad \delta \to 0.\]

Then $F \in \mathcal{R}(R)$ and

\[(4.57C) \quad \int_R F(x) dV(x) = B.\]

Furthermore, if there exist $\Psi, \Phi \in \mathcal{R}(R)$ such that $(4.57A)$ holds and

\[(4.57D) \quad \int_R (\Phi(x) - \Psi(x)) dV \leq \delta \to 0,\]

then there exists $B$ such that $(4.57B)$ holds. Hence $F \in \mathcal{R}(R)$ and $(4.57C)$ holds.

The most frequently invoked case of the change of variable formula, in the case $n = 2$, involves the following change from Cartesian to polar coordinates:

\[(4.58) \quad x = r \cos \theta, \quad y = r \sin \theta.\]

Thus, take $G(r, \theta) = (r \cos \theta, r \sin \theta)$. We have

\[(4.59) \quad DG(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}, \quad \det DG(r, \theta) = r.\]

For example, if $\rho \in (0, \infty)$ and

\[(4.60) \quad D_{\rho} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq \rho^2 \},\]

then, for $f \in C(D_{\rho})$,

\[(4.61) \quad \int_{D_{\rho}} f(x, y) dA = \int_0^\rho \int_0^{2\pi} f(r \cos \theta, r \sin \theta) r \, d\theta \, dr.\]

To get this, we first apply Proposition 4.14, with $\mathcal{O} = [\varepsilon, \rho] \times [0, 2\pi - \varepsilon]$, then apply Theorem 4.9, then let $\varepsilon \searrow 0$.

We next use Lemma 4.16 to establish the following useful result on products of Riemann integrable functions.
Proposition 4.17. Given $f_1, f_2 \in \mathcal{R}(R)$, we have $f_1 f_2 \in \mathcal{R}(R)$.

Proof. It suffices to prove this when $f_j \geq 0$. Take partitions $P_\nu$ and functions $\psi_{j\nu}, \varphi_{j\nu} \geq 0$, constant in the interior of each cell in $P_\nu$, such that

$$0 \leq \psi_{j\nu} \leq f_j \leq \varphi_{j\nu} \leq B,$$

and

$$\int \psi_{j\nu} dV, \quad \int \varphi_{j\nu} dV \rightarrow \int f_j dV.$$

We apply Lemma 4.16 with $F = f_1 f_2, \quad \Psi_\nu = \psi_{1\nu} \psi_{2\nu}, \quad \Phi_\nu = \varphi_{1\nu} \varphi_{2\nu}$.

Note that

$$\Phi_\nu - \Psi_\nu = \varphi_{1\nu}(\varphi_{2\nu} - \psi_{2\nu}) + \psi_{2\nu}(\varphi_{1\nu} - \psi_{1\nu})$$

$$\leq B(\varphi_{2\nu} - \psi_{2\nu}) + B(\varphi_{1\nu} - \psi_{1\nu}).$$

Hence (4.57D) holds, giving $F \in \mathcal{R}(R)$.

As a consequence of Proposition 4.17, we can make the following construction. Assume $R$ is a cell and $S \subset R$ is a contented set. If $f \in \mathcal{R}(R)$, we have $\chi_S f \in \mathcal{R}(R)$, by Proposition 4.17. We define

$$(4.61B) \quad \int_S f(x) dV(x) = \int_R \chi_S(x) f(x) dV(x).$$

Note how this extends the scope of (4.24).

Integrals over $\mathbb{R}^n$

It is often useful to integrate a function whose support is not bounded. Generally, given a bounded function $f : \mathbb{R}^n \to \mathbb{R}$, we say

$$f \in \mathcal{R}(\mathbb{R}^n)$$

provided $f|_R \in \mathcal{R}(R)$ for each cell $R \subset \mathbb{R}^n$, and

$$\int_R |f| dV \leq C,$$

for some $C < \infty$, independent of $R$. If $f \in \mathcal{R}(\mathbb{R}^n)$, we set

$$(4.62) \quad \int_{\mathbb{R}^n} f dV = \lim_{s \to \infty} \int_{R_s} f dV, \quad R_s = \{ x \in \mathbb{R}^n : |x_j| \leq s, \forall \ j \}.$$
The existence of the limit in (4.62) can be established as follows. If $M < N$, then
\[ \int_{R_N} f \, dV - \int_{R_M} f \, dV = \int_{R_N \setminus R_M} f \, dV, \]
which is dominated in absolute value by $\int_{R_N \setminus R_M} |f| \, dV$. If $f \in \mathcal{R}(\mathbb{R}^n)$, then $a_N = \int_{R_N} |f| \, dV$ is a bounded monotone sequence, which hence converges, so
\[ \int_{R_N \setminus R_M} |f| \, dV = \int_{R_N} |f| \, dV - \int_{R_M} |f| \, dV \to 0, \quad \text{as } M, N \to \infty. \]

The following simple but useful result is an exercise.

**Proposition 4.18.** If $K_\nu$ is any sequence of compact contented subsets of $\mathbb{R}^n$ such that each $R_s$, for $s < \infty$, is contained in all $K_\nu$ for $\nu$ sufficiently large, i.e., $\nu \geq N(s)$, then, whenever $f \in \mathcal{R}(\mathbb{R}^n)$,
\[ (4.63) \quad \int_{\mathbb{R}^n} f \, dV = \lim_{\nu \to \infty} \int_{K_\nu} f \, dV. \]

Change of variables formulas and Fubini’s Theorem extend to this case. For example, the limiting case of (4.61) as $\rho \to \infty$ is
\[ (4.64) \quad f \in C(\mathbb{R}^2) \cap \mathcal{R}(\mathbb{R}^2) \implies \int_{\mathbb{R}^2} f(x, y) \, dA = \int_0^\infty \int_0^{2\pi} f(r \cos \theta, r \sin \theta) r \, d\theta \, dr. \]

To see this, use Proposition 4.17 with $K_\nu = D_\nu$, defined as in (4.60), to write
\[ (4.64A) \quad \int_{\mathbb{R}^2} f(x, y) \, dA = \lim_{\nu \to \infty} \int_{D_\nu} f(x, y) \, dA, \]
and apply (4.61) to write the integral on the right as
\[ \int_0^\nu \int_0^{2\pi} f(r \cos \theta, r \sin \theta) r \, d\theta \, dr. \]
You get the right side of (4.64) in the limit $\nu \to \infty$.

The following is a good example. Take $f(x, y) = e^{-x^2 - y^2}$. We have
\[ (4.65) \quad \int_{\mathbb{R}^2} e^{-x^2 - y^2} \, dA = \int_0^\infty \int_0^{2\pi} e^{-r^2} r \, d\theta \, dr = 2\pi \int_0^\infty e^{-r^2} r \, dr. \]
Now, methods of \S 0 allow the substitution $s = r^2$, so

$$
\int_0^\infty e^{-r^2} r \, dr = \frac{1}{2} \int_0^\infty e^{-s} \, ds = \frac{1}{2}.
$$

Hence

$$
(4.66) \quad \int_{\mathbb{R}^2} e^{-x^2-y^2} \, dA = \pi.
$$

On the other hand, Theorem 4.9 extends to give

$$
(4.67) \quad \int_{\mathbb{R}^2} e^{-x^2-y^2} \, dA = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2} \, dy \, dx
= \left( \int_{-\infty}^{\infty} e^{-x^2} \, dx \right) \left( \int_{-\infty}^{\infty} e^{-y^2} \, dy \right).
$$

Note that the two factors in the last product are equal. We deduce that

$$
(4.68) \quad \int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi}.
$$

We can generalize (4.67), to obtain (via (4.68))

$$
(4.69) \quad \int_{\mathbb{R}^n} e^{-|x|^2} \, dV_n = \left( \int_{-\infty}^{\infty} e^{-x^2} \, dx \right)^n = \pi^{n/2}.
$$

The integrals (4.65)–(4.69) are called Gaussian integrals, and their evaluation has many uses. We shall see some in \S 5.

We record the following additivity result for the integral over $\mathbb{R}^n$, whose proof is also an exercise.

**Proposition 4.19.** If $f, g \in \mathcal{R}(\mathbb{R}^n)$, then $f + g \in \mathcal{R}(\mathbb{R}^n)$, and

$$
(4.70) \quad \int_{\mathbb{R}^n} (f + g) \, dV = \int_{\mathbb{R}^n} f \, dV + \int_{\mathbb{R}^n} g \, dV.
$$

Unbounded integrable functions
There are lots of unbounded functions we would like to be able to integrate. For example, consider \( f(x) = x^{-1/2} \) on \((0, 1]\) (defined any way you like at \( x = 0 \)). Since, for \( \varepsilon \in (0, 1) \),
\[
\int_{\varepsilon}^{1} x^{-1/2} \, dx = 2 - 2\sqrt{\varepsilon},
\]
this has a limit as \( \varepsilon \searrow 0 \), and it is natural to set
\[
\int_{0}^{1} x^{-1/2} \, dx = 2.
\]
Sometimes (4.72) is called an “improper integral,” but we do not consider that to be a proper designation. We aim for a treatment of the integral for a natural class of unbounded functions. To this end, we define a class \( \mathcal{R}^{\#}(I) \) of not necessarily bounded “integrable” functions on \( I \). The set \( I \) will stand for either \( \mathbb{R}^n \) or a cell in \( \mathbb{R}^n \).

To start, assume \( f \geq 0 \) on \( I \), and for \( A \in (0, \infty) \), set
\[
(4.73) \quad f_A(x) = \begin{cases} f(x) & \text{if } f(x) \leq A, \\ A & \text{if } f(x) > A. \end{cases}
\]
(We hereby abandon the use of \( f_A \) as in the proof of Proposition 4.10.) We say \( f \in \mathcal{R}^{\#}(I) \) provided
\[
(4.74) \quad f_A \in \mathcal{R}(I), \quad \forall A < \infty, \quad \text{and} \quad \exists \text{ uniform bound } \int_{I} f_A \, dV \leq M.
\]
If \( f \geq 0 \) satisfies (4.74), then \( \int_{I} f_A \, dV \) increases monotonically to a finite limit as \( A \nearrow +\infty \), and we call the limit \( \int_{I} f \, dV \):
\[
(4.75) \quad \int_{I} f_A \, dV \nearrow \int_{I} f \, dV, \quad \text{for } f \in \mathcal{R}^{\#}(I), \ f \geq 0.
\]
If \( I \) is understood, we might just write \( \int f \, dV \).

**Remark.** If \( f \in \mathcal{R}(I) \) is \( \geq 0 \), then \( f_A \in \mathcal{R}(I) \) for all \( A < \infty \). See the easy part of Exercise 15.

It is valuable to have the following.

**Proposition 4.20.** If \( f, g : I \to \mathbb{R}^+ \) are in \( \mathcal{R}^{\#}(I) \), then \( f + g \in \mathcal{R}^{\#}(I) \), and
\[
(4.76) \quad \int_{I} (f + g) \, dV = \int_{I} f \, dV + \int_{I} g \, dV.
\]
Proof. To start, note that \((f + g)_A \leq f_A + g_A\). In fact,

\[
(4.77) \quad (f + g)_A = (f_A + g_A)_A.
\]

Hence \((f + g)_A \in \mathcal{R}(I)\) and \(\int (f + g)_A \, dV \leq \int f_A \, dV + \int g_A \, dV \leq \int f \, dV + \int g \, dV\), so we have \(f + g \in \mathcal{R}^\#(I)\) and

\[
(4.78) \quad \int (f + g) \, dV \leq \int f \, dV + \int g \, dV.
\]

On the other hand, if \(B > 2A\), then \((f + g)_B \geq f_A + g_A\), so

\[
(4.79) \quad \int (f + g) \, dV \geq \int f_A \, dV + \int g_A \, dV,
\]

for all \(A < \infty\), and hence

\[
(4.80) \quad \int (f + g) \, dV \geq \int f \, dV + \int g \, dV.
\]

Together, (4.78) and (4.80) yield (4.76).

Next, we take \(f : I \to \mathbb{R}\) and set

\[
(4.81) \quad f = f^+ - f^-, \quad f^+(x) = f(x) \quad \text{if} \quad f(x) \geq 0, \quad 0 \quad \text{if} \quad f(x) < 0.
\]

Then we say

\[
(4.82) \quad f \in \mathcal{R}^\#(I) \iff f^+, f^- \in \mathcal{R}^\#(I),
\]

and set

\[
(4.83) \quad \int_I f \, dV = \int_I f^+ \, dV - \int_I f^- \, dV,
\]

where the two terms on the right are defined as in (4.75). To extend the additivity, we begin as follows

**Lemma 4.21.** Assume that \(g \in \mathcal{R}^\#(I)\) and that \(g_j \geq 0, \ g_j \in \mathcal{R}^\#(I)\), and

\[
(4.84) \quad g = g_0 - g_1.
\]

Then

\[
(4.85) \quad \int_I g \, dV = \int_I g_0 \, dV - \int_I g_1 \, dV.
\]
**Proof.** Take $g = g^+ - g^-$ as in (4.81). Then (4.84) implies

\[(4.86)\quad g^+ + g_1 = g_0 + g^-;\]

which by Proposition 4.19 yields

\[(4.87)\quad \int g^+ dV + \int g_1 dV = \int g_0 dV + \int g^- dV.\]

This implies

\[(4.88)\quad \int g^+ dV - \int g^- dV = \int g_0 dV - \int g_1 dV,\]

which yields (4.85)

We now extend additivity.

**Proposition 4.22.** Assume $f_1, f_2 \in \mathcal{R}^\#(I)$. Then $f_1 + f_2 \in \mathcal{R}^\#(I)$ and

\[(4.89)\quad \int (f_1 + f_2) dV = \int f_1 dV + \int f_2 dV.\]

**Proof.** If $g = f_1 + f_2 = (f_1^+ - f_1^-) + (f_2^+ - f_2^-)$, then

\[(4.90)\quad g = g_0 - g_1, \quad g_0 = f_1^+ + f_2^+, \quad g_1 = f_1^- + f_2^-;\]

We have $g_j \in \mathcal{R}^\#(I)$, and then

\[(4.91)\quad \int (f_1 + f_2) dV = \int g_0 dV - \int g_1 dV = \int (f_1^+ + f_2^+) dV - \int (f_1^- + f_2^-) dV = \int f_1^+ dV + \int f_2^+ dV - \int f_1^- dV - \int f_2^- dV,\]

the first equality by Lemma 4.21, the second tautologically, and the third by Proposition 4.20. Since

\[(4.92)\quad \int f_j dV = \int f_j^+ dV - \int f_j^- dV;\]

this gives (4.89).

If $f : I \to \mathbb{C}$, we set $f = f_1 + i f_2$, $f_j : I \to \mathbb{R}$, and say $f \in \mathcal{R}^\#(I)$ if and only if $f_1$ and $f_2$ belong to $\mathcal{R}^\#(I)$. Then we set

\[(4.93)\quad \int f dV = \int f_1 dV + i \int f_2 dV.\]

Similar comments apply to $f : I \to \mathbb{R}^n$.

We next establish a useful result on products.
Proposition 4.23. Assume \( f \in \mathcal{R}^\#(\mathbb{R}^n) \), \( g \in \mathcal{R}(\mathbb{R}^n) \), and \( f, g \geq 0 \). Then \( fg \in \mathcal{R}^\#(\mathbb{R}^n) \) and
\[
\int f_{A} g_{A} \, dV \nearrow \int fg \, dV \quad \text{as} \quad A \nearrow +\infty.
\]

Proof. Given the additivity properties just established, it would be equivalent to prove this with \( g \) replaced by \( g + 1 \), so we will assume from here that \( g \geq 1 \). Then
\[
(fg)_{A} = (f_{A}g)_{A}.
\]

By Proposition 4.17, \( f_{A}g_{R} \in \mathcal{R}(R) \) for each cell \( R \). Hence (e.g., by the easy part of Exercise 15), \( (f_{A}g)_{A} \in \mathcal{R}(R) \) for each cell \( R \). Thus
\[
(fg)_{A} \in R(R).
\]

Now there exists \( K < \infty \) such that \( 1 \leq g \leq K \), so
\[
f_{A}g \leq K f_{A}, \quad \text{hence} \quad (fg)_{A} \leq K f_{A}.
\]

The hypothesis \( f \in \mathcal{R}^\#(\mathbb{R}^n) \) implies there exists \( M < \infty \) such that
\[
\int f_{A} \, dV \leq M,
\]
for all \( A < \infty \) and each cell \( R \). Hence, by (4.97),
\[
\sup_{A} \int (fg)_{A} \, dV \leq MK,
\]
independent of \( R \). This implies \( fg \in \mathcal{R}^\#(\mathbb{R}^n) \). By definition,
\[
\int (fg)_{A} \, dV \nearrow \int fg \, dV, \quad \text{as} \quad A \nearrow +\infty.
\]

Meanwhile, clearly \( f_{A}g \nearrow A \nearrow \), so the estimate (4.97) implies
\[
\int f_{A}g \, dV \nearrow L, \quad \text{as} \quad A \nearrow +\infty,
\]
for some \( L \in \mathbb{R}^+ \). It remains to identify the limits in (4.100) and (4.101). Now (4.95) implies
\[
(fg)_{A} \leq f_{A}g, \quad \text{hence} \quad \int fg \, dV \leq L.
\]

Finally, since \( f_{A}g \leq fg \) and \( f_{A}g \leq KA \), we have
\[
f_{A}g \leq (fg)_{B} \quad \text{for} \quad B \geq KA.
\]

This implies
\[
L \leq \sup_{B} \int (fg)_{B} \, dV = \int fg \, dV,
\]
and hence we have (4.94).

We now extend the change of variable formula in Theorem 4.15 to unbounded functions. It is convenient to introduce the following notation. Given an open set \( \Omega \subset \mathbb{R}^n \), we say \( f \in \mathcal{R}^\#(\Omega) \) provided \( f \in \mathcal{R}^\#(\mathbb{R}^n) \) and \( f \) is supported on a compact subset of \( \Omega \).
**Proposition 4.24.** Let $\mathcal{O}$ and $\Omega$ be open in $\mathbb{R}^n$, $G : \mathcal{O} \to \Omega$ a $C^1$ diffeomorphism. If $f \in \mathcal{R}^\#(\Omega)$, then $f \circ G \in \mathcal{R}^\#(\mathcal{O})$ and

\[
\int_{\Omega} f(y) \, dV(y) = \int_{\mathcal{O}} f(G(x)) |\det DG(x)| \, dV(x).
\]

**Proof.** It suffices to establish this in case $f \geq 0$, which we assume from here. Then

\[
\int_{\Omega} f_A \, dV \nearrow \int_{\Omega} f \, dV.
\]

We set $\varphi = f \circ G$ and note that $f_A \circ G = \varphi_A$. Hence, by Theorem 4.15, for each $A \in (0, \infty)$,

\[
\int_{\Omega} f_A(y) \, dV(y) = \int_{\mathcal{O}} \varphi_A(x) |\det DG(x)| \, dV(x).
\]

If $f$ is supported on a compact set $K \subset \Omega$, then $\varphi_A$ is supported on $G^{-1}(K) \subset \mathcal{O}$, also compact, hence on which $|\det DG|$ has a positive lower bound. Hence an upper bound on the right side of (4.107) implies an upper bound on $\int \varphi_A \, dV$, independent of $A$, so $\varphi \in \mathcal{R}^\#(\mathbb{R}^n)$. Then Proposition 4.23 implies $\varphi |\det DG| \in \mathcal{R}^\#(\mathbb{R}^n)$ and

\[
\int \varphi_A(x) |\det DG(x)| \, dV(x) \nearrow \int \varphi(x) |\det DG(x)| \, dV(x).
\]


One also has versions of Proposition 4.24 where $f$ need not have compact support. See Exercise 13 below for an example.

Our next result on a class of elements of $\mathcal{R}^\#(I)$ ties in closely with the example in (4.71). As before, $I$ is either $\mathbb{R}^n$ or a cell in $\mathbb{R}^n$.

**Proposition 4.25.** Let $f : I \to [0, \infty)$ and assume $f_A \in \mathcal{R}(I)$ for each $A < \infty$. Assume there are nested contented subsets of $I$:

\[
U_1 \supset U_2 \supset U_3 \supset \cdots, \quad V(U_\nu) \to 0.
\]

Assume $f(1 - \chi_{U_\nu}) \in \mathcal{R}(I)$ for each $\nu$ and that there exists $C < \infty$ such that

\[
\int_{I \setminus U_\nu} f \, dV = J_\nu \leq C, \quad \forall \nu.
\]

Then $f \in \mathcal{R}^\#(I)$ and

\[
J_\nu \nearrow \int_I f \, dV.
\]
Proof. The hypothesis (4.110) implies $J_{\nu} \not\nearrow J$ for some $J \in [0, \infty)$. Also, since $0 \leq f_A \leq f$, we have

\[(4.112) \quad \int_{I \setminus U_\nu} f_A dV \leq J, \quad \forall \nu, A.\]

Furthermore,

\[(4.113) \quad \int f_A dV \leq AV(U_\nu), \quad \forall \nu, A,\]

so

\[(4.114) \quad \int f_A dV \leq J + AV(U_\nu), \quad \forall \nu, A,\]

hence

\[(4.115) \quad \int f_A dV \leq J, \quad \forall A.\]

It follows that $f \in \mathcal{R}^\#(I)$ and

\[(4.116) \quad \int_I f dV \leq J.\]

On the other hand,

\[(4.117) \quad \int_I f dV \geq \int_{I \setminus U_\nu} f dV = J_{\nu},\]

for each $\nu$, so we have (4.111).

Monotone convergence theorem

We aim to establish a circle of results known as monotone convergence theorems. Here is the first result.

Proposition 4.26. Let $R \subset \mathbb{R}^n$ be a cell. Assume $f_k \in \mathcal{R}(R)$. Then

\[(4.118) \quad f_k(x) \searrow 0 \quad \forall x \in R \implies \int_R f_k dV \searrow 0.\]
Proof. It suffices to assume $V(R) = 1$. Say $0 \leq f_1 \leq K$ on $R$, so also $0 \leq f_k \leq K$. We have

\begin{equation}
(4.119) \int_R f_k \, dV \searrow \alpha,
\end{equation}

for some $\alpha \geq 0$, and we want to show that $\alpha = 0$. Suppose $\alpha > 0$. Pick a partition $\mathcal{P}_k$ of $R$ such that $\mathcal{L}_{\mathcal{P}_k}(f_k) \geq \alpha/2$. Thus $f_k \geq \varphi_k \geq 0$ for some $\varphi_k \in \text{PK}(R)$, constant on the interior of each cell in $\mathcal{P}_k$, with integral $\geq \alpha/2$. The contribution to $\int_R \varphi_k \, dV$ from the cells on which $\varphi_k \leq \alpha/4$ is $\leq \alpha/4$, so the contribution from the cells on which $\varphi_k \geq \alpha/4$ must be $\geq \alpha/4$. Since $\varphi_k \leq K$ on $R$, it follows that the latter class of cells must have total volume $\geq \alpha/4K$. Consequently, for each $k$, there exists $S_k \subset R$, a finite union of cells in $\mathcal{P}_k$, such that

\begin{equation}
(4.120) \quad V(S_k) \geq \frac{\alpha}{4K}, \quad \text{and} \quad f_k \geq \frac{\alpha}{4} \text{ on } S_k.
\end{equation}

Then $f_\ell \geq \alpha/4$ on $S_k$ for all $\ell \leq k$. Hence, with

\begin{equation}
(4.121) \quad \mathcal{O}_\ell = \bigcup_{k \geq \ell} S_k,
\end{equation}

we have

\begin{equation}
(4.122) \quad \text{cont}^-(\mathcal{O}_\ell) \geq \frac{\alpha}{4K}, \quad f_\ell \geq \frac{\alpha}{4} \text{ on } \mathcal{O}_\ell.
\end{equation}

The hypothesis $f_\ell \searrow 0$ on $R$ implies

\begin{equation}
(4.123) \quad \mathcal{O}_\ell \searrow 0 \quad \text{as } \ell \nearrow \infty.
\end{equation}

Without loss of generality, we can take $S_k$ open in (4.120), hence each $\mathcal{O}_\ell$ is open. The conclusion of Proposition 4.26 is hence a consequence of the following, which implies that (4.122) and (4.123) are contradictory.

**Proposition 4.27.** If $\mathcal{O}_\ell \subset R$ are open sets, for $\ell \in \mathbb{N}$, then

\begin{equation}
(4.124) \quad \mathcal{O}_\ell \searrow 0 \quad \Longrightarrow \quad \text{cont}^-(\mathcal{O}_\ell) \searrow 0.
\end{equation}

**Proof.** Assume $\mathcal{O}_\ell \searrow 0$. If the conclusion of (4.124) fails, then

\begin{equation}
(4.125) \quad \text{cont}^-(\mathcal{O}_\ell) \searrow b
\end{equation}

for some $b > 0$. Passing to a subsequence if necessary, we can assume

\begin{equation}
(4.126) \quad \text{cont}^-(\mathcal{O}_\ell) \leq b + \delta_\ell, \quad \delta_\ell < 2^{-\ell} \cdot 10^{-9} \cdot b.
\end{equation}
Then we can pick $K_\ell \subset \mathcal{O}_\ell$, a compact union of finitely many cells in a partition of $R$, such that

$$V(K_\ell) \geq b - \delta_\ell. \tag{4.127}$$

We claim that $\cap_\ell K_\ell \neq \emptyset$, which will provide a contradiction.

Place $K_1 \cup K_2$ in a finite union $C_1$ of cells, contained in $\mathcal{O}_1$. We then have

$$V(K_1 \cap K_2) = V(K_1) - V(C_1 \setminus K_2) \geq b - (2\delta_1 + \delta_2), \tag{4.128}$$

since $V(C_1 \setminus K_2) = V(C_1) - V(K_2) \leq \text{cont}^{-}(\mathcal{O}_1) - V(K_2) \leq \delta_1 + \delta_2$. Next, place $(K_1 \cap K_2) \cup K_3$ in a finite union $C_2$ of cells, contained in $\mathcal{O}_2$. Then

$$V(K_1 \cap K_2 \cap K_3) = V(K_1 \cap K_2) - V(C_2 \setminus K_3) \geq b - (2\delta_1 + \delta_2) - (2\delta_2 + \delta_3), \tag{4.129}$$

since $V(C_2 \setminus K_3) = V(C_2) - V(K_3) \leq \text{cont}^{-}(\mathcal{O}_2) - V(K_3) \leq \delta_2 + \delta_3$. Proceeding in this fashion, we get

$$V\left(\bigcap_{\ell=1}^{k} K_\ell\right) \geq b - \sum_{\ell=1}^{k} (2\delta_\ell + \delta_{\ell+1}) > 0, \quad \forall k. \tag{4.130}$$

Thus, $\tilde{K}_k = \bigcap_{\ell=1}^{k} K_\ell$ is a decreasing sequence of nonempty compact sets. Hence

$$\bigcap_{\ell \geq 1} \mathcal{O}_\ell \supset \bigcap_{\ell \geq 1} K_\ell \neq \emptyset, \tag{4.131}$$

contradicting the hypothesis of (4.124).

Having Proposition 4.26, we proceed to the following significant improvement.

**Proposition 4.28.** Assume $f_k \in \mathcal{R}^\#(R)$. Then

$$f_k(x) \downarrow 0 \quad \forall x \in R \implies \int_{R} f_k \, dV \downarrow 0. \tag{4.132}$$

**Proof.** Again we have (4.119) for some $\alpha \geq 0$ and again we want to show that $\alpha = 0$. For each $A \in (0, \infty)$ and each $k \in \mathbb{N}$, form $(f_k)_A$, as in (4.73). Thus $(f_k)_A \in \mathcal{R}(R)$, and the hypothesis of (4.132) implies $(f_k)_A \downarrow 0$ as $k \nearrow \infty$. Thus, by Proposition 4.26,

$$\int_{R} (f_k)_A \, dV \downarrow 0 \quad \text{as} \quad k \to \infty, \quad \text{for each} \quad A < \infty. \tag{4.133}$$
We note that

\[(4.134)\quad f_{k+1}(x) - (f_{k+1})_A(x) \leq f_k(x) - (f_k)_A(x)\]

for each \(x \in R, \ k \in \mathbb{N}\). In fact, if \(f_k(x) \leq A\) (so \(f_{k+1}(x) \leq A\)), both sides of (4.134) are 0, if \(f_{k+1}(x) \geq A\) (so \(f_k(x) \geq A\)), we get \(f_{k+1}(x) - A \leq f_k(x) - A\), and if \(f_{k+1}(x) < A < f_k(x)\), we get \(0 \leq f_k(x) - A\). It follows that, for each \(A < \infty\),

\[(4.135)\quad \int_R [f_k - (f_k)_A] \, dV \searrow \alpha, \quad \text{as} \quad k \to \infty.\]

However, for each \(\delta > 0\), there exists \(A = A(\delta) < \infty\) such that \(\int_R [f_1 - (f_1)_A] \, dV \leq \delta\). This forces \(\alpha = 0\), and proves Proposition 4.28.

Applying Proposition 4.28 to \(f_k = g - g_k\), we have the following.

**Corollary 4.29.** Assume \(g, g_k \in \mathcal{R}^#(R)\). Then

\[(4.136)\quad g_k(x) \nearrow g(x) \quad \forall \ x \in R \implies \int_R g_k \, dV \nearrow \int_R g \, dV.\]

Finally, we remove the support constraint.

**Proposition 4.30.** Assume \(g, g_k \in \mathcal{R}^#(\mathbb{R}^n)\). Then

\[(4.137)\quad g_k(x) \nearrow g(x) \quad \forall \ x \in \mathbb{R}^n \implies \int_{\mathbb{R}^n} g_k \, dV \nearrow \int_{\mathbb{R}^n} g \, dV.\]

**Proof.** Clearly

\[(4.138)\quad \int_{\mathbb{R}^n} g_k \, dV \nearrow c, \quad \text{and} \quad c \leq \int_{\mathbb{R}^n} g \, dV.\]

Now, given \(\varepsilon > 0\), there is a cell \(R \subset \mathbb{R}^n\) such that

\[(4.139)\quad \int_{\mathbb{R}^n \setminus R} (|g| + |g_1|) \, dV < \varepsilon,\]

and Corollary 4.29 gives

\[(4.140)\quad \int_R g_k \, dV \nearrow \int_R g \, dV.\]
We deduce that \( c \geq \int_{\mathbb{R}^n} g \, dV - \varepsilon \) for all \( \varepsilon > 0 \), so (4.137) holds.

In the Lebesgue theory of integration, there is a stronger result. Namely, if \( g_k \) are integrable on \( \mathbb{R}^n \) and \( g_k(x) \nearrow g(x) \) for each \( x \), and if there is a uniform upper bound \( \int_{\mathbb{R}^n} g_k \, dx \leq B < \infty \), then \( g \) is integrable on \( \mathbb{R}^n \) and the conclusion of (4.137) holds. Such a result can be found in [T2].

**Upper content and outer measure**

Given a bounded set \( S \subset \mathbb{R}^n \), its upper content is defined in (4.13) and an equivalent characterization given in (4.15). A related quantity is the *outer measure* of \( S \), defined by

\[
(4.141) \quad m^*(S) = \inf \left\{ \sum_{k \geq 1} V(R_k) : R_k \subset \mathbb{R}^n \text{ cells}, S \subset \bigcup_{k \geq 1} R_k \right\}.
\]

The difference between (4.15) and (4.141) is that in (4.15) we require the cover of \( S \) by cells to be finite and in (4.141) we allow any *countable* cover of \( S \) by cells. Clearly (4.141) is an inf over a larger collection of objects than (4.15), so

\[
(4.142) \quad m^*(S) \leq \cont^+(S).
\]

We get the same result in (4.141) if we require

\[
(4.143) \quad S \subset \bigcup_{k \geq 1} \overset{\circ}{R_k}
\]

(just expand each \( R_k \) by a factor of \((1 + 2^{-k}\varepsilon)\)). Since any open cover of a compact set has a finite subcover (see Appendix A), it follows that

\[
(4.144) \quad S \text{ compact } \implies m^*(S) = \cont^+(S).
\]

On the other hand, it is readily verified from (4.141) that

\[
(4.145) \quad S \text{ countable } \implies m^*(S) = 0.
\]

For example, if \( R = \{ x \in \mathbb{R}^n : 0 \leq x_j \leq 1, \ \forall j \} \), then

\[
(4.146) \quad m^*(R \cap \mathbb{Q}^n) = 0, \quad \text{but} \quad \cont^+(R \cap \mathbb{Q}^n) = 1,
\]

the latter result by (4.16).

We now establish the following sharpening of Proposition 4.6.
Proposition 4.31. Let \( f : R \to \mathbb{R} \) be bounded, and let \( S \subset R \) be the set of points of discontinuity of \( f \). Then

\[
m^*(S) = 0 \implies f \in \mathcal{R}(R).
\]

Proof. Assume \( |f| \leq M \) and pick \( \varepsilon > 0 \). Take a countable collection \( \{R_k\} \) of cells that are open (in \( R \)) such that \( S \subset \bigcup_{k \geq 1} R_k \) and \( \sum_{k \geq 1} V(R_k) < \varepsilon \). Now \( f \) is continuous at each \( p \in R \setminus S \), so there exists a cell \( R_p^\# \), open (in \( R \)), containing \( p \), such that \( \sup_{R_p^\#} f - \inf_{R_p^\#} f < \varepsilon \).

Then \( \{R_k : k \in \mathbb{N}\} \cup \{R_p^\# : p \in R \setminus S\} \) is an open cover of \( R \). Since \( R \) is compact, there is a finite subcover, which we denote \( \{R_1, \ldots, R_N, R_1^\#, \ldots, R_M^\#\} \). We have

\[
\sum_{k=1}^N V(R_k) < \varepsilon, \quad \text{and} \quad \sup_{R_j^\#} f - \inf_{R_j^\#} f < \varepsilon, \quad \forall j \in \{1, \ldots, M\}.
\]

Recall that \( R = I_1 \times \cdots \times I_n \) is a product of \( n \) closed, bounded intervals. Also each cell \( R_k \) and \( R_j^\# \) is a product of intervals. For each \( \nu \in \{1, \ldots, n\} \), take the collection of all endpoints in the \( \nu \)th factor of each of these cells, and use these to form a partition of \( I_\nu \).

Taking products yields a partition \( \mathcal{P} \) of \( R \). We can write

\[
\mathcal{P} = \{L_k : 1 \leq k \leq \mu\} = \left( \bigcup_{k \in A} L_k \right) \cup \left( \bigcup_{k \in B} L_k \right),
\]

where we say \( k \in A \) provided \( L_k \) is contained in a cell of the form \( R_j^\# \) for some \( j \in \{1, \ldots, M\} \), as in (4.148). Consequently, if \( k \in B \), then \( L_k \subset R_\ell \) for some \( \ell \in \{1, \ldots, N\} \), so

\[
\bigcup_{k \in B} L_k \subset \bigcup_{\ell=1}^N R_\ell.
\]

We therefore have

\[
\sum_{k \in B} V(L_k) < \varepsilon, \quad \text{and} \quad \sup_{L_j} f - \inf_{L_j} f < \varepsilon, \quad \forall j \in A.
\]

It follows that

\[
0 \leq I_\mathcal{P}(f) - I_\mathcal{P}(f) < \sum_{k \in B} 2M V(L_k) + \sum_{j \in A} \varepsilon V(L_j) < 2\varepsilon M + \varepsilon V(R).
\]

Since \( \varepsilon \) can be taken arbitrarily small, this establishes that \( f \in \mathcal{R}(R) \).
Remark. The condition (4.147) is sharp. That is, given \( f : R \to \mathbb{R} \) bounded, \( f \in \mathcal{R}(R) \) \( \iff m^*(S) = 0 \). Proofs of this can be found in standard measure theory texts, such as [T2].

Exercises

1. Show that any two partitions of a cell \( R \) have a common refinement.

   \textit{Hint.} Consider the argument given for the one-dimensional case in §0.

2. Write down a careful proof of the identity (4.16), i.e., \( \text{cont}^+(S) = \text{cont}^+(\overline{S}) \).

3. Write down the details of the argument giving (4.25), on the independence of the integral from the choice of cell.

4. Write down a direct proof that the transformation formula (4.35) holds for those linear transformations of the form \( A_1 \) and \( A_2 \) in (4.36). Compare Exercise 1 of §0.

5. Consider spherical polar coordinates on \( \mathbb{R}^3 \), given by

\[
    x = \rho \sin \varphi \cos \theta, \quad y = \rho \sin \varphi \sin \theta, \quad z = \rho \cos \varphi,
\]

i.e., take \( G(\rho, \varphi, \theta) = (\rho \sin \varphi \cos \theta, \rho \sin \varphi \sin \theta, \rho \cos \theta) \). Show that

\[
    \det DG(\rho, \varphi, \theta) = \rho^2 \sin \varphi.
\]

Use this to compute the volume of the unit ball in \( \mathbb{R}^3 \).

6. If \( B \) is the unit ball in \( \mathbb{R}^3 \), show that Theorem 4.9 implies

\[
    V(B) = 2 \int_D \sqrt{1 - |x|^2} \, dA(x),
\]

where \( D = \{ x \in \mathbb{R}^2 : |x| \leq 1 \} \) is the unit disk. Use polar coordinates, as in (4.58)–(4.61), to compute this integral. Compare the result with that of Exercise 6.

7. Apply Corollary 4.13 and the answer to Exercises 5 and 6 to compute the volume of the ellipsoidal region in \( \mathbb{R}^3 \) defined by

\[
    \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1,
\]

given \( a, b, c \in (0, \infty) \).
8. Prove Lemma 4.16.

9. If $R$ is a cell and $S \subset R$ is a contented set, and $f \in \mathcal{R}(R)$, we have, via Proposition 4.17,

$$\int_S f(x)\,dV(x) = \int_R \chi_S(x)f(x)\,dV(x).$$

Show that, if $S_j \subset R$ are contented and they are disjoint (or more generally $\text{cont}^+(S_1 \cap S_2) = 0$), then, for $f \in \mathcal{R}(R)$,

$$\int_{S_1 \cup S_2} f(x)\,dV(x) = \int_{S_1} f(x)\,dV(x) + \int_{S_2} f(x)\,dV(x).$$

10. Establish the convergence result (4.63), for all $f \in \mathcal{R}(\mathbb{R}^n)$.

11. Take $B = \{x \in \mathbb{R}^n : |x| \leq 1/2\}$, and let $f : B \to \mathbb{R}^+$. Assume $f$ is continuous on $B \setminus 0$. Show that

$$f \in \mathcal{R}^\#(B) \iff \int_{|x| > \varepsilon} f\,dV \text{ is bounded as } \varepsilon \searrow 0.$$

12. With $B \subset \mathbb{R}^n$ as in Exercise 11, define $q_b : B \to \mathbb{R}$ by

$$q_b(x) = \frac{1}{x^n |\log x|^b},$$

for $x \neq 0$. Say $q_b(0) = 0$. Show that $q_b \in \mathcal{R}^\#(B) \iff b > 1$.

13. Show that

$$f(x) = |x|^{-a}e^{-|x|^2} \in \mathcal{R}^\#(\mathbb{R}^n) \iff a < n.$$

14. Theorem 4.9, relating multiple integrals and iterated integrals, played the following role in the proof of the change of variable formula (4.47). Namely, it was used to establish the identity (4.50) for the volume of the parallelipiped $H_\alpha$, via an appeal to Corollary 4.13, hence to Proposition 4.10, whose proof relied on Theorem 4.9.

Try to establish Corollary 4.13 directly, without using Theorem 4.9, in the case when $\Sigma$ is either a cell or the image of a cell under an element of $Gl(n, \mathbb{R})$.

In preparation for the next three exercises, review the proof of Proposition 0.12.

15. Assume $f \in \mathcal{R}(R)$, $|f| \leq M$, and let $\varphi : [-M, M] \to \mathbb{R}$ be Lipschitz and monotone. Show directly from the definition that $\varphi \circ f \in \mathcal{R}(R)$. 

16. If $\varphi : [-M, M] \rightarrow \mathbb{R}$ is continuous and piecewise linear, show that you can write $\varphi = \varphi_1 - \varphi_2$ with $\varphi_j$ Lipschitz and monotone. Deduce that $f \in \mathcal{R}(R) \Rightarrow \varphi \circ f \in \mathcal{R}(R)$ when $\varphi$ is piecewise linear.

17. Assume $u_\nu \in \mathcal{R}(R)$ and that $u_\nu \rightarrow u$ uniformly on $R$. Show that $u \in \mathcal{R}(R)$. Deduce that if $f \in \mathcal{R}(R)$, $|f| \leq M$, and $\psi : [-M, M] \rightarrow \mathbb{R}$ is continuous, then $\psi \circ f \in \mathcal{R}(R)$.

18. Let $R \subset \mathbb{R}^n$ be a cell and let $f, g : R \rightarrow \mathbb{R}$ be bounded. Show that

$$
\bar{T}(f + g) \leq \bar{T}(f) + \bar{T}(g), \quad \underline{I}(f + g) \geq \underline{I}(f) + \underline{I}(g).
$$

**Hint.** Look at the proof of Proposition 0.1.

19. Let $R \subset \mathbb{R}^n$ be a cell and let $f : R \rightarrow \mathbb{R}$ be bounded. Assume that for each $\varepsilon > 0$, there exist bounded $f_\varepsilon, g_\varepsilon$ such that

$$
f = f_\varepsilon + g_\varepsilon, \quad f_\varepsilon \in \mathcal{R}(R), \quad \bar{T}(|g_\varepsilon|) \leq \varepsilon.
$$

Show that $f \in \mathcal{R}(R)$ and

$$
\int_R f_\varepsilon \, dV \rightarrow \int_R f \, dV.
$$

**Hint.** Use Exercise 18.

20. Use the result of Exercise 19 to produce another proof of Proposition 4.6.

21. Behind (4.45) is the assertion that if $R$ is a cell, $g$ is supported on $K \subset R$, and $|g| \leq M$, then $\bar{T}(|g|) \leq M \text{ cont}^+(K)$. Prove this. More generally, if $g, h : R \rightarrow \mathbb{R}$ are bounded and $|g| \leq M$, show that $\bar{T}(|gh|) \leq M\bar{T}(|h|)$.

22. Establish the following Fubini-type theorem, and compare it with Theorem 4.9.

**Proposition 4.9A.** Let $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$ be cells, and take $f \in \mathcal{R}(A \times B)$. For $x \in A$, define $f_x : B \rightarrow \mathbb{R}$ by $f_x(y) = f(x, y)$. Define $L_f, U_f : A \rightarrow \mathbb{R}$ by

$$
L_f(x) = \underline{I}(f_x), \quad U_f(x) = \bar{T}(f_x).
$$

Then $L_f$ and $U_f$ belong to $\mathcal{R}(A)$, and

$$
\int_{A \times B} f \, dV = \int_A L_f(x) \, dx = \int_A U_f(x) \, dx.
$$

**Hint.** Given $\varepsilon > 0$, use Proposition 4.11 to take $\varphi, \psi \in \mathcal{PK}(A \times B)$ such that

$$
\varphi \leq f \leq \psi, \quad \int \psi \, dV - \int \varphi \, dV < \varepsilon.
$$
With definitions of $\varphi_x$ and $\psi_x$ analogous to that of $f_x$, show that

$$\int_{A \times B} \varphi \, dV = \int_A \varphi_x \, dx \leq I(L_f) \leq \mathcal{I}(U_f) \leq \int_A \psi_x \, dx = \int_{A \times B} \psi \, dV.$$ 

Deduce that

$$I(L_f) = \mathcal{I}(U_f),$$

and proceed.

**Exercises on row reduction and matrix products**

We consider the following three types of row operations on an $n \times n$ matrix $A = (a_{jk})$.

If $\sigma$ is a permutation of $\{1, \ldots, n\}$, let

$$\rho_\sigma(A) = (a_{\sigma(j)k}).$$

If $c = (c_1, \ldots, c_j)$, and all $c_j$ are nonzero, set

$$\mu_c(A) = (c_j^{-1}a_{jk}).$$

Finally, if $c \in \mathbb{R}$ and $\mu \neq \nu$, define

$$\epsilon_{\mu \nu c}(A) = (b_{jk}), \quad b_{\nu k} = a_{\nu k} - ca_{\mu k}, \quad b_{jk} = a_{jk} \quad \text{for} \quad j \neq \nu.$$ 

We relate these operations to left multiplication by matrices $P_\sigma, M_c$, and $E_{\mu \nu c}$, defined by the following actions on the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$:

$$P_\sigma e_j = e_{\sigma(j)}, \quad M_c e_j = c_j e_j,$$

and

$$E_{\mu \nu c} e_\mu = e_\mu + c e_\nu, \quad E_{\mu \nu c} e_j = e_j \quad \text{for} \quad j \neq \mu.$$ 

1. Show that

$$A = P_\sigma \rho_\sigma(A), \quad A = M_c \mu_c(A), \quad A = E_{\mu \nu c} \epsilon_{\mu \nu c}(A).$$

2. Show that $P_\sigma^{-1} = P_{\sigma^{-1}}$.

3. Show that, if $\mu \neq \nu$, then $E_{\mu \nu c} = P_\sigma^{-1} E_{21c} P_\sigma$, for some permutation $\sigma$. 
4. If $B = \rho_\alpha(A)$ and $C = \mu_e(B)$, show that $A = P_\alpha M_c C$. Generalize this to other cases where a matrix $C$ is obtained from a matrix $A$ via a sequence of row operations.

5. If $A$ is an invertible, real $n \times n$ matrix (i.e., $A \in Gl(n, \mathbb{R})$), then the rows of $A$ form a basis of $\mathbb{R}^n$. Use this to show that $A$ can be transformed to the identity matrix via a sequence of row operations. Deduce that any $A \in Gl(n, \mathbb{R})$ can be written as a finite product of matrices of the form $P_\alpha$, $M_c$ and $E_{\mu_e}$, hence as a finite product of matrices of the form listed in (4.36).
5. Surfaces and surface integrals

A smooth $m$-dimensional surface $M \subset \mathbb{R}^n$ is characterized by the following property. Given $p \in M$, there is a neighborhood $U$ of $p$ in $M$ and a smooth map $\varphi : \mathcal{O} \to U$, from an open set $\mathcal{O} \subset \mathbb{R}^m$ bijectively to $U$, with injective derivative at each point. Such a map $\varphi$ is called a coordinate chart on $M$. We call $U \subset M$ a coordinate patch. If all such maps $\varphi$ are smooth of class $C^k$, we say $M$ is a surface of class $C^k$. In §8 we will define analogous notions of a $C^k$ surface with boundary, and of a $C^k$ surface with corners.

There is an abstraction of the notion of a surface, namely the notion of a manifold, which we will discuss at the end of this section. Examples include projective spaces and other spaces obtained as quotients of surfaces.

If $\varphi : \mathcal{O} \to U$ is a $C^k$ coordinate chart, such as described above, or more generally $\varphi : \mathcal{O} \to \mathbb{R}^n$ is a $C^k$ map with injective derivative, and $\varphi(x_0) = p$, we set

$$T_pM = \text{Range } D\varphi(x_0),$$

a linear subspace of $\mathbb{R}^n$ of dimension $m$, and we denote by $N_pM$ its orthogonal complement. It is useful to consider the following map. Pick a linear isomorphism $A : \mathbb{R}^{n-m} \to N_pM$, and define

$$\Phi : \mathcal{O} \times \mathbb{R}^{n-m} \to \mathbb{R}^n, \quad \Phi(x, z) = \varphi(x) + Az.$$

Thus $\Phi$ is a $C^k$ map defined on an open subset of $\mathbb{R}^n$. Note that

$$D\Phi(x_0, 0) \begin{pmatrix} v \\ w \end{pmatrix} = D\varphi(x_0)v + Aw,$$

so $D\Phi(x_0, 0) : \mathbb{R}^n \to \mathbb{R}^n$ is surjective, hence bijective, so the Inverse Function Theorem applies; $\Phi$ maps some neighborhood of $(x_0, 0)$ diffeomorphically onto a neighborhood of $p \in \mathbb{R}^n$.

Suppose there is another $C^k$ coordinate chart, $\psi : \Omega \to U$. Since $\varphi$ and $\psi$ are by hypothesis one-to-one and onto, it follows that $F = \psi^{-1} \circ \varphi : \mathcal{O} \to \Omega$ is a well defined map, which is one-to-one and onto. See Fig. 5.1. In fact, we can say more.

**Lemma 5.1.** Under the hypotheses above, $F$ is a $C^k$ diffeomorphism.

**Proof.** It suffices to show that $F$ and $F^{-1}$ are $C^k$ on a neighborhood of $x_0$ and $y_0$, respectively, where $\varphi(x_0) = \psi(y_0) = p$. Let us define a map $\Psi$ in a fashion similar to (5.2). To be precise, we set $\tilde{T}_pM = \text{Range } D\psi(y_0)$, and let $\tilde{N}_pM$ be its orthogonal complement. (Shortly we will show that $\tilde{T}_pM = T_pM$, but we are not quite ready for that.) Then pick a linear isomorphism $B : \mathbb{R}^{n-m} \to \tilde{N}_pM$ and consider

$$\Psi : \Omega \times \mathbb{R}^{n-m} \to \mathbb{R}^n, \quad \Psi(y, z) = \psi(y) + Bz.$$
Again, $\Psi$ is a $C^k$ diffeomorphism from a neighborhood of $(y_0, 0)$ onto a neighborhood of $p$.

It follows that $\Psi^{-1} \circ \Phi$ is a $C^k$ diffeomorphism from a neighborhood of $(x_0, 0)$ onto a neighborhood of $(y_0, 0)$. Now note that, for $x$ close to $x_0$ and $y$ close to $y_0$,

\begin{equation}
\Psi^{-1} \circ \Phi(x, 0) = (F(x), 0), \quad \Phi^{-1} \circ \Psi(y, 0) = (F^{-1}(y), 0).
\end{equation}

These identities imply that $F$ and $F^{-1}$ have the desired regularity.

Thus, when there are two such coordinate charts, $\varphi : \mathcal{O} \to U$, $\psi : \Omega \to U$, we have a $C^k$ diffeomorphism $F : \mathcal{O} \to \Omega$ such that

\begin{equation}
\varphi = \psi \circ F.
\end{equation}

By the chain rule,

\begin{equation}
D\varphi(x) = D\psi(y) \, DF(x), \quad y = F(x).
\end{equation}

In particular this implies that $\text{Range} \, D\varphi(x_0) = \text{Range} \, D\psi(y_0)$, so $T_pM$ in (5.1) is independent of the choice of coordinate chart. It is called the tangent space to $M$ at $p$.

**Remark.** An application of the inverse function theorem related to the proof of Lemma 5.1 can be used to show that if $\mathcal{O} \subset \mathbb{R}^m$ is open, $m < n$, and $\varphi : \mathcal{O} \to \mathbb{R}^n$ is a $C^k$ map such that $D\varphi(p) : \mathbb{R}^m \to \mathbb{R}^n$ is injective, $(p \in \mathcal{O})$, then there is a neighborhood $\tilde{\mathcal{O}}$ of $p$ in $\mathcal{O}$ such that the image of $\tilde{\mathcal{O}}$ under $\varphi$ is a $C^k$ surface in $\mathbb{R}^n$. Compare Exercise 11 in §2.

We next define an object called the **metric tensor** on $M$. Given a coordinate chart $\varphi : \mathcal{O} \to U$, there is associated an $m \times m$ matrix $G(x) = (g_{jk}(x))$ of functions on $\mathcal{O}$, defined in terms of the inner product of vectors tangent to $M$:

\begin{equation}
g_{jk}(x) = D\varphi(x)e_j \cdot D\varphi(x)e_k = \frac{\partial \varphi}{\partial x_j} \cdot \frac{\partial \varphi}{\partial x_k} = \sum_{\ell=1}^{n} \frac{\partial \varphi_{\ell}}{\partial x_j} \frac{\partial \varphi_{\ell}}{\partial x_k},
\end{equation}

where $\{e_j : 1 \leq j \leq m\}$ is the standard orthonormal basis of $\mathbb{R}^m$. Equivalently,

\begin{equation}
G(x) = D\varphi(x)^t \, D\varphi(x).
\end{equation}

We call $(g_{jk})$ the metric tensor of $M$, on $U$, with respect to the coordinate chart $\varphi : \mathcal{O} \to U$. Note that this matrix is positive-definite. From a coordinate-independent point of view, the metric tensor on $M$ specifies inner products of vectors tangent to $M$, using the inner product of $\mathbb{R}^n$.

If we take another coordinate chart $\psi : \Omega \to U$, we want to compare $(g_{jk})$ with $H = (h_{jk})$, given by

\begin{equation}
h_{jk}(y) = D\psi(y)e_j \cdot D\psi(y)e_k, \quad \text{i.e.,} \quad H(y) = D\psi(y)^t \, D\psi(y).
\end{equation}
As seen above we have a diffeomorphism $F : \mathcal{O} \to \Omega$ such that (5.5)–(5.6) hold. Consequently,

$$G(x) = DF(x)^t H(y) DF(x), \quad \text{for } y = F(x),$$

or equivalently,

$$g_{jk}(x) = \sum_{i,\ell} \frac{\partial F_i}{\partial x_j} \frac{\partial F_{\ell}}{\partial x_k} h_{\ell\ell}(y).$$

We now define the notion of surface integral on $M$. If $f : M \to \mathbb{R}$ is a continuous function supported on $U$, we set

$$\int_M f \, dS = \int_{\mathcal{O}} f \circ \varphi(x) \sqrt{g(x)} \, dx,$$

where

$$g(x) = \det G(x).$$

We need to know that this is independent of the choice of coordinate chart $\varphi : \mathcal{O} \to U$. Thus, if we use $\psi : \Omega \to U$ instead, we want to show that (5.12) is equal to $\int_{\Omega} f \circ \psi(y) \sqrt{h(y)} \, dy$, where $h(y) = \det H(y)$. Indeed, since $f \circ \psi \circ F = f \circ \varphi$, we can apply the change of variable formula, Theorem 4.14, to get

$$\int_{\Omega} f \circ \psi(y) \sqrt{h(y)} \, dy = \int_{\mathcal{O}} f \circ \varphi(x) \sqrt{h(F(x))} \, |\det DF(x)| \, dx.$$

Now, (5.10) implies that

$$\sqrt{g(x)} = |\det DF(x)| \sqrt{h(y)},$$

so the right side of (5.14) is seen to be equal to (5.12), and our surface integral is well defined, at least for $f$ supported in a coordinate patch. More generally, if $f : M \to \mathbb{R}$ has compact support, write it as a finite sum of terms, each supported on a coordinate patch, and use (5.12) on each patch. Using (4.11), one readily verifies that

$$\int_M (f_1 + f_2) \, dV = \int_M f_1 \, dV + \int_M f_2 \, dV,$$

if $f_j : M \to \mathbb{R}$ are continuous functions with compact support.

Let us pause to consider the special cases $m = 1$ and $m = 2$. For $m = 1$, we are considering a curve in $\mathbb{R}^n$, say $\varphi : [a, b] \to \mathbb{R}^n$. Then $G(x)$ is a $1 \times 1$ matrix, namely
\(G(x) = |\varphi'(x)|^2\). If we denote the curve in \(\mathbb{R}^n\) by \(\gamma\), rather than \(M\), the formula (5.12) becomes the \textit{arc length} integral

\[(5.16) \quad \int f \, ds = \int_a^b f \circ \varphi(x) \, |\varphi'(x)| \, dx.\]

In case \(m = 2\), let us consider a surface \(M \subset \mathbb{R}^3\), with a coordinate chart \(\varphi : \mathcal{O} \to U \subset M\). For \(f\) supported in \(U\), an alternative way to write the surface integral is

\[(5.17) \quad \int_M f \, dS = \int_{\partial U} f \circ \varphi(x) \, |\partial_1 \varphi \times \partial_2 \varphi| \, dx_1 dx_2,\]

where \(u \times v\) is the cross product of vectors \(u\) and \(v\) in \(\mathbb{R}^3\). To see this, we compare this integrand with the one in (5.12). In this case,

\[(5.18) \quad g = \det \begin{pmatrix} \partial_1 \varphi \cdot \partial_1 \varphi & \partial_1 \varphi \cdot \partial_2 \varphi \\ \partial_2 \varphi \cdot \partial_1 \varphi & \partial_2 \varphi \cdot \partial_2 \varphi \end{pmatrix} = |\partial_1 \varphi|^2 |\partial_2 \varphi|^2 - (\partial_1 \varphi \cdot \partial_2 \varphi)^2.\]

Recall from (1.94) that

\[|u \times v| = |u| |v| \sin \theta,\]

where \(\theta\) is the angle between \(u\) and \(v\). Equivalently, since \(u \cdot v = |u| |v| \cos \theta\),

\[(5.19) \quad |u \times v|^2 = |u|^2 |v|^2 (1 - \cos^2 \theta) = |u|^2 |v|^2 - (u \cdot v)^2.\]

Thus we see that

\[|\partial_1 \varphi \times \partial_2 \varphi| = \sqrt{g},\]

in this case, and (5.17) is equivalent to (5.12).

An important class of surfaces is the class of graphs of smooth functions. Let \(u \in C^1(\Omega)\), for an open \(\Omega \subset \mathbb{R}^{n-1}\), and let \(M\) be the graph of \(z = u(x)\). The map \(\varphi(x) = (x, u(x))\) provides a natural coordinate system, in which the metric tensor formula (5.7) becomes

\[(5.20) \quad g_{jk}(x) = \delta_{jk} + \frac{\partial u}{\partial x_j} \frac{\partial u}{\partial x_k}.\]

If \(u\) is \(C^1\), we see that \(g_{jk}\) is continuous. To calculate \(g = \det(g_{jk})\), at a given point \(p \in \Omega\), if \(\nabla u(p) \neq 0\), rotate coordinates so that \(\nabla u(p)\) is parallel to the \(x_1\) axis. We obtain

\[(5.21) \quad \sqrt{g} = (1 + |\nabla u|^2)^{1/2}.\]

(See Exercise 28 for another take on this formula.) In particular, the \((n - 1)\)-dimensional volume of the surface \(M\) is given by

\[(5.22) \quad V_{n-1}(M) = \int_M dS = \int_{\Omega} (1 + |\nabla u(x)|^2)^{1/2} \, dx.\]

Particularly important examples of surfaces are the unit spheres \(S^{n-1}\) in \(\mathbb{R}^n\),

\[(5.23) \quad S^{n-1} = \{ x \in \mathbb{R}^n : |x| = 1 \}. \]
Spherical polar coordinates on $\mathbb{R}^n$ are defined in terms of a smooth diffeomorphism

\begin{equation}
R : (0, \infty) \times S^{n-1} \longrightarrow \mathbb{R}^n \setminus 0, \quad R(r, \omega) = r\omega.
\end{equation}

Let $(h_{\ell m})$ denote the metric tensor on $S^{n-1}$ (induced from its inclusion in $\mathbb{R}^n$) with respect to some coordinate chart $\varphi : \mathcal{O} \rightarrow U \subset S^{n-1}$. Then we have a coordinate chart $\Phi : (0, \infty) \times \mathcal{O} \rightarrow \mathcal{U} \subset \mathbb{R}^n$ given by $\Phi(r, y) = r\varphi(y)$. Take $y_0 = r$, $y = (y_1, \ldots, y_{n-1})$. In the coordinate system $\Phi$ the Euclidean metric tensor $(e_{jk})$ is given by

\begin{align*}
e_{00} &= \partial_0 \Phi \cdot \partial_0 \Phi = \varphi(y) \cdot \varphi(y) = 1, \\
e_{0j} &= \partial_0 \Phi \cdot \partial_j \Phi = \varphi(y) \cdot \partial_j \varphi(y) = 0, \quad 1 \leq j \leq n-1, \\
e_{jk} &= r^2 \partial_j \varphi \cdot \partial_k \varphi = r^2 h_{jk}, \quad 1 \leq j, k \leq n-1.
\end{align*}

The fact that $\varphi(y) \cdot \partial_j \varphi(y) = 0$ follows by applying $\partial/\partial y_j$ to the identity $\varphi(y) \cdot \varphi(y) = 0$. To summarize,

\begin{equation}
(e_{jk}) = \begin{pmatrix} 1 \\
r^2 h_{\ell m} \end{pmatrix}.
\end{equation}

Now (5.25) yields

\begin{equation}
\sqrt{e} = r^{n-1}\sqrt{h}.
\end{equation}

We therefore have the following result for integrating a function in spherical polar coordinates.

\begin{equation}
\int_{\mathbb{R}^n} f(x) \, dx = \int_{S^{n-1}} \left[ \int_0^\infty f(r\omega) r^{n-1} \, dr \right] dS(\omega).
\end{equation}

We next compute the $(n - 1)$-dimensional area $A_{n-1}$ of the unit sphere $S^{n-1} \subset \mathbb{R}^n$, using (5.27) together with the computation

\begin{equation}
\int_{\mathbb{R}^n} e^{-|x|^2} \, dx = \pi^{n/2},
\end{equation}

from (4.69). First note that, whenever $f(x) = \varphi(|x|)$, (5.27) yields

\begin{equation}
\int_{\mathbb{R}^n} \varphi(|x|) \, dx = A_{n-1} \int_0^\infty \varphi(r) r^{n-1} \, dr.
\end{equation}

In particular, taking $\varphi(r) = e^{-r^2}$ and using (5.28), we have

\begin{equation}
\pi^{n/2} = A_{n-1} \int_0^\infty e^{-r^2} r^{n-1} \, dr = \frac{1}{2} A_{n-1} \int_0^\infty e^{-s^{n/2-1}} \, ds,
\end{equation}
where we used the substitution $s = r^2$ to get the last identity. We hence have

$$A_{n-1} = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}, \tag{5.31}$$

where $\Gamma(z)$ is Euler’s Gamma function, defined for $z > 0$ by

$$\Gamma(z) = \int_0^\infty e^{-s}s^{z-1} \, ds. \tag{5.32}$$

We need to complement (5.31) with some results on $\Gamma(z)$ allowing a computation of $\Gamma(n/2)$ in terms of more familiar quantities. Of course, setting $z = 1$ in (5.32), we immediately get

$$\Gamma(1) = 1. \tag{5.33}$$

Also, setting $n = 1$ in (5.30), we have

$$\pi^{1/2} = 2\int_0^\infty e^{-r^2} \, dr = \int_0^\infty e^{-s}s^{-1/2} \, ds,$$

or

$$\Gamma\left(\frac{1}{2}\right) = \pi^{1/2}. \tag{5.34}$$

We can proceed inductively from (5.33)–(5.34) to a formula for $\Gamma(n/2)$ for any $n \in \mathbb{Z}^+$, using the following.

**Lemma 5.2.** For all $z > 0$,

$$\Gamma(z + 1) = z\Gamma(z). \tag{5.35}$$

**Proof.** We can write

$$\Gamma(z + 1) = -\int_0^\infty \left(\frac{d}{ds}e^{-s}\right)s^z \, ds = \int_0^\infty e^{-s} \frac{d}{ds}(s^z) \, ds,$$

the last identity by integration by parts. The last expression here is seen to equal the right side of (5.35).

Consequently, for $k \in \mathbb{Z}^+$,

$$\Gamma(k) = (k - 1)!, \quad \Gamma\left(k + \frac{1}{2}\right) = \left(k - \frac{1}{2}\right)\cdots \left(\frac{1}{2}\right)\pi^{1/2}. \tag{5.36}$$
Thus (5.31) can be rewritten

\[(5.37) \quad A_{2k-1} = \frac{2\pi^k}{(k-1)!}, \quad A_{2k} = \frac{2\pi^k}{(k-\frac{1}{2}) \cdot \cdots \cdot (\frac{1}{2})}.\]

We discuss another important example of a smooth surface, in the space \(M(n, \mathbb{R}) \approx \mathbb{R}^{n^2}\) of real \(n \times n\) matrices, namely \(SO(n)\), the set of matrices \(T \in M(n, \mathbb{R})\) satisfying \(T^t T = I\) and \(\det T > 0\) (hence \(\det T = 1\)). The exponential map \(\text{Exp}: M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R})\) defined by \(\text{Exp}(A) = e^A\) has the property

\[(5.38) \quad \text{Exp} : \text{Skew}(n) \rightarrow SO(n),\]

where \(\text{Skew}(n)\) is the set of skew-symmetric matrices in \(M(n, \mathbb{R})\). As seen in (2.28)–(2.30),

\[(5.39) \quad D \text{Exp}(0)Y = Y, \quad \forall Y \in M(n, \mathbb{R}),\]

and hence the Inverse Function Theorem implies that there is a ball \(\Omega\) centered at 0 in \(M(n, \mathbb{R})\) that is mapped diffeomorphically by \(\text{Exp}\) onto a neighborhood \(\tilde{\Omega}\) of \(I\) in \(M(n, \mathbb{R})\). From the identities

\(\text{Exp} X^t = (\text{Exp} X)^t, \quad \text{Exp}(-X) = (\text{Exp} X)^{-1},\)

we see that, given \(X \in \Omega, \quad A = \text{Exp} X \in \tilde{\Omega},\)

\(A \in SO(n) \iff X \in \text{Skew}(n).\)

Thus there is a neighborhood \(\mathcal{O}\) of 0 in \(\text{Skew}(n)\) that is mapped by \(\text{Exp}\) diffeomorphically onto a smooth surface \(U \subset M(n, \mathbb{R})\), of dimension \(m = n(n-1)/2\). Furthermore, \(U\) is a neighborhood of \(I\) in \(SO(n)\). For general \(T \in SO(n)\), we can define maps

\[(5.40) \quad \varphi_T : \mathcal{O} \rightarrow SO(n), \quad \varphi_T(A) = T \text{Exp}(A),\]

and obtain coordinate charts in \(SO(n)\), which is consequently a smooth surface of dimension \(n(n-1)/2\) in \(M(n, \mathbb{R})\). Note that \(SO(n)\) is a closed bounded subset of \(M(n, \mathbb{R})\); hence it is compact.

We use the inner product on \(M(n, \mathbb{R})\) computed componentwise; equivalently,

\[(5.41) \quad \langle A, B \rangle = \text{Tr}(B^t A) = \text{Tr}(BA^t).\]

This produces a metric tensor on \(SO(n)\). The surface integral over \(SO(n)\) has the following important invariance property.
Proposition 5.3. Given \( f \in C(SO(n)) \), if we set

\[
\rho_T f(X) = f(XT), \quad \lambda_T f(X) = f(TX),
\]

for \( T, X \in SO(n) \), we have

\[
\int_{SO(n)} \rho_T f \, dS = \int_{SO(n)} \lambda_T f \, dS = \int_{SO(n)} f \, dS.
\]

Proof. Given \( T \in SO(n) \), the maps \( R_T, L_T : M(n, \mathbb{R}) \to M(n, \mathbb{R}) \) defined by \( R_T(X) = XT \), \( L_T(X) = TX \) are easily seen from (5.41) to be isometries. Thus they yield maps of \( SO(n) \) to itself which preserve the metric tensor, proving (5.43).

Since \( SO(n) \) is compact, its total volume \( V(SO(n)) = \int_{SO(n)} 1 \, dS \) is finite. We define the integral with respect to “Haar measure”

\[
\int_{SO(n)} f(g) \, dg = \frac{1}{V(SO(n))} \int_{SO(n)} f \, dS.
\]

This is used in many arguments involving “averaging over rotations.” Examples of such averaging arise in §§10 and 12.

Extended notion of coordinates

Basic calculus as developed in this text so far has involved maps from one Euclidean space to another, of the type \( F : \mathbb{R}^n \to \mathbb{R}^m \). It is convenient and useful to extend our setting to \( F : V \to W \), where \( V \) and \( W \) are general finite-dimensional real vector spaces. There is the following notion of the derivative.

Let \( V \) and \( W \) be as above, and let \( \Omega \subset V \) be open. We say \( F : \Omega \to W \) is differentiable at \( x \in \Omega \) provided there exists a linear map \( L : V \to W \) such that, for \( y \in V \) small,

\[
F(x + y) = F(x) + Ly + r(x, y),
\]

with \( r(x, y) \to 0 \) faster than \( y \to 0 \), i.e.,

\[
\frac{\|r(x, y)\|}{\|y\|} \to 0 \quad \text{as} \quad y \to 0.
\]

For this to be meaningful, we need norms on \( V \) and \( W \). Often these norms come from inner products. See Appendix H for a discussion of inner product spaces. If (5.45)–(5.46) hold, we set \( DF(x) = L \), and call the linear map

\[
DF(x) : V \to W
\]
the derivative of \( F \) at \( x \). We say \( F \) is \( C^1 \) if \( DF(x) \) is continuous in \( x \). Notions of \( F \) in \( C^k \) are produced in analogy with the situation in §1. Of course, we can reduce all this to the setting of §1 by picking bases of \( V \) and \( W \).

Often such \( V \) and \( W \) arise as linear subspaces of \( \mathbb{R}^n \), such as \( T_p M \) in (5.1), or \( V = N_p M \), mentioned right below that. As noted there, we can take a linear isomorphism of such \( V \) with \( \mathbb{R}^k \) for some \( k \), and keep working in the context of maps between such standard Euclidean spaces, as in (5.2). However, it can be convenient to avoid this distraction, and, for example, replace (5.2) by

\[
\Phi : \mathcal{O} \times N_p M \longrightarrow \mathbb{R}^n, \quad \Phi(x, z) = \varphi(x) + z,
\]

and (5.3) by

\[
D\Phi(x_0, 0)\begin{pmatrix} v \\ w \end{pmatrix} = D\varphi(x_0)v + w.
\]

In order to carry out Lemma 5.1 in this setting, we want the following version of the Inverse Function Theorem.

**Proposition 5.4.** Let \( V \) and \( W \) be real vector spaces, each of dimension \( n \). Let \( F \) be a \( C^k \) map from an open neighborhood \( \Omega \) of \( p_0 \in V \) to \( W \), with \( q_0 = F(p_0) \), \( k \geq 1 \). Assume the derivative

\[
DF(p_0) : V \rightarrow W \text{ is an isomorphism.}
\]

Then there exist a neighborhood \( U \) of \( p_0 \) and a neighborhood \( \tilde{U} \) of \( q_0 \) such that \( F : U \rightarrow \tilde{U} \) is one-to-one and onto, and \( F^{-1} : \tilde{U} \rightarrow U \) is a \( C^k \) map.

While Proposition 5.4 is apparently an extension of Theorem 2.1, there is no extra work required to prove it. One can simply take linear isomorphisms \( A : \mathbb{R}^n \rightarrow V \) and \( B : \mathbb{R}^n \rightarrow W \) and apply Theorem 2.1 to the map \( G(x) = B^{-1}F(Ax) \). Thus Proposition 5.4 is not a technical improvement of Theorem 2.1, but it is a useful conceptual extension.

With this in mind, we can define the notion of an \( m \)-dimensional surface \( M \subset V \) (an \( n \)-dimensional vector space) as follows. Take a vector space \( W \), of dimension \( m \). Given \( p \in M \), we require there to be a neighborhood \( U \) of \( p \) in \( M \) and a smooth map \( \varphi : \mathcal{O} \rightarrow U \), from an open set \( \mathcal{O} \subset W \) bijectively to \( U \), with an injective derivative at each point. We call such a map a coordinate chart. If all such maps are smooth of class \( C^k \), we say \( M \) is a surface of class \( C^k \). As a further wrinkle, we could take different vector spaces \( W_p \) for different \( p \in M \), as long as they all have dimension \( m \). The reader is invited to formulate the appropriate modification of Lemma 5.1 in this setting.

**Submersions**

Let \( V \) and \( W \) be finite dimensional real vector spaces, \( \Omega \subset V \) open, and \( F : \Omega \rightarrow W \) a \( C^k \) map, \( k \geq 1 \). We say \( F \) is a submersion provided that, for each \( x \in \Omega \), \( DF(x) : V \rightarrow W \) is surjective. (This requires \( \dim V \geq \dim W \).) We establish the following Submersion
Mapping Theorem, which the reader might recognize as a variant of the Implicit Function Theorem. In the statement, \( \ker T \) denotes the null space
\[
\ker T = \{ v \in V :Tv = 0 \},
\]
if \( T : V \to W \) is a linear transformation.

**Proposition 5.5.** With \( V, W, \) and \( \Omega \subset V \) as above, assume \( F : \Omega \to W \) is a \( C^k \) map, \( k \geq 1 \). Fix \( p \in W \), and consider
\[
S = \{ x \in V : F(x) = p \}.
\]
Assume that, for each \( x \in S \), \( DF(x) : V \to W \) is surjective. Then \( S \) is a \( C^k \) surface in \( \Omega \). Furthermore, for each \( x \in S \),
\[
T_xS = \ker DF(x).
\]

**Proof.** Given \( q \in S \), set \( K_q = \ker DF(q) \) and define
\[
G_q : V \to W \oplus K_q, \quad G_q(x) = (F(x), P_q(x - q)),
\]
where \( P_q \) is a projection of \( V \) onto \( K_q \). Note that
\[
G_q(q) = (F(q), 0) = (p, 0).
\]
Also
\[
DG_q(x) = (DF(x), P_q), \quad x \in V.
\]
We claim that
\[
DG_q(q) = (DF(q), P_q) : V \to W \oplus K_q \text{ is an isomorphism.}
\]
This is a special case of the following general observation.

**Lemma 5.6.** If \( A : V \to W \) is a surjective linear map and \( P \) is a projection of \( V \) onto \( \ker A \), then
\[
(A, P) : V \to W \oplus \ker A \text{ is an isomorphism.}
\]

We postpone the proof of this lemma and proceed with the proof of Proposition 5.5. Having (5.54), we can apply the Inverse Function Theorem (Proposition 5.4) to obtain a neighborhood \( U \) of \( q \) in \( V \) and a neighborhood \( \mathcal{O} \) of \( (p, 0) \) in \( W \oplus K_q \) such that \( G_q : U \to \mathcal{O} \) is bijective, with \( C^k \) inverse
\[
G_q^{-1} : \mathcal{O} \to U, \quad G_q^{-1}(p, 0) = q.
\]
By (5.51), given $x \in U$,

\[(5.57) \quad x \in S \iff G_q(x) = (p, v), \text{ for some } v \in K_q.\]

Hence $S \cap U$ is the image under the $C^k$ diffeomorphism $G_q^{-1}$ of $O \cap \{(p, v) : v \in K_q\}$. Hence $S$ is smooth of class $C^k$ and $\dim T_q S = \dim K_q$. It follows from the chain rule that $T_q S \subset K_q$, so the dimension count yields $T_q S = K_q$. This proves Proposition 5.5. Note that we have the following coordinate chart on a neighborhood of $q \in S$:

\[(5.58) \quad \psi_q(v) = G_q^{-1}(p, v), \quad \psi_q : \Omega_q \to S,\]

where $\Omega_q$ is a neighborhood of $0$ in $T_q S = K_q = \ker DF(q)$.

It remains to prove Lemma 5.6. Indeed, given that $A : V \to W$ is surjective, the fundamental theorem of linear algebra implies $\dim V = \dim(W \oplus \ker A)$, and it is clear that $(A, P)$ in (5.55) is injective, so the isomorphism property follows.

**Remark.** In case $V = \mathbb{R}^n$ and $W = \mathbb{R}$, $DF(x)$ is typically denoted $\nabla F(x)$, the hypothesis on $DF(x)$ becomes $\nabla F(x) \neq 0$, and (5.50) is equivalent to the assertion that $\dim S = n - 1$ and, for $x \in S$,

\[(5.59) \quad \nabla F(x) \perp T_x S.\]

Compare the discussion following Proposition 2.6.

We illustrate Proposition 5.5 with another proof that

\[(5.60) \quad SO(n) \subset M(n, \mathbb{R})\]

is a smooth surface, different from the argument involving (5.38)–(5.40). To get this, we take

\[(5.61) \quad V = M(n, \mathbb{R}), \quad W = \{A \in M(n, \mathbb{R}) : A = A^t\},\]

and

\[(5.62) \quad F : V \to W, \quad F(X) = X^t X.\]

Now, given $X, Y \in V$, $Y$ small,

\[(5.63) \quad F(X + Y) = X^t X + X^t Y + Y^t X + O(\|Y\|^2),\]

so

\[(5.64) \quad DF(X)Y = X^t Y + Y^t X.\]
We claim that
\[(5.65) \quad X \in SO(n) \implies DF(X) : M(n, \mathbb{R}) \to W \text{ is surjective.}\]
Indeed, given \(A \in W\), i.e., \(A \in M(n, \mathbb{R})\) and \(A^t = A\), and \(X \in SO(n)\), we have
\[(5.66) \quad Y = \frac{1}{2} X A \implies DF(X)Y = A.\]
This establishes (5.65), so Proposition 5.5 applies. Again we conclude that \(SO(n)\) is a smooth surface in \(M(n, \mathbb{R})\).

**Riemann integrable functions on a surface**

Let \(M \subset \mathbb{R}^n\) be an \(m\)-dimensional surface, smooth of class \(C^1\). We define the class \(\mathcal{R}_c(M)\) of compactly supported Riemann integrable functions as follows, guided by Proposition 4.11. If \(f : M \to \mathbb{R}\) is bounded and has compact support, we set
\[(5.67) \quad \bar{T}(f) = \inf \left\{ \int_M g \, dS : g \in C_c(M), \, g \geq f \right\},\]
\[(5.68) \quad \bar{I}(f) = \sup \left\{ \int_M h \, dS : h \in C_c(M), \, h \leq f \right\},\]
where \(C_c(M)\) denotes the set of continuous functions on \(M\) with compact support. Then
\[(5.69) \quad f \in \mathcal{R}_c(M) \iff \bar{T}(f) = \bar{I}(f),\]
and if such is the case, we denote the common value by \(\int_M f \, dS\). It follows readily from the definition and arguments produced in §4 that
\[(5.70) \quad f_1, f_2 \in \mathcal{R}_c(M) \implies f_1 + f_2 \in \mathcal{R}_c(M) \quad \text{and} \quad \int_M (f_1 + f_2) \, dS = \int_M f_1 \, dS + \int_M f_2 \, dS.\]
In fact, using (5.15A) for functions that are continuous on \(M\) with compact support, one obtains from the definition (5.67) that, if \(f_j : M \to \mathbb{R}\) are bounded and have compact support,
\[(5.71) \quad \bar{T}(f_1 + f_2) \leq \bar{T}(f_1) + \bar{T}(f_2), \quad \bar{I}(f_1 + f_2) \geq \bar{I}(f_1) + \bar{I}(f_2),\]
which yields (5.69). Also one can modify the proof of Proposition 4.17 to show that
\[(5.72) \quad f \in \mathcal{R}_c(M), \, u \in C(M) \implies uf \in \mathcal{R}_c(M).\]
Furthermore, if \(\varphi : \mathcal{O} \to U \subset M\) is a coordinate chart and \(f \in \mathcal{R}_c(U)\), then an application of Proposition 4.11 gives
\[(5.73) \quad f \circ \varphi \in \mathcal{R}_c(\mathcal{O}), \quad \text{and} \quad \int_M f \, dS = \int_{\mathcal{O}} f(\varphi(x)) \sqrt{g(x)} \, dx,\]
with \( g(x) \) as in (5.12)-(5.13). Given any \( f \in \mathcal{R}_c(M) \), we can take a continuous partition of unity \( \{u_j\} \), write \( f = \sum_j f_j = \sum_j u_j f \), and use (5.69)-(5.71) to express \( \int_M f \, dS \) as a sum of integrals over coordinate charts.

If \( \Sigma \subset M \) has compact closure, then

\[
\text{cont}^+ \Sigma = \mathcal{I}(\chi_{\Sigma}),
\]

and \( \Sigma \) is contented if and only if \( \chi_{\Sigma} \in \mathcal{R}_c(M) \). In such a case, (5.72) is the area of \( \Sigma \). Given \( f : M \to \mathbb{R} \), bounded and compactly supported, in parallel with (4.39) we say

\[
f \in \mathcal{C}_c(M) \Longleftrightarrow \text{the set } \Sigma \text{ of points of discontinuity of } f \text{ satisfies } \text{cont}^+ \Sigma = 0.\]

We have

\[
\mathcal{C}_c(M) \subset \mathcal{R}_c(M),
\]

and (again parallel to Proposition 4.11) if \( f : M \to \mathbb{R} \) is bounded and compactly supported,

\[
\mathcal{I}(f) = \inf \left\{ \int_M g \, dS : g \in \mathcal{C}_c(M), \ g \geq f \right\},
\]

\[
\mathcal{I}(f) = \sup \left\{ \int_M h \, dS : h \in \mathcal{C}_c(M), \ h \leq f \right\}.\]

One can proceed from here to define the spaces

\[
\mathcal{R}(M), \quad \mathcal{R}^\#(M),
\]

and establish properties of functions in these spaces, in analogy with work in §4 on \( \mathcal{R}(\mathbb{R}^n) \) and \( \mathcal{R}^\#(\mathbb{R}^n) \). We leave such an investigation to the reader.

**Vector fields and flows on surfaces**

Let \( M \subset \mathbb{R}^n \) be a smooth, \( m \)-dimensional surface. A smooth vector field \( X \) on \( M \) (sometimes called a tangent vector field) is a smooth map

\[
X : M \longrightarrow \mathbb{R}^n \text{ such that } X(p) \in T_p M, \ \forall p \in M.
\]

If \( \varphi : \mathcal{O} \to U \subset M \) is a coordinate chart, then there is a unique smooth vector field \( X_{\varphi} : \mathcal{O} \to \mathbb{R}^m \) such that

\[
X(\varphi(x)) = D\varphi(x)X_{\varphi}(x).
\]
The vector field $X$ generates a flow $\mathcal{F}_X^t$ on $M$, satisfying
\begin{equation}
(5.79) \quad \mathcal{F}_X^t(\varphi(x)) = \varphi(\mathcal{F}_X^t(x)),
\end{equation}
at least for small $|t|$. As before, we have the defining property
\begin{equation}
(5.80) \quad \frac{d}{dt} \mathcal{F}_X^t(x) = X(\mathcal{F}_X^t(x)).
\end{equation}
Note also that
\begin{equation}
(5.81) \quad \mathcal{F}_X^{s+t}(x) = \mathcal{F}_X^s \circ \mathcal{F}_X^t(x).
\end{equation}
With slight abuse of notation we will use the same symbol $X$ for the vector field $X$ on $M$ and for the associated vector field $X_\varphi$ on a coordinate patch.

Valuable information on the behavior of the flow $\mathcal{F}_X^t$ can be obtained by investigating the $t$-derivative of
\begin{equation}
(5.82) \quad v_t(x) = v(\mathcal{F}_X^t(x)),
\end{equation}
given $v \in C^1_0(M)$, i.e., $v$ is of class $C^1$ and vanishes outside some compact subset of $M$. In fact, we take $v \in C^1_0(U)$, and identify this with $v \in C^1_0(\mathcal{O})$, with $\mathcal{O} \subset \mathbb{R}^m$ as above. The chain rule plus (5.80) yields
\begin{equation}
(5.83) \quad \frac{d}{dt} v_t(x) = X(\mathcal{F}_X^t(x)) \cdot \nabla v(\mathcal{F}_X^t(x)),
\end{equation}
In particular,
\begin{equation}
(5.84) \quad \frac{d}{ds} v(\mathcal{F}_X^s(x))|_{s=0} = X(x) \cdot \nabla v(x).
\end{equation}
Here $\nabla v$ is the gradient of $v$, given by $\nabla v = (\partial v/\partial x_1, \ldots, \partial v/\partial x_m)$. A useful alternative formula to (5.83) is
\begin{equation}
(5.85) \quad \frac{d}{dt} v_t(x) = \frac{d}{ds} v_t(\mathcal{F}_X^s(x))|_{s=0} = X(x) \cdot \nabla v_t(x),
\end{equation}
the first equality following from (5.81) and the second from (5.84), with $v$ replaced by $v_t$.

One significant consequence of (5.85), which will lead to the formula (5.90) below, is that, for $v \in C^1_0(\mathcal{O})$,
\begin{equation}
(5.86) \quad \frac{d}{dt} \int_\mathcal{O} v(\mathcal{F}_X^t(x)) \sqrt{g} \, dx = \int_\mathcal{O} X(x) \cdot \nabla v_t(x) \sqrt{g} \, dx = -\int_\mathcal{O} \text{div} X(x)v(\mathcal{F}_X^t(x)) \sqrt{g} \, dx.
\end{equation}
Here, \( \text{div} \, X(x) \) is the **divergence** of the vector field \( X(x) = (X_1(x), \ldots, X_m(x)) \), defined (in local coordinates) by

\[
(5.87) \quad \text{div} \, X(x) = g^{-1/2} \sum_j \frac{\partial}{\partial x_j} (g^{1/2} X_j(x)).
\]

The last equality in (5.86) follows by integration by parts,

\[
\int_G G_k(x) \frac{\partial v_l}{\partial x_k} dx = - \int_G \frac{\partial G_k}{\partial x_k} v_l(x) dx, \quad G_k(x) = \sqrt{g} X_k(x),
\]

followed by summation over \( k \). We restate (5.86) in global terms:

\[
(5.88) \quad \frac{d}{dt} \int_M v(F^t_X(x)) \, dV = - \int_M \text{div} \, X(x) v(F^t_X(x)) \, dV.
\]

So far, we have (5.88) for \( v \in C^1_0(M) \). We extend this to less regular functions. First, note that (5.88) implies

\[
(5.89) \quad \int_M v(F^t_X(x)) \, dV - \int_M v(x) \, dV = - \int_0^t \int_M \text{div} \, X(x) v(F^s_X(x)) \, dV \, ds.
\]

Basic results on the integral allow one to pass from \( v \in C^1_0(M) \) in (5.57) to more general \( v \), including \( v = \chi_\Omega \) (the characteristic function of \( \Omega \), defined to be equal to 1 on \( \Omega \) and 0 on \( M \setminus \Omega \)), for smoothly bounded compact \( \Omega \subset M \).

In more detail, if \( \overline{\Omega} \subset M \) is a compact, smoothly bounded subset, let \( B_\delta = \{ x \in M : \text{dist}(x, \overline{\Omega}) \leq \delta \} \). There exists \( \delta_0 > 0 \) such that \( B_\delta \subset M \) for \( \delta \in (0, \delta_0] \). For such \( \delta \), one can produce \( v_\delta \in C^1_0(M) \) such that

\[
v_\delta = 1 \text{ on } B_\delta, \quad 0 \leq v_\delta \leq 1, \quad v_\delta = 0 \text{ on } M \setminus B_\delta.
\]

Then

\[
\left| \int_M \chi_\Omega(x) \, dV - \int_M v_\delta(x) \, dV \right| \leq \text{Vol}(B_\delta \setminus \Omega) \to 0, \text{ as } \delta \to 0,
\]

so, as \( \delta \to 0 \),

\[
\int_M v_\delta(x) \, dV \to \int_M \chi_\Omega(x) \, dV.
\]

Similar arguments give

\[
\int_M v_\delta(F^t_X(x)) \, dV \to \int_M \chi_\Omega(F^t_X(x)) \, dV,
\]
and
\[ \int_0^t \int_M \text{div} X(x) v_\delta (F^s_X(x)) \, dV \, ds \to \int_0^t \int_M \text{div} X(x) \chi_\Omega (F^s_X(x)) \, dV \, ds. \]

These results allow one to take \( v = \chi_\Omega \) in (5.89).

Now one can pass from (5.89) back to (5.88), via the fundamental theorem of calculus. Note that
\[ \text{Vol} F^t_X(\Omega) = \int_M \chi_\Omega (F^{-t}_X(x)) \, dV. \]

We can apply (5.88), with \( t \) replaced by \(-t\), and \( v \) by \( \chi_\Omega \), and deduce the following.

**Proposition 5.7.** If \( X \) is a \( C^1 \) vector field, generating the flow \( F^t_X \), well defined on \( M \) for \( t \in I \), and \( \overline{\Omega} \subset M \) is compact and smoothly bounded, then, for \( t \in I \),
\[ \frac{d}{dt} \text{Vol} F^t_X(\Omega) = \int_{F^t_X(\Omega)} \text{div} X(x) \, dV. \]

This result is behind the notation \( \text{div} X \), i.e., the divergence of \( X \). Vector fields with positive divergence generate flows \( F^t_X \) that magnify volumes as \( t \) increases, while vector fields with negative divergence generate flows that shrink volumes as \( t \) increases. We will see more of the divergence operation on vector fields in Sections 9 and 11.

**Projective spaces, quotient surfaces, and manifolds**

Real projective space \( \mathbb{P}^{n-1} \) is obtained from the sphere \( S^{n-1} \) by identifying each pair of antipodal points:
\[ \mathbb{P}^{n-1} = S^{n-1} / \sim, \]
where
\[ x \sim y \iff x = \pm y, \]
for \( x, y \in S^{n-1} \subset \mathbb{R}^n \). More generally, if \( M \subset \mathbb{R}^n \) is an \( m \)-dimensional surface, smooth of class \( C^k \), satisfying
\[ 0 \notin M, \quad x \in M \implies -x \in M, \]
we define
\[ \mathbb{P}(M) = M / \sim, \]
using the equivalence relation (5.92). Note that \( M \) has the metric space structure \( d(x, y) = \|x - y\| \), and then \( \mathbb{P}(M) \) becomes a metric space with metric
\[ d([x], [y]) = \min \{d(x', y') : x' \in [x], y' \in [y]\}, \]
or, in view of (5.92),

\[(5.96) \quad d([x], [y]) = \min\{d(x, y), d(x, -y)\}.\]

Here, \(x \in M\) and \([x] \in \mathbb{P}(M)\) is its associated equivalence class. The map \(x \mapsto [x]\) is a continuous map

\[(5.97) \quad \rho : M \longrightarrow \mathbb{P}(M).\]

It has the following readily established property.

**Lemma 5.8.** Each \(p \in \mathbb{P}(M)\) has an open neighborhood \(U \subset \mathbb{P}(M)\) such that \(\rho^{-1}(U) = U_0 \cup U_1\) is the disjoint union of two open subsets of \(M\), and, for \(j = 0, 1\), \(\rho : U_j \rightarrow U\) is a homeomorphism, i.e., it is continuous, one-to-one, and onto, with continuous inverse.

Given \(p \in \mathbb{P}(M)\), \(\{p_0, p_1\} = \rho^{-1}(p)\), let \(U_0\) be a neighborhood of \(p_0\) in \(M\) for which there is a \(C^k\) coordinate chart \(\varphi_0 : \mathcal{O} \rightarrow U_0\) (\(\mathcal{O} \subset \mathbb{R}^m\) open). Then \(\varphi_1(x) = -\varphi_0(x)\) gives a coordinate chart \(\varphi_1 : \mathcal{O} \rightarrow U_1\) onto a neighborhood \(U_1\) of \(p_1 \in M\). If \(U_0\) is picked small enough, \(U_0\) and \(U_1\) are disjoint. The projection \(\rho\) maps \(U_0\) and \(U_1\) homeomorphically onto a neighborhood \(U\) of \(p\) in \(\mathbb{P}(M)\), and we have “coordinate charts”

\[(5.98) \quad \rho \circ \varphi_j : \mathcal{O} \longrightarrow U.\]

In fact, \(\rho \circ \varphi_1 = \rho \circ \varphi_0\). If \(\psi_0 : \Omega \rightarrow U_0\) is another \(C^k\) coordinate chart, then, as in Lemma 5.1, we have a \(C^k\) diffeomorphism \(F : \mathcal{O} \rightarrow \Omega\), \(\psi_1(x) = -\psi_0(x)\), and we have \(\rho \circ \psi_1 \circ F = \rho \circ \varphi_j\).

The structure just placed on the “quotient surface” \(\mathbb{P}(M)\) makes it a manifold, an object we now define.

Given a metric space \(X\), we say \(X\) has the structure of a \(C^k\) manifold of dimension \(m\) provided the following conditions hold. First, for each \(p \in X\), we have an open neighborhood \(U_p\) of \(p\) in \(X\), an open set \(\mathcal{O}_p \subset \mathbb{R}^m\), and a homeomorphism

\[(5.99) \quad \varphi_p : \mathcal{O}_p \longrightarrow U_p.\]

Next, if also \(q \in X\) and \(U_{pq} = U_p \cap U_q \neq \emptyset\), then the homeomorphism from \(\mathcal{O}_{pq} = \varphi_p^{-1}(U_{pq})\) to \(\mathcal{O}_{qp} = \varphi_q^{-1}(U_{pq})\),

\[(5.100) \quad F_{pq} = \varphi_q^{-1} \circ \varphi_p|_{\mathcal{O}_{pq}},\]

is a \(C^k\) diffeomorphism. As before, we call the maps \(\varphi_p : \mathcal{O}_p \rightarrow U_p \subset X\) coordinate charts on \(X\).

A metric tensor on a \(C^k\) manifold \(X\) is defined by positive-definite, symmetric \(m \times m\) matrices \(G_p \in C^{k-1}(\mathcal{O}_p)\), satisfying the compatibility condition

\[(5.101) \quad G_p(x) = DF_{pq}(x)^t G_q(y) D F_{pq}(x),\]
for
\begin{equation}
(5.102) \quad x \in \mathcal{O}_{pq} \subset \mathcal{O}_p, \quad y = F_{pq}(x) \in \mathcal{O}_{ap} \subset \mathcal{O}_q.
\end{equation}
We then set
\begin{equation}
(5.103) \quad g_p = \det G_p \in C^{k-1}(\mathcal{O}_p),
\end{equation}
satisfying
\begin{equation}
(5.104) \quad \sqrt{g_p(x)} = |\det DF_{pq}(x)| \sqrt{g_q(y)},
\end{equation}
for \(x\) and \(y\) as in (5.102). If \(f : X \to \mathbb{R}\) is a continuous function supported in \(U_p\), we set
\begin{equation}
(5.105) \quad \int_X f \, dS = \int_{\mathcal{O}_p} f(\varphi_p(x)) \sqrt{g_p(x)} \, dx.
\end{equation}
As in (5.14)–(5.15), this leads to a well defined integral \(\int_X f \, dS\) for \(f \in C_c(X)\), obtained by writing \(f\) as a finite sum of continuous functions supported on various coordinate patches \(U_p\). From here we can develop the class of functions \(\mathcal{R}_c(X)\) and their integrals over \(X\), in a fashion parallel to that done above when \(X\) is a surface in \(\mathbb{R}^n\).

The quotient surfaces \(\mathbb{P}(M)\) are examples of \(C^k\) manifolds as defined above. They get natural metric tensors with the property that \(\rho\) in (5.97) is a local isometry. In such a case,
\begin{equation}
(5.106) \quad \int_{\mathbb{P}(M)} f \, dS = \frac{1}{2} \int_M f \circ \rho \, dS.
\end{equation}
Another important quotient manifold is the “flat torus”
\begin{equation}
(5.107) \quad \mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n.
\end{equation}
Here the equivalence relation on \(\mathbb{R}^n\) is \(x \sim y \Leftrightarrow x - y \in \mathbb{Z}^n\). Natural local coordinates on \(\mathbb{T}^n\) are given by the projection \(\rho : \mathbb{R}^n \to \mathbb{T}^n\), restricted to sufficiently small open sets in \(\mathbb{R}^n\). The quotient \(\mathbb{T}^n\) gets a natural metric tensor for which \(\rho\) is a local isometry.

Given two \(C^k\) manifolds \(X\) and \(Y\), a continuous map \(\psi : X \to Y\) is said to be smooth of class \(C^k\) provided that for each \(p \in X\), there are neighborhoods \(U\) of \(p\) and \(\tilde{U}\) of \(q = \psi(p)\), and coordinate charts \(\varphi_1 : \mathcal{O} \to U\), \(\varphi_2 : \tilde{\mathcal{O}} \to \tilde{U}\), such that \(\varphi_2^{-1} \circ \psi \circ \varphi_1 : \mathcal{O} \to \tilde{\mathcal{O}}\) is a \(C^k\) map. We say \(\psi\) is a \(C^k\) diffeomorphism if it is one-to-one and onto and \(\psi^{-1} : Y \to X\) is a \(C^k\) map. If \(X\) is a \(C^k\) manifold and \(M \subset \mathbb{R}^n\) a \(C^k\) surface, a \(C^k\) diffeomorphism \(\psi : X \to M\) is called a \(C^k\) embedding of \(X\) into \(\mathbb{R}^n\).

Here is an embedding of \(\mathbb{T}^n\) into \(\mathbb{R}^{2n}\):
\begin{equation}
(5.108) \quad \psi(x) = \sum_{j=1}^n (\cos 2\pi x_j)e_j + \sum_{j=1}^n (\sin 2\pi x_j)e_{n+j}.
\end{equation}
A priori, \( \psi : \mathbb{R}^n \to \mathbb{R}^{2n} \), but \( \psi(x) = \psi(y) \) whenever \( x - y \in \mathbb{Z}^n \), so this naturally induces a smooth map \( T^n \to \mathbb{R}^{2n} \), which can be seen to be an embedding.

If \( M \subset \mathbb{R}^n \) is an \( m \)-dimensional surface satisfying (5.93), an embedding of \( \mathbb{P}(M) \) into \( M(n, \mathbb{R}) \) can be constructed via the map

\[
\psi : \mathbb{R}^n \to M(n, \mathbb{R}), \quad \psi(x) = xx^t.
\]  

Note that

\[
\begin{pmatrix}
\vdots \\
x_1 \\
\vdots \\
x_n
\end{pmatrix}
\implies
xx^t =
\begin{pmatrix}
x_1^2 & \cdots & x_1x_j & \cdots & x_1x_n \\
\vdots & & \vdots & & \vdots \\
x_nx_1 & \cdots & x_nx_j & \cdots & x_n^2
\end{pmatrix}.
\]

We need a couple of lemmas.

**Lemma 5.9.** For \( \psi \) as in (5.109), \( x, y \in \mathbb{R}^n \),

\[
\psi(x) = \psi(y) \iff x = \pm y.
\]

**Proof.** The map \( \psi \) is characterized by \( \psi(x)e_j = x_jx \), where \( x \) is as in (5.110) and \( \{e_j\} \) is the standard basis of \( \mathbb{R}^n \). It follows that if \( x \neq 0 \), \( \psi(x) \) has exactly one nonzero eigenvalue, namely \( |x|^2 \), and \( \psi(x)x = |x|^2x \). Thus \( \psi(x) = \psi(y) \) implies that \( |x|^2 = |y|^2 \) and that \( x \) and \( y \) are parallel. Thus \( x = ay \) and \( a = \pm 1 \).

**Lemma 5.10.** In the setting of Lemma 5.9, if \( x \neq 0 \),

\[
D\psi(x) : \mathbb{R}^n \to M(n, \mathbb{R}) \text{ is injective.}
\]

**Proof.** A calculation gives

\[
D\psi(x)v = xv^t + vx^t.
\]

Thus, if \( v \in \ker D\psi(x) \),

\[
xv^t = -vx^t.
\]

Both sides are rank 1 elements of \( M(n, \mathbb{R}) \). The range of the left side is spanned by \( x \) and that of the right side is spanned by \( v \), so \( v = ax \) for some \( a \in \mathbb{R} \). Then (5.114) becomes

\[
ax^t = -ax^t,
\]

which implies \( a = 0 \) if \( x \neq 0 \).

**Remark.** Here is a refinement of Lemma 5.10. Using the inner product on \( M(n, \mathbb{R}) \) given by (5.41), we can calculate

\[
\langle D\psi(x)v, D\psi(x)v \rangle = 2(|x|^2|v|^2 + (x \cdot v)^2).
\]
Lemmas 5.9 and 5.10 imply that if $M \subset \mathbb{R}^n$ is an $m$-dimensional surface satisfying (5.93), then $\psi|_M$ yields an embedding of $\mathbb{P}(M)$ into $M(n, \mathbb{R})$. Denote the image surface by $M^\#$. As we see from (5.116), this embedding is not typically an isometry. However, if $M = S^{n-1}$ and $v$ is tangent to $S^{n-1}$ at $x$, then $v \cdot x = 0$, and (5.116) implies that in this case the embedding of $\mathbb{P}^{n-1}$ into $M(n, \mathbb{R})$ is an isometry, up to a factor of 2.

It is the case that if $X$ is any $C^k$ manifold that is a countable union of compact sets, then $X$ can be embedded into $\mathbb{R}^n$ for some $n$. In case $X$ is compact, this is not very hard to prove, using local coordinate charts and smooth cutoffs, and the interested reader might take a crack at it. If $X$ is provided with a metric tensor, this embedding might not preserve this metric tensor. If it does, one calls it an isometric embedding. It is the case that any such manifold has an isometric embedding into $\mathbb{R}^n$ for some $n$ (if $k$ is sufficiently large). This result is the famous Nash embedding theorem, and its proof is quite difficult. For $X$ compact and $C^\infty$, a proof is given in Chapter 14 of [T].

**Polar decomposition of matrices**

We define the spaces $\text{Sym}(n)$ and $\mathcal{P}(n)$ by

\begin{align*}
\text{Sym}(n) &= \{A \in M(n, \mathbb{R}) : A = A^t\}, \\
\mathcal{P}(n) &= \{A \in \text{Sym}(n) : x \cdot Ax > 0, \ \forall x \in \mathbb{R}^n \setminus \{0\}\}.
\end{align*}

It is easy to show that $\mathcal{P}(n)$ is an open, convex subset of the linear space $\text{Sym}(n)$. We aim to prove the following result.

**Proposition 5.11.** Given $A \in GL_+(n, \mathbb{R})$, there exist unique $U \in SO(n)$ and $Q \in \mathcal{P}(n)$ such that

\begin{equation}
A = UQ.
\end{equation}

The representation (5.118) is called the polar decomposition of $A$. Note that

\begin{equation}
(UQ)^tUQ = QU^tUQ = Q^2,
\end{equation}

so if the identity (5.118) were to hold, we would have

\begin{equation}
A^tA = Q^2.
\end{equation}

Note also that

\begin{equation}
A \in GL(n, \mathbb{R}) \implies A^tA \in \mathcal{P}(n),
\end{equation}

since $x \cdot A^tAx = (Ax) \cdot (Ax) = |Ax|^2$.

To prove Proposition 5.11, we bring in the following basic result of linear algebra. See, e.g., §11 of [T7].
Proposition 5.12. Given $B \in \text{Sym}(n)$, there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B$, with eigenvalues $\lambda_j \in \mathbb{R}$. Equivalently, there exists $V \in SO(n)$ such that

$$B = VDV^{-1},$$

with

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

$\lambda_j \in \mathbb{R}$.

If $B \in \mathcal{P}(n)$, then each $\lambda_j > 0$. We can then set

$$Q = V \begin{pmatrix} \lambda_1^{1/2} & & \\ & \ddots & \\ & & \lambda_n^{1/2} \end{pmatrix} V^{-1},$$

and obtain the following.

Corollary 5.13. Given $B \in \mathcal{P}(n)$, there is a unique $Q \in \mathcal{P}(n)$ satisfying

$$Q^2 = B.$$

We say $Q = B^{1/2}$.

To obtain the decomposition (5.118), we set

$$Q = (A^tA)^{1/2}, \quad U = AQ^{-1}.$$

Note that

$$U^tU = Q^{-1}A^tAQ^{-1} = Q^{-1}Q^2Q^{-1} = I,$$

and $(\det U)(\det Q) = \det A > 0$, so $\det U > 0$, and hence $U \in SO(n)$, as desired. By (5.120) and Corollary 5.13, the factor $Q \in \mathcal{P}(n)$ in (5.118) is unique, and hence so is the factor $U$.

We can use Proposition 5.11 to prove the following.

Proposition 5.14. The set $Gl_+(n, \mathbb{R})$ is connected. In fact, given $A \in Gl_+(n, \mathbb{R})$, there is a smooth path $\gamma : [0, 1] \rightarrow Gl_+(n, \mathbb{R})$ such that $\gamma(0) = I$ and $\gamma(1) = A$.

Proof. To start, we have that

$$\exp : \text{Skew}(n) \rightarrow SO(n) \text{ is onto.}$$
See Exercise 14 below for this. Hence, with \( A = UQ \) as in (5.118), we have a smooth path \( \alpha(t) = \text{Exp}(tS) \), \( \alpha : [0, 1] \rightarrow SO(n) \), such that \( \alpha(0) = I \) and \( \alpha(1) = U \). Since \( P(n) \) is a convex subset of \( \text{Sym}(n) \), we can take \( \beta(t) = (1 - t)I + tQ \), obtaining a smooth path \( \beta : [0, 1] \rightarrow P(n) \), such that \( \beta(0) = I \) and \( \beta(1) = Q \). Then

\[
\gamma(t) = \alpha(t)\beta(t)
\]

(5.129)

does the trick.

**Exercises**

1. Define \( \varphi : [0, \theta] \rightarrow \mathbb{R}^2 \) to be \( \varphi(t) = (\cos t, \sin t) \). Show that, if \( 0 < \theta \leq 2\pi \), the image of \([0, \theta]\) under \( \varphi \) is an arc of the unit circle, of length \( \theta \). Deduce that the unit circle in \( \mathbb{R}^2 \) has total length \( 2\pi \). This result follows also from (5.37).

**Remark.** Use the definition of \( \pi \) given in the auxiliary problem set after §3.

This length formula provided the original definition of \( \pi \), in ancient Greek geometry.

2. Compute the volume of the unit ball \( B^n = \{ x \in \mathbb{R}^n : |x| \leq 1 \} \).

**Hint.** Apply (5.29) with \( \varphi = \chi_{[0,1]} \).

3. Taking the upper half of the sphere \( S^n \) to be the graph of \( x_{n+1} = (1 - |x|^2)^{1/2} \), for \( x \in B^n \), the unit ball in \( \mathbb{R}^n \), deduce from (5.22) and (5.29) that

\[
A_n = 2A_{n-1}\int_0^1 \frac{r^{n-1}}{\sqrt{1-r^2}} \, dr = 2A_{n-1}\int_0^{\pi/2} (\sin \theta)^{n-1} \, d\theta.
\]

Use this to get an alternative derivation of the formula (5.37) for \( A_n \).

**Hint.** Rewrite this formula as

\[
A_n = A_{n-1}b_{n-1}, \quad b_k = \int_0^\pi \sin^k \theta \, d\theta.
\]

To analyze \( b_k \), you can write, on the one hand,

\[
b_{k+2} = b_k - \int_0^\pi \sin^k \theta \cos^2 \theta \, d\theta,
\]

and on the other, using integration by parts,

\[
b_{k+2} = \int_0^\pi \cos \theta \frac{d}{d\theta} \sin^{k+1} \theta \, d\theta.
\]

Deduce that

\[
b_{k+2} = \frac{k+1}{k+2} b_k.
\]
4. Suppose $M$ is a surface in $\mathbb{R}^n$ of dimension 2, and $\varphi : \mathcal{O} \to U \subset M$ is a coordinate chart, with $\mathcal{O} \subset \mathbb{R}^2$. Set $\varphi_{jk}(x) = (\varphi_j(x), \varphi_k(x))$, so $\varphi_{jk} : \mathcal{O} \to \mathbb{R}^2$. Show that the formula (5.12) for the surface integral is equivalent to

$$\int_M f \, dS = \int_\mathcal{O} f \circ \varphi(x) \sqrt{\sum_{j<k} (\det D\varphi_{jk}(x))^2} \, dx.$$

*Hint.* Show that the quantity under $\sqrt{\text{—}}$ is equal to (5.18).

5. If $M$ is an $m$-dimensional surface, $\varphi : \mathcal{O} \to M \subset M$ a coordinate chart, for $J = (j_1, \ldots, j_m)$ set

$$\varphi_J(x) = (\varphi_{j_1}(x), \ldots, \varphi_{j_m}(x)), \quad \varphi_J : \mathcal{O} \to \mathbb{R}^m.$$

Show that the formula (5.12) is equivalent to

$$\int_M f \, dS = \int_\mathcal{O} f \circ \varphi(x) \sqrt{\sum_{j_1 < \cdots < j_m} (\det D\varphi_J(x))^2} \, dx.$$

*Hint.* Reduce to the following. For fixed $x_0 \in \mathcal{O}$, the quantity under $\sqrt{\text{—}}$ is equal to $g(x)$ at $x = x_0$, in the case $D\varphi(x_0) = (D\varphi_1(x_0), \ldots, D\varphi_m(x_0), 0, \ldots, 0)$. Reconsider this problem when working on the exercises for §6.

6. Let $M$ be the graph in $\mathbb{R}^{n+1}$ of $x_{n+1} = u(x)$, $x \in \mathcal{O} \subset \mathbb{R}^n$. Show that, for $p = (x, u(x)) \in M$, $T_pM$ (given as in (5.1)) has a 1-dimensional orthogonal complement $N_pM$, spanned by $(-\nabla u(x), 1)$. We set $N = (1 + |\nabla u|^2)^{-1/2}(-\nabla u, 1)$, and call it the (upward-pointing) unit normal to $M$.

7. Let $M$ be as in Exercise 6, and define $N$ as done there. Show that, for a continuous function $f : M \to \mathbb{R}^{n+1},$

$$\int_M f \cdot N \, dS = \int_\mathcal{O} f(x, u(x)) \cdot (-\nabla u(x), 1) \, dx.$$

The left side is often denoted $\int_M f \cdot dS$.

8. Let $M$ be a 2-dimensional surface in $\mathbb{R}^3$, covered by a single coordinate chart, $\varphi : \mathcal{O} \to M$. Suppose $f : M \to \mathbb{R}^3$ is continuous. Show that, if $\int_M f \cdot dS$ is defined as in Exercise 7, then

$$\int_M f \cdot dS = \int_\mathcal{O} f(\varphi(x)) \cdot (\partial_1 \varphi \times \partial_2 \varphi) \, dx.$$

9. Consider a symmetric $n \times n$ matrix $A = (a_{jk})$ of the form $a_{jk} = v_j v_k$. Show that the
range of $A$ is the one-dimensional space spanned by $v = (v_1, \ldots, v_n)$ (if this is nonzero). Deduce that $A$ has exactly one nonzero eigenvalue, namely $\lambda = |v|^2$. Use this to give another derivation of (5.21) from (5.20).

**Hint.** Show that $Ae_j = v_jv$, for each $j$.

10. Let $\Omega \subset \mathbb{R}^n$ be open and $u : \Omega \to \mathbb{R}$ be a $C^k$ map. Fix $c \in \mathbb{R}$ and consider

$$ S = \{ x \in \Omega : u(x) = c \}. $$

Assume $S \neq \emptyset$ and that $\nabla u(x) \neq 0$ for all $x \in S$.

As seen after Proposition 5.6, $S$ is a $C^k$ surface of dimension $n - 1$, and, for each $p \in S$, $T_pS$ has a 1-dimensional orthogonal complement $N_pS$ spanned by $\nabla u(p)$. Assume now that there is a $C^k$ map $\varphi : O \to \mathbb{R}$, with $O \subset \mathbb{R}^{n-1}$ open, such that $u(x', \varphi(x')) = c$, and that $x' \mapsto (x', \varphi(x'))$ parametrizes $S$. Show that

$$ \int_S f dS = \int_O f \frac{|\nabla u|}{|\partial_n u|} \, dx', $$

where the functions in the integrand on the right are evaluated at $(x', \varphi(x'))$.

**Hint.** Compare the formula in Exercise 6 for $N$ with the fact that $N = \nabla u/|\nabla u|$, and keep in mind the formula (5.22).

In the next exercises, we study $\exp tJ = e^{tJ}$, where

$$ J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. $$

11. Show that if $v \in \mathbb{R}^2$, then

$$ \frac{d}{dt} \|e^{tJ}v\|^2 = 2e^{tJ}v \cdot J e^{tJ}v = 0, $$

and deduce that $\|e^{tJ}v\| = \|v\|$ for all $v \in \mathbb{R}^2$, $t \in \mathbb{R}$.

12. Define $c(t)$ and $s(t)$ by

$$ e^{tJ} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c(t) \\ s(t) \end{pmatrix}. $$

Show that the identity $(d/dt)e^{tJ} = Je^{tJ}$ implies

$$ c'(t) = -s(t), \quad s'(t) = c(t). $$

Deduce that $(c(t), s(t))$ is a unit speed curve, starting at $(c(0), s(0)) = (1, 0)$, with initial velocity $(c'(0), s'(0)) = (0, 1)$, and tracing out the unit circle $x^2 + y^2 = 1$ in $\mathbb{R}^2$. See Fig. 5.2. Compare the derivation of (3.94)–(3.96).
13. Using Exercise 12 and (5.16), show that for \( t > 0 \), the curve \( \gamma : [0,t] \rightarrow \mathbb{R}^2 \) given by \( \gamma(t) = (c(t),s(t)) \) has length \( t \). In trigonometry, the line segments from \((0,0)\) to \((1,0)\) and from \((0,0)\) to \((c(t),s(t))\) are said to meet at an angle, measured in radians, equal to the length of this curve, i.e., to \( t \) radians. Then the geometric definitions of the trigonometric functions \( \cos t \) and \( \sin t \) yield

\[
\cos t = c(t), \quad \sin t = s(t). \tag{5.130}
\]

Deduce that

\[
e^{tJ} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \tag{5.131}
\]

and from this, using \( e^{tJ}J = Je^{tJ} \), that

\[
e^{tJ} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \tag{5.132}
\]

Compare (3.107). However, the derivation of (5.132) here is completely independent of the one in §3, which used (3.91)–(3.92). It shows that the definition of \( \cos t \) and \( \sin t \) given here is equivalent to that given in §3, and again leads to Euler’s formula (3.92).

14. The following is a basic result of linear algebra. See, e.g., §12 of [T7].

**Proposition.** If \( A : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is orthogonal, so \( A^t A = I \), then \( \mathbb{R}^n \) has an orthonormal basis in which the matrix representation of \( A \) consists of blocks

\[
\begin{pmatrix}
c_j & -s_j \\
s_j & c_j
\end{pmatrix}, \quad c_j^2 + s_j^2 = 1,
\]

plus perhaps an identity matrix block if 1 is an eigenvalue of \( A \), and a block that is \(-I\) if \(-1\) is an eigenvalue of \( A \).

Use this and (5.132) to prove that

\[
\text{Exp} : \text{Skew}(n) \rightarrow SO(n) \quad \text{is onto.} \tag{5.133}
\]

In the next exercise, \( T \) denotes the “inner tube” obtained as follows. Take the circle in the \((y,z)\)-plane, centered at \( y = a, z = 0 \), of radius \( b \), with \( 0 < b < a \). Rotate this circle about the \( z \)-axis. Then \( T \) is the surface so swept out. See Fig. 5.3.

15. Define \( \psi : \mathbb{R}^2 \rightarrow \mathbb{R}^3 \) by \( \psi(\theta, \varphi) = (x(\theta, \varphi), y(\theta, \varphi), z(\theta, \varphi)) \), with

\[
\begin{align*}
x(\theta, \varphi) &= (a + b \cos \varphi) \cos \theta, \\
y(\theta, \varphi) &= (a + b \cos \varphi) \sin \theta, \\
z(\theta, \varphi) &= b \sin \varphi.
\end{align*}
\]
Show that $\psi$ maps $[0, 2\pi] \times [0, 2\pi]$ onto $T$.

Show that $|\partial_\theta \psi \times \partial_\phi \psi| = b(a + b \cos \varphi)$.

Using (5.17), show that

$$\text{Area } T = 4\pi^2 ab.$$  

16. In the setting of Exercise 15, compute the following integrals.

$$\int_T x^2 \, dS, \quad \int_T y^2 \, dS, \quad \int_T z^2 \, dS.$$  

In the next exercise, $M$ is a surface of revolution, obtained by taking the graph of a function $y = f(x), \ 0 \leq x \leq b$ (assuming $f > 0$) and rotating it about the $x$-axis, in $\mathbb{R}^3$.

17. Define $\psi : [a, b] \times \mathbb{R} \to \mathbb{R}^3$ by $\psi(s, t) = (s, f(s) \cos t, f(s) \sin t)$.

Show that $\psi$ maps $[a, b] \times [0, 2\pi]$ onto $M$.

Show that $|\partial_s \psi \times \partial_t \psi| = f(s) \sqrt{1 + f'(s)^2}$.

Using (5.17), show that if $u : M \to \mathbb{R}$ is continuous,

$$\int_M u \, dS = \int_0^{2\pi} \int_a^b u(s, f(s) \cos t, f(s) \sin t) f(s) \sqrt{1 + f'(s)^2} \, ds \, dt.$$  

In particular,

$$\text{Area } M = 2\pi \int_a^b f(s) \sqrt{1 + f'(s)^2} \, ds.$$  

18. Consider the ellipsoid of revolution $E_a$, given for $a > 0$ by

$$\frac{x^2}{a^2} + y^2 + z^2 = 1.$$  

Use the method of Exercise 17 to show that

$$\text{Area } E_a = 4\pi \int_0^a \sqrt{1 - \beta s^2} \, ds, \quad \beta = \frac{1}{a^2} - \frac{1}{a^4}.$$  

19. Given $a, b, c > 0$, consider the ellipsoid $E(a, b, c)$, given by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$  

Using (5.22), write down a formula for the area of $E(a, b, c)$ as an integral over the region

$$E_{a,b} = \left\{(x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}.$$
In the next exercise, $M$ is an $n$-dimensional “surface of revolution” given by a smooth map

$$\psi : [a, b] \times S^{n-1} \rightarrow M \subset \mathbb{R} \times \mathbb{R}^n$$

of the form

$$\psi(s, \omega) = (s, f(s)\omega).$$

A coordinate chart $\varphi : \Omega \rightarrow S^{n-1}$, with $\Omega$ open in $\mathbb{R}^{n-1}$, gives rise to a coordinate chart

$$\tilde{\psi} : [a, b] \times \Omega \rightarrow M, \quad \tilde{\psi}(s, x) = (s, f(s)\varphi(x)).$$

We set $x_0 = s$, $x = (x_1, \ldots, x_{n-1})$.

20. Show that, in $\tilde{\psi}$-coordinates, the metric tensor of $M$ takes the form $(g_{jk})$, for $0 \leq j, k \leq n - 1$, with the following components:

$$g_{00} = \frac{\partial \tilde{\psi}}{\partial x_0} \cdot \frac{\partial \tilde{\psi}}{\partial x_0} = 1 + f'(s)^2,$$

$$g_{0j} = \frac{\partial \tilde{\psi}}{\partial x_0} \cdot \frac{\partial \tilde{\psi}}{\partial x_j} = 0, \quad \text{for } 1 \leq j \leq n - 1,$$

$$g_{jk} = \frac{\partial \tilde{\psi}}{\partial x_j} \cdot \frac{\partial \tilde{\psi}}{\partial x_k} = f(s)^2 h_{jk}, \quad \text{for } 1 \leq j, k \leq n - 1,$$

where $(h_{\ell m})$ is the metric tensor of $S^{n-1}$ in the $\varphi$-coordinates. Otherwise said,

$$(g_{jk}) = \begin{pmatrix} 1 + f'(s)^2 & f(s)^2 h_{\ell m} \\ f(s)^2 h_{\ell m} & f(s)^2 h_{\ell m} \end{pmatrix}.$$

Compare (5.25).

21. In the setting of Exercise 20, deduce that if $u : M \rightarrow R$ is continuous,

$$\int_M u \, dS = \int_a^b \int_{S^{n-1}} u(s, f(s)\omega) f(s)^{n-1} \sqrt{1 + f'(s)^2} \, dS(\omega) \, ds.$$

In particular, with $A_{n-1}$ as in (5.29)–(5.31),

$$\text{Area } M = A_{n-1} \int_a^b f(s)^{n-1} \sqrt{1 + f'(s)^2} \, ds.$$

Note how this generalizes the conclusion of Exercise 17.
21A. In the setting of Exercises 20–21, let $M = S^n$, with $f(s) = \sqrt{1 - s^2}$. Show that
\[
\int_{S^n} u(x_0) dS(x) = A_{n-1} \int_{-1}^{1} u(s)(1 - s^2)^{(n-2)/2} ds.
\]

22. Let $\psi : SO(n) \to M(k, \mathbb{R})$ be continuous and satisfy the following properties:
\[
\psi(gh) = \psi(g)\psi(h), \quad \psi(g^{-1}) = \psi(g)^{-1},
\]
for all $g, h \in SO(n)$. We say $\psi$ is a representation of $SO(n)$ on $\mathbb{R}^k$. Form
\[
P = \int_{SO(n)} \psi(g) dg, \quad P \in M(k, \mathbb{R}),
\]
using the integral (5.44) (but here with a matrix valued integrand). Show that
\[
P : \mathbb{R}^k \to V, \text{ and } v \in V \Rightarrow P v = v,
\]
where
\[
V = \{v \in \mathbb{R}^k : \psi(g)v = v, \forall g \in SO(n)\}.
\]
Thus $P$ is a projection of $\mathbb{R}^k$ onto $V$.
*Hint.* With $h \in SO(n)$ arbitrary, express $\psi(h)P$ as the integral $\int \psi(hg) dg$, and apply (5.43).

23. In the setting of Exercise 22, show that
\[
\dim V = \int_{SO(n)} \chi(g) dg, \quad \chi(g) = \text{Tr } \psi(g).
\]

24. Given $u \in C(\mathbb{R}^n)$, define $Au \in C(\mathbb{R}^n)$ by
\[
Au(x) = \int_{SO(n)} u(gx) dg.
\]
Show that $Au$ is a radial function, in fact
\[
Au(x) = Su(|x|), \quad \text{with } Su(r) = \frac{1}{A_{n-1}} \int_{S^{n-1}} u(r\omega) dS(\omega).
\]

25. In the setting of Exercise 24, show that if $u(x)$ is a polynomial in $x$, then $Su(r)$ is a
polynomial in $r$.

**Hint.** Show that $Au(x)$ is a polynomial in $x$. Look at $Au(re_1)$.

26. Let $M$ be a $C^1$ surface, $K \subset M$ compact. Let $\varphi_j : O_j \to U_j$ be coordinate charts on $M$ and assume $K \subset \bigcup_{j=1}^k U_j$. Take $v_j \in C_c(U_j)$ such that $\sum_1^k v_j > 0$ on $K$.

Let $f : M \to \mathbb{R}$ be bounded and supported on $K$. Show that $f \in \mathcal{R}_c(M) \iff (v_j f) \circ \varphi_j \in \mathcal{R}_c(O_j)$, for each $j$.

Here, use the definition (5.67)-(5.68) for $\mathcal{R}_c(M)$ and define $\mathcal{R}_c(O_j)$ as in §4.

27. Let $M \subset \mathbb{R}^n$ be a compact, $m$-dimensional, $C^1$ surface. We define a contented partition of $M$ to be a finite collection $\mathcal{P} = \{\Sigma_k\}$ of contented subsets of $M$ such that

$$M = \bigcup_k \Sigma_k, \quad \text{cont}^+ (\Sigma_j \cap \Sigma_k) = 0, \quad \forall j \neq k.$$ 

We say

$$\text{maxsize } \mathcal{P} = \max_k \text{diam} (\Sigma_k),$$

where $\text{diam} \Sigma_k = \sup_{x,y \in \Sigma_k} \|x - y\|$. Establish the following variant of the Darboux theorem (Proposition 4.1).

**Proposition.** Let $\mathcal{P}_\nu = \{\Sigma_{k\nu} : 1 \leq k \leq N(\nu)\}$ be a sequence of contented partitions of $M$ such that $\text{maxsize } \mathcal{P}_\nu \to 0$. Pick points $\xi_{k\nu} \in \Sigma_{k\nu}$. Then, given $f \in \mathcal{R}(M)$, we have

$$\int_M f \, dS = \lim_{\nu \to \infty} \sum_{k=1}^{N(\nu)} f(\xi_{k\nu}) V(\Sigma_{k\nu}),$$

where $V(\Sigma_{k\nu}) = \int_M \chi_{\Sigma_{k\nu}} \, dS$ is the content of $\Sigma_{k\nu}$.

**Hint.** First treat the case $f \in C(M)$. Use the material in (5.67)-(5.75) to extend this to $f \in \mathcal{R}(M)$.

28. We desire to compute $\det G$ when $G = (g_{jk})$ is an $m \times m$ matrix given by

$$g_{jk} = \delta_{jk} + v_j v_k.$$ 

Compare (5.20). In other words,

$$G = I + T, \quad T = (t_{jk}), \quad t_{jk} = v_j v_k.$$

(a) Let $v \in \mathbb{R}^m$ have components $v_j$. Show that, for $w \in \mathbb{R}^m$, $Tw = (v \cdot w)v$.

(b) Deduce that $T$ has one nonzero eigenvalue, $|v|^2$.

(c) Deduce that one eigenvalue of $G$ is $1 + |v|^2$, and the other $m - 1$ eigenvalues are $1$.

(d) Deduce that $g = \det G = 1 + |v|^2$, so $\sqrt{g} = \sqrt{1 + |v|^2}$. Compare (5.21), with $v = \nabla u$. 


6. Differential forms

It is very desirable to be able to make constructions that depend as little as possible on
a particular choice of coordinate system. The calculus of differential forms, whose study
we now take up, is one convenient set of tools for this purpose.

We start with the notion of a 1-form. It is an object that gets integrated over a curve;
formally, a 1-form on $\Omega \subset \mathbb{R}^n$ is written

$$\alpha = \sum_j a_j(x) \, dx_j.$$  \hspace{1cm} (6.1)

If $\gamma : [a, b] \to \Omega$ is a smooth curve, we set

$$\int_\gamma \alpha = \int_a^b \sum_j a_j(\gamma(t)) \gamma_j'(t) \, dt.$$  \hspace{1cm} (6.2)

In other words,

$$\int_\gamma \alpha = \int_I \gamma^* \alpha$$  \hspace{1cm} (6.3)

where $I = [a, b]$ and

$$\gamma^* \alpha = \sum_j a_j(\gamma(t)) \gamma_j'(t) \, dt$$

is the pull-back of $\alpha$ under the map $\gamma$. More generally, if $F : \mathcal{O} \to \Omega$ is a smooth map
($\mathcal{O} \subset \mathbb{R}^m$ open), the pull-back $F^* \alpha$ is a 1-form on $\mathcal{O}$ defined by

$$F^* \alpha = \sum_{j,k} a_j(F(y)) \frac{\partial F_j}{\partial y_k} \, dy_k.$$  \hspace{1cm} (6.4)

The usual change of variable for integrals gives

$$\int_\gamma \alpha = \int_{F \circ \sigma} F^* \alpha$$  \hspace{1cm} (6.5)

if $\gamma$ is the curve $F \circ \sigma$.

If $F : \mathcal{O} \to \Omega$ is a diffeomorphism, and

$$X = \sum b^j(x) \frac{\partial}{\partial x_j}$$  \hspace{1cm} (6.6)
is a vector field on $\Omega$, recall from (3.40) that we have the vector field on $\mathcal{O}$:

$$F_\# X(y) = (DF^{-1}(p))X(p), \quad p = F(y).$$

If we define a pairing between 1-forms and vector fields on $\Omega$ by

$$\langle X, \alpha \rangle = \sum_j b^j(x)a_j(x) = b \cdot a,$$

a simple calculation gives

$$\langle F_\# X, F^* \alpha \rangle = \langle X, \alpha \rangle \circ F.$$

Thus, a 1-form on $\Omega$ is characterized at each point $p \in \Omega$ as a linear transformation of the space of vectors at $p$ to $\mathbb{R}$.

More generally, we can regard a $k$-form $\alpha$ on $\Omega$ as a $k$-multilinear map on vector fields:

$$\alpha(X_1, \ldots, X_k) \in C^\infty(\Omega);$$

we impose the further condition of anti-symmetry when $k \geq 2$:

$$\alpha(X_1, \ldots, X_j, \ldots, X_\ell, \ldots, X_k) = -\alpha(X_1, \ldots, X_\ell, \ldots, X_j, \ldots, X_k).$$

Let us note that a 0-form is simply a function.

There is a special notation we use for $k$-forms. If $1 \leq j_1 < \cdots < j_k \leq n$, $j = (j_1, \ldots, j_k)$, we set

$$\alpha = \frac{1}{k!} \sum_j a_j(x) \, dx_{j_1} \wedge \cdots \wedge dx_{j_k}$$

where

$$a_j(x) = \alpha(D_{j_1}, \ldots, D_{j_k}), \quad D_j = \partial/\partial x_j.$$

More generally, we assign meaning to (6.12) summed over all $k$-indices $(j_1, \ldots, j_k)$, where we identify

$$dx_{j_1} \wedge \cdots \wedge dx_{j_k} = (\text{sgn } \sigma) \, dx_{j_{\sigma(1)}} \wedge \cdots \wedge dx_{j_{\sigma(k)}},$$

$\sigma$ being a permutation of $\{1, \ldots, k\}$. If any $j_m = j_\ell$ ($m \neq \ell$), then (6.14) vanishes. A common notation for the statement that $\alpha$ is a $k$-form on $\Omega$ is

$$\alpha \in \Lambda^k(\Omega).$$
In particular, we can write a 2-form $\beta$ as

\begin{equation}
\beta = \frac{1}{2} \sum b_{jk}(x) \, dx_j \wedge dx_k
\end{equation}

and pick coefficients satisfying $b_{jk}(x) = -b_{kj}(x)$. According to (6.12)–(6.13), if we set $U = \sum u_j(x) \partial / \partial x_j$ and $V = \sum v_j(x) \partial / \partial x_j$, then

\begin{equation}
\beta(U, V) = \sum b_{jk}(x) u^j(x) v^k(x).
\end{equation}

If $b_{jk}$ is not required to be antisymmetric, one gets

\begin{equation}
\beta(U, V) = (1/2) \sum (b_{jk} - b_{kj}) u^j v^k.
\end{equation}

If $F : \mathcal{O} \to \Omega$ is a smooth map as above, we define the pull-back $F^* \alpha$ of a $k$-form $\alpha$, given by (6.12), to be

\begin{equation}
F^* \alpha = \sum_j a_j(F(y)) (F^* dx_{j1}) \wedge \cdots \wedge (F^* dx_{jk})
\end{equation}

where

\begin{equation}
F^* dx_j = \sum_{\ell} \frac{\partial F_j}{\partial y_\ell} dy_\ell,
\end{equation}

the algebraic computation in (6.18) being performed using the rule (6.14). Extending (6.9), if $F$ is a diffeomorphism, we have

\begin{equation}
(F^* \alpha)(F_\# X_1, \ldots, F_\# X_k) = \alpha(X_1, \ldots, X_k) \circ F.
\end{equation}

If $B = (b_{jk})$ is an $n \times n$ matrix, then, by (6.14),

\begin{equation}
\left( \sum_k b_{1k} dx_k \right) \wedge \left( \sum_k b_{2k} dx_k \right) \wedge \cdots \wedge \left( \sum_k b_{nk} dx_k \right)
= \sum_{k_1, \ldots, k_n} b_{1k_1} \cdots b_{nk_n} dx_{k_1} \wedge \cdots \wedge dx_{k_n}
= \left( \sum_{\sigma \in S_n} (\text{sgn} \, \sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} \right) dx_1 \wedge \cdots \wedge dx_n
= (\det B) \, dx_1 \wedge \cdots \wedge dx_n.
\end{equation}

Here $S_n$ denotes the set of permutations of $\{1, \ldots, n\}$, and the last identity is the formula for the determinant presented in (1.101). It follows that if $F : \mathcal{O} \to \Omega$ is a $C^1$ map between two domains of dimension $n$, and

\begin{equation}
\alpha = A(x) \, dx_1 \wedge \cdots \wedge dx_n
\end{equation}
is an $n$-form on $\Omega$, then

$$F^* \alpha = \det DF(y) \ A(F(y)) \ dy_1 \wedge \cdots \wedge dy_n.$$  

(6.23)

Comparison with the change of variable formula for multiple integrals suggests that one has an intrinsic definition of $\int_{\Omega} \alpha$ when $\alpha$ is an $n$-form on $\Omega$, $n = \dim \Omega$. To implement this, we need to take into account that $\det DF(y)$ rather than $|\det DF(y)|$ appears in (6.21). We say a smooth map $F : \mathcal{O} \to \Omega$ between two open subsets of $\mathbb{R}^n$ preserves orientation if $\det DF(y)$ is everywhere positive. The object called an “orientation” on $\Omega$ can be identified as an equivalence class of nowhere vanishing $n$-forms on $\Omega$, two such forms being equivalent if one is a multiple of another by a positive function in $C^\infty(\Omega)$; the standard orientation on $\mathbb{R}^n$ is determined by $dx_1 \wedge \cdots \wedge dx_n$. If $S$ is an $n$-dimensional surface in $\mathbb{R}^{n+k}$, an orientation on $S$ can also be specified by a nowhere vanishing form $\omega \in \Lambda^n(S)$. If such a form exists, $S$ is said to be orientable. The equivalence class of positive multiples $a(x)\omega$ is said to consist of “positive” forms. A smooth map $\psi : S \to M$ between oriented $n$-dimensional surfaces preserves orientation provided $\psi^* \sigma$ is positive on $S$ whenever $\sigma \in \Lambda^n(M)$ is positive. If $S$ is oriented, one can choose coordinate charts which are all orientation preserving. We mention that there exist surfaces that cannot be oriented, such as the famous “Möbius strip,” and also the projective space $\mathbb{P}^2$, discussed in §5.

We define the integral of an $n$-form over an oriented $n$-dimensional surface as follows. First, if $\alpha$ is an $n$-form supported on an open set $\Omega \subset \mathbb{R}^n$, given by (6.22), then we set

$$\int_{\Omega} \alpha = \int_{\Omega} A(x) \ dV(x),$$  

(6.24)

the right side defined as in §4. If $\mathcal{O}$ is also open in $\mathbb{R}^n$ and $F : \mathcal{O} \to \Omega$ is an orientation preserving diffeomorphism, we have

$$\int_{\mathcal{O}} F^* \alpha = \int_{\Omega} \alpha,$$  

(6.25)

as a consequence of (6.23) and the change of variable formula (4.47). More generally, if $S$ is an $n$-dimensional surface with an orientation, say the image of an open set $\mathcal{O} \subset \mathbb{R}^n$ by $\varphi : \mathcal{O} \to S$, carrying the natural orientation of $\mathcal{O}$, we can set

$$\int_S \alpha = \int_{\mathcal{O}} \varphi^* \alpha$$  

(6.26)

for an $n$-form $\alpha$ on $S$. If it takes several coordinate patches to cover $S$, define $\int_S \alpha$ by writing $\alpha$ as a sum of forms, each supported on one patch.

We need to show that this definition of $\int_S \alpha$ is independent of the choice of coordinate system on $S$ (as long as the orientation of $S$ is respected). Thus, suppose $\varphi : \mathcal{O} \to U \subset S$
and $\psi : \Omega \to U \subset S$ are both coordinate patches, so that $F = \psi^{-1} \circ \varphi : \mathcal{O} \to \Omega$ is an orientation-preserving diffeomorphism, as in Fig. 5.1 of the last section. We need to check that, if $\alpha$ is an $n$-form on $S$, supported on $U$, then

\begin{equation}
\int_{\mathcal{O}} \varphi^* \alpha = \int_{\Omega} \psi^* \alpha.
\end{equation}

To establish this, we first show that, for any form $\alpha$ of any degree,

\begin{equation}
\psi \circ F = \varphi \implies \varphi^* \alpha = F^* \psi^* \alpha.
\end{equation}

It suffices to check (6.28) for $\alpha = dx_j$. Then (6.19) gives $\psi^* dx_j = \sum (\partial \psi_j / \partial x_\ell) dx_\ell$, so

\begin{equation}
F^* \psi^* dx_j = \sum_{\ell,m} \frac{\partial F_\ell}{\partial x_m} \frac{\partial \psi_j}{\partial x_\ell} dx_m, \quad \varphi^* dx_j = \sum_m \frac{\partial \varphi_j}{\partial x_m} dx_m;
\end{equation}

but the identity of these forms follows from the chain rule:

\begin{equation}
D \varphi = (D\psi)(DF) \implies \frac{\partial \varphi_j}{\partial x_m} = \sum_\ell \frac{\partial \psi_j}{\partial x_\ell} \frac{\partial F_\ell}{\partial x_m}.
\end{equation}

Now that we have (6.28), we see that the left side of (6.27) is equal to

\begin{equation}
\int_{\mathcal{O}} F^*(\psi^* \alpha),
\end{equation}

which is equal to the right side of (6.27), by (6.25). Thus the integral of an $n$-form over an oriented $n$-dimensional surface is well defined.

**Exercises**

1. If $F : U_0 \to U_1$ and $G : U_1 \to U_2$ are smooth maps and $\alpha \in \Lambda^k(U_2)$, then (6.26) implies

\begin{equation}
(G \circ F)^* \alpha = F^* (G^* \alpha) \text{ in } \Lambda^k(U_0).
\end{equation}

In the special case that $U_j = \mathbb{R}^n$ and $F$ and $G$ are linear maps, and $k = n$, show that this identity implies

\begin{equation}
\det(GF) = (\det F)(\det G).
\end{equation}

Compare this with the derivation of (1.87).
2. Let $\Lambda^k \mathbb{R}^n$ denote the space of $k$-forms (6.12) with constant coefficients. Show that

$$\dim_{\mathbb{R}} \Lambda^k \mathbb{R}^n = \binom{n}{k}. \tag{6.34}$$

If $T : \mathbb{R}^m \to \mathbb{R}^n$ is linear, then $T^*$ preserves this class of spaces; we denote the map

$$\Lambda^k T^* : \Lambda^k \mathbb{R}^n \to \Lambda^k \mathbb{R}^m. \tag{6.35}$$

Similarly, replacing $T$ by $T^*$ yields

$$\Lambda^k T : \Lambda^k \mathbb{R}^m \to \Lambda^k \mathbb{R}^n. \tag{6.36}$$

3. Show that $\Lambda^k T$ is uniquely characterized as a linear map from $\Lambda^k \mathbb{R}^m$ to $\Lambda^k \mathbb{R}^n$ which satisfies

$$(\Lambda^k T)(v_1 \land \cdots \land v_k) = (Tv_1) \land \cdots \land (Tv_k), \quad v_j \in \mathbb{R}^m. \tag{6.37}$$

4. Show that if $S, T : \mathbb{R}^n \to \mathbb{R}^n$ are linear maps, then

$$\Lambda^k (ST) = (\Lambda^k S) \circ (\Lambda^k T). \tag{6.38}$$

Relate this to (6.28).

If $\{e_1, \ldots, e_n\}$ is the standard orthonormal basis of $\mathbb{R}^n$, define an inner product on $\Lambda^k \mathbb{R}^n$ by declaring an orthonormal basis to be

$$\{e_{j_1} \land \cdots \land e_{j_k} : 1 \leq j_1 < \cdots < j_k \leq n\}. \tag{6.39}$$

If $A : \Lambda^k \mathbb{R}^n \to \Lambda^k \mathbb{R}^n$ is a linear map, define $A^t : \Lambda^k \mathbb{R}^n \to \Lambda^k \mathbb{R}^n$ by

$$\langle A\alpha, \beta \rangle = \langle \alpha, A^t \beta \rangle, \quad \alpha, \beta \in \Lambda^k \mathbb{R}^n, \tag{6.40}$$

where $\langle \ , \ \rangle$ is the inner product on $\Lambda^k \mathbb{R}^n$ defined above.

5. Show that, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is linear, with transpose $T^t$, then

$$\Lambda^k T^t = \Lambda^k (T^t). \tag{6.41}$$

Hint. Check the identity $\langle (\Lambda^k T)\alpha, \beta \rangle = \langle \alpha, (\Lambda^k T^t)\beta \rangle$ when $\alpha$ and $\beta$ run over the orthonormal basis (6.38). That is, show that if $\alpha = e_{j_1} \land \cdots \land e_{j_k}$, $\beta = e_{i_1} \land \cdots \land e_{i_k}$, then

$$\langle Te_{j_1} \land \cdots \land Te_{j_k}, e_{i_1} \land \cdots \land e_{i_k} \rangle = \langle e_{j_1} \land \cdots \land e_{j_k}, T^t e_{i_1} \land \cdots \land T^t e_{i_k} \rangle. \tag{6.42}$$
Hint. Say $T = (t_{ij})$. In the spirit of (6.21), expand $T e_{j_1} \wedge \cdots \wedge T e_{j_k}$, and show that the left side of (6.41) is equal to

$$\sum_{\sigma \in S_k} (\text{sgn } \sigma) t_{i_{\sigma(1)} j_1} \cdots t_{i_{\sigma(k)} j_k},$$

where $S_k$ denotes the set of permutations of $\{1, \ldots, k\}$. Similarly, show that the right side of (6.41) is equal to

$$\sum_{\tau \in S_k} (\text{sgn } \tau) t_{i_{\tau(1)} j_1} \cdots t_{i_{\tau(k)} j_k}. \tag{6.43}$$

To compare these two formulas, see the hint for (1.106) in §1.

6. Show that if $\{u_1, \ldots, u_n\}$ is any orthonormal basis of $\mathbb{R}^n$, then the set $\{u_{j_1} \wedge \cdots \wedge u_{j_k} : 1 \leq j_1 < \cdots < j_k \leq n\}$ is an orthonormal basis of $\Lambda^k \mathbb{R}^n$.

Hint. Use Exercises 4 and 5 to show that if $T : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation on $\mathbb{R}^n$ (i.e., preserves the inner product) then $\Lambda^k T$ is an orthogonal transformation on $\Lambda^k \mathbb{R}^n$.

7. Let $v_j, w_j \in \mathbb{R}^n$, $1 \leq j \leq k \ (k < n)$. Form the matrices $V$, whose $k$ columns are the column vectors $v_1, \ldots, v_k$, and $W$, whose $k$ columns are the column vectors $w_1, \ldots, w_k$. Show that

$$\langle v_1 \wedge \cdots \wedge v_k, w_1 \wedge \cdots \wedge w_k \rangle = \det W^t V = \det V^t W. \tag{6.44}$$

Hint. Show that both sides are linear in each $v_j$ and in each $w_j$. (To treat the right side, see the exercises on determinants, in §1.) Use this to reduce the problem to verifying (6.44) when each $v_j$ and each $w_j$ is chosen from among the set of basis vectors $\{e_1, \ldots, e_n\}$. Use anti-symmetries to reduce the problem further.

8. Deduce from Exercise 7 that if $v_j, w_j \in \mathbb{R}^n$, then

$$\langle v_1 \wedge \cdots \wedge v_k, w_1 \wedge \cdots \wedge w_k \rangle = \sum_{\pi} (\text{sgn } \pi) \langle v_1, w_{\pi(1)} \rangle \cdots \langle v_k, w_{\pi(k)} \rangle, \tag{6.45}$$

where $\pi$ ranges over the set of permutations of $\{1, \ldots, k\}$.

9. Show that the conclusion of Exercise 6 also follows from (6.45).

10. Let $A, B : \mathbb{R}^k \to \mathbb{R}^n$ be linear maps and set $\omega = e_1 \wedge \cdots \wedge e_k \in \Lambda^k \mathbb{R}^k$. We have $\Lambda^k A \omega, \Lambda^k B \omega \in \Lambda^k \mathbb{R}^n$. Deduce from (6.44) that

$$\langle \Lambda^k A \omega, \Lambda^k B \omega \rangle = \det B^t A. \tag{6.46}$$
11. Let $\varphi: \mathcal{O} \to \mathbb{R}^n$ be smooth, with $\mathcal{O} \subset \mathbb{R}^m$ open. Deduce from Exercise 10 that, for each $x \in \mathcal{O}$,

$$\|\Lambda^m D\varphi(x) \omega\|^2 = \det D\varphi(x)^t D\varphi(x),$$

(6.47)

where $\omega = e_1 \wedge \cdots \wedge e_m$. Deduce that if $\varphi: \mathcal{O} \to U \subset M$ is a coordinate patch on a smooth $m$-dimensional surface $M \subset \mathbb{R}^n$ and $f \in C(M)$ is supported on $U$, then

$$\int_M f \, dS = \int_{\mathcal{O}} f(\varphi(x)) \|\Lambda^m D\varphi(x) \omega\| \, dx.$$  

(6.48)

12. Show that the result of Exercise 5 in §5 follows from (6.48), via (6.41)–(6.42).

13. Recall the projective spaces $\mathbb{P}^n$, constructed in §5. Show that $\mathbb{P}^n$ is orientable if and only if $n$ is odd.
7. Products and exterior derivatives of forms

Having discussed the notion of a differential form as something to be integrated, we now consider some operations on forms. There is a wedge product, or exterior product, characterized as follows. If \( \alpha \in \Lambda^k(\Omega) \) has the form (6.12) and if

\[
(7.1) \quad \beta = \sum_i b_i(x) \, dx_{i_1} \wedge \cdots \wedge dx_{i_\ell} \in \Lambda^\ell(\Omega),
\]

define

\[
(7.2) \quad \alpha \wedge \beta = \sum_{j,i} a_j(x) b_i(x) \, dx_{j_1} \wedge \cdots \wedge dx_{j_k} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_\ell}
\]

in \( \Lambda^{k+\ell}(\Omega) \). A special case of this arose in (6.18)–(6.21). We retain the equivalence (6.14). It follows easily that

\[
(7.3) \quad \alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha.
\]

In addition, there is an interior product if \( \alpha \in \Lambda^k(\Omega) \) with a vector field \( X \) on \( \Omega \), producing \( \iota_X \alpha = \alpha| X \in \Lambda^{k-1}(\Omega) \), defined by

\[
(7.4) \quad (\alpha| X)(X_1, \ldots, X_{k-1}) = \alpha(X, X_1, \ldots, X_{k-1}).
\]

Consequently, if \( \alpha = dx_{j_1} \wedge \cdots \wedge dx_{j_k} \), \( D_i = \partial/\partial x_i \), then

\[
(7.5) \quad \alpha| D_{j_i} = (-1)^{i-1} dx_{j_1} \wedge \cdots \wedge \widehat{dx_{j_i}} \wedge \cdots \wedge dx_{j_k}
\]

where \( \widehat{dx_{j_i}} \) denotes removing the factor \( dx_{j_i} \). Furthermore,

\[
i \notin \{j_1, \ldots, j_k\} \implies \alpha| D_i = 0.
\]

If \( F : \mathcal{O} \to \Omega \) is a diffeomorphism and \( \alpha, \beta \) are forms and \( X \) a vector field on \( \Omega \), it is readily verified that

\[
(7.6) \quad F^*(\alpha \wedge \beta) = (F^*\alpha) \wedge (F^*\beta), \quad F^*(\alpha| X) = (F^*\alpha)| (F^\# X).
\]

We make use of the operators \( \wedge_k \) and \( \iota_k \) on forms:

\[
(7.7) \quad \wedge_k \alpha = dx_k \wedge \alpha, \quad \iota_k \alpha = \alpha| D_k.
\]

There is the following useful anticommutation relation:

\[
(7.8) \quad \wedge_k \iota_\ell + \iota_\ell \wedge_k = \delta_{k\ell},
\]
where $\delta_{k\ell}$ is 1 if $k = \ell$, 0 otherwise. This is a fairly straightforward consequence of (7.5). We also have

\[(7.9) \quad \wedge_j \wedge_k + \wedge_k \wedge_j = 0, \quad \tau_j \tau_k + \tau_k \tau_j = 0.\]

From (7.8)–(7.9) one says that the operators $\{\tau_j, \wedge_j : 1 \leq j \leq n\}$ generate a “Clifford algebra.”

Another important operator on forms is the exterior derivative:

\[(7.10) \quad d : \Lambda^k(\Omega) \rightarrow \Lambda^{k+1}(\Omega),\]

defined as follows. If $\alpha \in \Lambda^k(\Omega)$ is given by (6.12), then

\[(7.11) \quad d\alpha = \sum_{j,\ell} \frac{\partial a_j}{\partial x_\ell} dx_\ell \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_k}.\]

Equivalently,

\[(7.12) \quad d\alpha = \sum_{\ell=1}^n \partial_\ell \wedge_\ell \alpha\]

where $\partial_\ell = \partial / \partial x_\ell$ and $\wedge_\ell$ is given by (7.7). The antisymmetry $dx_m \wedge dx_\ell = -dx_\ell \wedge dx_m$, together with the identity $\partial^2 a_j / \partial x_\ell \partial x_m = \partial^2 a_j / \partial x_m \partial x_\ell$, implies

\[(7.13) \quad d(d\alpha) = 0,\]

for any differential form $\alpha$. We also have a product rule:

\[(7.14) \quad d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta), \quad \alpha \in \Lambda^k(\Omega), \beta \in \Lambda^j(\Omega).\]

The exterior derivative has the following important property under pull-backs:

\[(7.15) \quad F^*(d\alpha) = dF^*\alpha,\]

if $\alpha \in \Lambda^k(\Omega)$ and $F : \mathcal{O} \rightarrow \Omega$ is a smooth map. To see this, extending (7.14) to a formula for $d(\alpha \wedge \beta_1 \wedge \cdots \wedge \beta_\ell)$ and using this to apply $d$ to $F^*\alpha$, we have

\[(7.16) \quad dF^*\alpha = \sum_{j,\ell} \frac{\partial}{\partial x_\ell} (a_j \circ F(x)) dx_\ell \wedge (F^*dx_{j_1}) \wedge \cdots \wedge (F^*dx_{j_\ell})
\]

\[+ \sum_{j,\nu} (\pm) a_j (F(x))(F^*dx_{j_1}) \wedge \cdots \wedge d(F^*dx_{j_\nu}) \wedge \cdots \wedge (F^*dx_{j_\ell}).\]

Now the definition (6.18)–(6.19) of pull-back gives directly that

\[(7.17) \quad F^*dx_i = \sum_\ell \frac{\partial F_i}{\partial x_\ell} dx_\ell = dF_i,\]
and hence \( d(F^*dx_i) = df_i = 0 \), so only the first sum in (7.16) contributes to \( dF^*\alpha \).

Meanwhile,

\[
(F^*d\alpha) = \sum_{j,m} \frac{\partial a_j}{\partial x_m} (F(x)) \left( F^*dx_m \right) \left( F^*dx_{j_1} \right) \cdots \left( F^*dx_{j_k} \right),
\]

so (7.15) follows from the identity

\[
\sum_{\ell} \frac{\partial}{\partial x_\ell} (a_j \circ F(x)) \ dx_\ell = \sum_m \frac{\partial a_j}{\partial x_m} (F(x)) F^*dx_m,
\]

which in turn follows from the chain rule.

If \( d\alpha = 0 \), we say \( \alpha \) is closed; if \( \alpha = d\beta \) for some \( \beta \in \Lambda^{k-1}(\Omega) \), we say \( \alpha \) is exact. Formula (7.13) implies that every exact form is closed. The converse is not always true globally. Consider the multi-valued angular coordinate \( \theta \) on \( \mathbb{R}^2 \setminus (0,0) ; d\theta \) is a single valued closed form on \( \mathbb{R}^2 \setminus (0,0) \) which is not globally exact. An important result of H. Poincaré, is that every closed form is locally exact. A proof is given in §11. (A special case is established in §8.)

Exercises

1. If \( \alpha \) is a \( k \)-form, verify the formula (7.14), i.e., \( d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta \).

If \( \alpha \) is closed and \( \beta \) is exact, show that \( \alpha \wedge \beta \) is exact.

2. Let \( F \) be a vector field on \( U \), open in \( \mathbb{R}^3 \), \( F = \sum_1^3 f_j(x) \partial/\partial x_j \). The vector field \( G = \text{curl} \ F \) is classically defined as a formal determinant

\[
\text{curl} \ F = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ \partial_1 & \partial_2 & \partial_3 \\ f_1 & f_2 & f_3 \end{pmatrix},
\]

where \( \{e_j\} \) is the standard basis of \( \mathbb{R}^3 \). Consider the 1-form \( \varphi = \sum_1^3 f_j(x)dx_j \). Show that \( d\varphi \) and \( \text{curl} \ F \) are related in the following way:

\[
\text{curl} \ F = \sum_1^3 g_j(x) \partial/\partial x_j,
\]

\[
d\varphi = g_1(x) \ dx_2 \wedge dx_3 + g_2(x) \ dx_3 \wedge dx_1 + g_3(x) \ dx_1 \wedge dx_2.
\]

See (9.30)–(9.37) for more on this connection.

3. If \( F \) and \( \varphi \) are related as in Exercise 2, show that \( \text{curl} \ F \) is uniquely specified by the relation

\[
d\varphi \wedge \alpha = \langle \text{curl} \ F, \alpha \rangle \omega
\]
for all 1-forms $\alpha$ on $U \subset \mathbb{R}^3$, where $\omega = dx_1 \wedge dx_2 \wedge dx_3$ is the volume form.

4. Let $B$ be a ball in $\mathbb{R}^3$, $F$ a smooth vector field on $B$. Show that
   \begin{equation}
   \exists \ u \in C^\infty(B) \text{ s.t. } F = \text{grad } u \implies \text{curl } F = 0.
   \end{equation}
   \textbf{Hint.} Compare $F = \text{grad } u$ with $\varphi = du$.

5. Let $B$ be a ball in $\mathbb{R}^3$ and $G$ a smooth vector field on $B$. Show that
   \begin{equation}
   \exists \text{ vector field } F \text{ s.t. } G = \text{curl } F \implies \text{div } G = 0.
   \end{equation}
   \textbf{Hint.} If $G = \sum_{j=1}^{3} g_j(x) \partial/\partial x_j$, consider
   \begin{equation}
   \psi = g_1(x) \, dx_2 \wedge dx_3 + g_2(x) \, dx_3 \wedge dx_1 + g_3(x) \, dx_1 \wedge dx_2.
   \end{equation}
   Show that
   \begin{equation}
   d\psi = (\text{div } G) \, dx_1 \wedge dx_2 \wedge dx_3.
   \end{equation}

6. Show that the 1-form $d\theta$ mentioned below (7.19) is given by
   \begin{equation}
   d\theta = \frac{x \, dy - y \, dx}{x^2 + y^2}.
   \end{equation}

For the next set of exercises, let $\Omega$ be a planar domain, $X = f(x, y) \partial/\partial x + g(x, y) \partial/\partial y$ a nonvanishing vector field on $\Omega$. Consider the 1-form $\alpha = g(x, y) \, dx - f(x, y) \, dy$.

7. Let $\gamma : I \to \Omega$ be a smooth curve, $I = (a, b)$. Show that the image $C = \gamma(I)$ is the image of an integral curve of $X$ if and only if $\gamma^* \alpha = 0$. Consequently, with slight abuse of notation, one describes the integral curves by $g \, dx - f \, dy = 0$.
   If $\alpha$ is exact, i.e., $\alpha = du$, conclude the level curves of $u$ are the integral curves of $X$.

8. A function $\varphi$ is called an integrating factor if $\tilde{\alpha} = \varphi \alpha$ is exact, i.e., if $d(\varphi \alpha) = 0$, provided $\Omega$ is simply connected. Show that an integrating factor always exists, at least locally. Show that $\varphi = e^v$ is an integrating factor if and only if $Xv = - \text{div } X$.
   Find an integrating factor for $\alpha = (x^2 + y^2 - 1) \, dx - 2xy \, dy$.

9. Define the radial vector field $R = x_1 \partial/\partial x_1 + \cdots + x_n \partial/\partial x_n$, on $\mathbb{R}^n$. Show that
   \begin{equation}
   \omega = dx_1 \wedge \cdots \wedge dx_n \implies \omega \lrcorner R = \sum_{j=1}^{n} (-1)^{j-1} x_j \, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n.
   \end{equation}
   Show that
   \begin{equation}
   d(\omega \lrcorner R) = n \omega.
   \end{equation}

10. Show that if $F : \mathbb{R}^n \to \mathbb{R}^n$ is a linear rotation (i.e., $F \in SO(n)$) then $\beta = \omega \lrcorner R$ in Exercise 9 has the property that $F^* \beta = \beta$. 

8. The general Stokes formula

The Stokes formula involves integrating a k-form over a k-dimensional surface with boundary. We first define that concept. Let $S$ be a smooth k-dimensional surface (say in $\mathbb{R}^N$), and let $M$ be an open subset of $S$, such that its closure $\overline{M}$ (in $\mathbb{R}^N$) is contained in $S$. Its boundary is $\partial M = \overline{M} \setminus M$. We say $\overline{M}$ is a smooth surface with boundary if also $\partial M$ is a smooth $(k-1)$-dimensional surface. In such a case, any $p \in \partial M$ has a neighborhood $U \subset S$ with a coordinate chart $\varphi : \mathcal{O} \to U$, where $\mathcal{O}$ is an open neighborhood of 0 in $\mathbb{R}^k$, such that $\varphi(0) = p$ and $\varphi$ maps $\{x \in \mathcal{O} : x_1 = 0\}$ onto $U \cap \partial M$.

If $S$ is oriented, then $\overline{M}$ is oriented, and $\partial M$ inherits an orientation, uniquely determined by the following requirement: if

$$(8.1) \quad \overline{M} = \mathbb{R}^k_+ = \{x \in \mathbb{R}^k : x_1 \leq 0\},$$

then $\partial M = \{(x_2, \ldots, x_k)\}$ has the orientation determined by $dx_2 \wedge \cdots \wedge dx_k$.

We can now state the Stokes formula.

**Proposition 8.1.** Given a compactly supported $(k-1)$-form $\beta$ of class $C^1$ on an oriented k-dimensional surface $\overline{M}$ (of class $C^2$) with boundary $\partial M$, with its natural orientation,

$$(8.2) \quad \int_M d\beta = \int_{\partial M} \beta.$$  

**Proof.** Using a partition of unity and invariance of the integral and the exterior derivative under coordinate transformations, it suffices to prove this when $\overline{M}$ has the form (8.1). In that case, we will be able to deduce (8.2) from the Fundamental Theorem of Calculus. Indeed, if

$$(8.3) \quad \beta = b_j(x) \, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_k,$$

with $b_j(x)$ of bounded support, we have

$$(8.4) \quad d\beta = (-1)^{j-1} \frac{\partial b_j}{\partial x_j} dx_1 \wedge \cdots \wedge dx_k.$$ 

If $j > 1$, we have

$$(8.5) \quad \int_M d\beta = (-1)^{j-1} \int \left\{ \int_{-\infty}^{\infty} \frac{\partial b_i}{\partial x_j} \, dx_j \right\} dx' = 0,$$
and also $\kappa^* \beta = 0$, where $\kappa : \partial M \to \overline{M}$ is the inclusion. On the other hand, for $j = 1$, we have

$$
\int_{\partial M} \beta = \int \left\{ \int_{-\infty}^{0} \frac{\partial b_1}{\partial x_1} \, dx_1 \right\} \, dx_2 \cdots dx_k
$$

$$
= \int b_1(0, x') \, dx'
$$

$$
= \int_{\partial M} \beta.
$$

This proves Stokes’ formula (8.2).

It is useful to allow singularities in $\partial M$. We say a point $p \in \overline{M}$ is a corner of dimension $\nu$ if there is a neighborhood $U$ of $p$ in $\overline{M}$ and a $C^2$ diffeomorphism of $U$ onto a neighborhood of 0 in

$$
K = \{ x \in \mathbb{R}^k : x_j \leq 0, \text{ for } 1 \leq j \leq k - \nu \},
$$

where $k$ is the dimension of $M$. If $M$ is a $C^2$ surface and every point $p \in \partial M$ is a corner (of some dimension), we say $\overline{M}$ is a $C^2$ surface with corners. In such a case, $\partial M$ is a locally finite union of $C^2$ surfaces with corners. The following result extends Proposition 8.1.

**Proposition 8.2.** If $\overline{M}$ is a $C^2$ surface of dimension $k$, with corners, and $\beta$ is a compactly supported $(k - 1)$-form of class $C^1$ on $\overline{M}$, then (8.2) holds.

**Proof.** It suffices to establish this when $\beta$ is supported on a small neighborhood of a corner $p \in \partial M$, of the form $U$ described above. Hence it suffices to show that (8.2) holds whenever $\beta$ is a $(k - 1)$-form of class $C^1$, with compact support on $K$ in (8.7); and we can take $\beta$ to have the form (8.3). Then, for $j > k - \nu$, (8.5) still holds, while, for $j \leq k - \nu$, we have, as in (8.6),

$$
\int_{K} \beta = (-1)^{j-1} \int \left\{ \int_{-\infty}^{0} \frac{\partial b_j}{\partial x_j} \, dx_j \right\} \, dx_1 \cdots \widehat{dx_j} \cdots dx_k
$$

$$
= (-1)^{j-1} \int b_j(x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_k) \, dx_1 \cdots \widehat{dx_j} \cdots dx_k
$$

$$
= \int_{\partial K} \beta.
$$

This completes the proof.

The reason we required $\overline{M}$ to be a surface of class $C^2$ (with corners) in Propositions 8.1 and 8.2 is the following. Due to the formulas (6.18)–(6.19) for a pull-back, if $\beta$ is of class $C^j$ and $F$ is of class $C^\ell$, then $F^* \beta$ is generally of class $C^\mu$, with $\mu = \min(j, \ell - 1)$. 
Thus, if \( j = \ell = 1 \), \( F^* \beta \) might be only of class \( C^0 \), so there is not a well-defined notion of a differential form of class \( C^1 \) on a \( C^1 \) surface, though such a notion is well defined on a \( C^2 \) surface. This problem can be overcome, and one can extend Propositions 8.1–8.2 to the case where \( \overline{M} \) is a \( C^1 \) surface (with corners), and \( \beta \) is a \((k-1)\)-form with the property that both \( \beta \) and \( d\beta \) are continuous. We will not go into the details. Substantially more sophisticated generalizations are given in [Fed].

We will mention one useful extension of the scope of Proposition 8.2, to surfaces with piecewise smooth boundary that do not satisfy the corner condition. An example is illustrated in Fig. 8.1. There the point \( p \) is a singular point of \( \partial M \) that is not a corner, according to the definition using (8.7). However, in many cases \( M \) can be divided into pieces (\( \overline{M}_1 \) and \( \overline{M}_2 \) for the example presented in Fig. 8.1) and each piece \( M_j \) is a surface with corners. Then Proposition 8.2 applies to each piece separately:

\[
\int_{M_j} d\beta = \int_{\partial M_j} \beta,
\]

and one can sum over \( j \) to get (8.2) in this more general setting.

We next apply Proposition 8.2 to prove the following special case of the Poincaré lemma, which will be used in §10.

**Proposition 8.3.** If \( \alpha \) is a 1-form on \( B = \{ x \in \mathbb{R}^2 : |x| < 1 \} \) and \( d\alpha = 0 \), then there exists a real valued \( u \in C^\infty(B) \) such that \( \alpha = du \).

In fact, let us set

\[
(8.9) \quad u_j(x) = \int_{\gamma_j(x)} \alpha,
\]

where \( \gamma_j(x) \) is a path from 0 to \( x = (x_1, x_2) \) which consists of two line segments. The path first goes from 0 to \((0, x_2)\) and then from \((0, x_2)\) to \( x \), if \( j = 1 \), while if \( j = 2 \) it first goes from 0 to \((x_1, 0)\) and then from \((x_1, 0)\) to \( x \). See Fig. 8.2. It is easy to verify that \( \partial u_j / \partial x_j = \alpha_j(x) \). We claim that \( u_1 = u_2 \), or equivalently that

\[
(8.10) \quad \int_{\sigma(x)} \alpha = 0,
\]

where \( \sigma(x) \) is a closed path consisting of \( \gamma_2(x) \) followed by \( \gamma_1(x) \), in reverse. In fact, Stokes’ formula, Proposition 8.2, implies that

\[
(8.11) \quad \int_{\sigma(x)} \alpha = \int_{R(x)} d\alpha,
\]

where \( R(x) \) is the rectangle whose boundary is \( \sigma(x) \). If \( d\alpha = 0 \), then (8.11) vanishes, and we have the desired function \( u : u = u_1 = u_2 \).
Exercises

1. In the setting of Proposition 8.1, show that

$$\partial M = \emptyset \implies \int_M d\beta = 0.$$ 

2. Consider the region $\Omega \subset \mathbb{R}^2$ defined by

$$\Omega = \{(x, y) : 0 \leq y \leq x^2, 0 \leq x \leq 1\}.$$ 

Show that the boundary points $(1, 0)$ and $(1, 1)$ are corners, but $(0, 0)$ is not a corner. The boundary of $\Omega$ is too sharp at $(0, 0)$ to be a corner; it is called a “cusp.” Extend Proposition 8.2 to treat this region.

3. Suppose $U \subset \mathbb{R}^n$ is an open set with smooth boundary $M = \partial U$, and $U$ has the standard orientation, determined by $dx_1 \wedge \cdots \wedge dx_n$. (See the paragraph above (6.23).) Let $\varphi \in C^1(\mathbb{R}^n)$ satisfy $\varphi(x) < 0$ for $x \in U$, $\varphi(x) > 0$ for $x \in \mathbb{R}^n \setminus \overline{U}$, and grad $\varphi(x) \neq 0$ for $x \in \partial U$, so grad $\varphi$ points out of $U$. Show that the natural orientation on $\partial U$, as defined just before Proposition 8.1, is the same as the following. The equivalence class of forms $\beta \in \Lambda^{n-1}(\partial U)$ defining the orientation on $\partial U$ satisfies the property that $d\varphi \wedge \beta$ is a positive multiple of $dx_1 \wedge \cdots \wedge dx_n$, on $\partial U$.

4. Suppose $U = \{x \in \mathbb{R}^n : x_n < 0\}$. Show that the orientation on $\partial U$ described above is that of $(-1)^{n-1} dx_1 \wedge \cdots \wedge dx_{n-1}$. If $V = \{x \in \mathbb{R}^n : x_n > 0\}$, what orientation does $\partial V$ inherit?

5. Extend the special case of Poincaré’s Lemma given in Proposition 8.3 to the case where $\alpha$ is a closed 1-form on $B = \{x \in \mathbb{R}^n : |x| < 1\}$, i.e., from the case dim $B = 2$ to higher dimensions.

6. Define $\beta \in \Lambda^{n-1}(\mathbb{R}^n)$ by

$$\beta = \sum_{j=1}^{n} (-1)^{j-1} x_j \, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n.$$ 

Let $\Omega \subset \mathbb{R}^n$ be a smoothly bounded compact subset. Show that

$$\frac{1}{n} \int_{\partial \Omega} \beta = \text{Vol}(\Omega).$$
7. In the setting of Exercise 6, show that if \( f \in C^1(\overline{\Omega}) \), then

\[
\int_{\partial \Omega} f = \int_{\Omega} (Rf + nf) \, dx,
\]

where

\[
Rf = \sum_{j=1}^{n} x_j \frac{\partial f}{\partial x_j}.
\]

8. In the setting of Exercises 6–7, and with \( S^{n-1} \subset \mathbb{R}^n \) the unit sphere, show that

\[
\int_{S^{n-1}} f \beta = \int_{S^{n-1}} f \, dS.
\]

**Hint.** Let \( B \subset \mathbb{R}^{n-1} \) be the unit ball, and define \( \varphi : B \to S^{n-1} \) by \( \varphi(x') = (x', \sqrt{1 - |x'|^2}) \). Compute \( \varphi^* \beta \). Compare surface area formulas derived in §5.

Another approach. The unit sphere \( S^{n-1} \hookrightarrow \mathbb{R}^n \) has a volume form (cf. (9.13)), it must be a scalar multiple \( g(j^* \beta) \), and, by Exercise 10 of §7, \( g \) must be constant. Then Exercise 6 identifies this constant, in light of results from §5.

See the exercises in §9 for more on this.

9. Given \( \beta \) as in Exercise 6, show that the \((n-1)\)-form

\[
\omega = |x|^{-n} \beta
\]

on \( \mathbb{R}^n \setminus 0 \) is closed. Use Exercise 6 to show that \( \int_{S^{n-1}} \omega \neq 0 \), and hence \( \omega \) is not exact.

10. Let \( \Omega \subset \mathbb{R}^n \) be a compact, smoothly bounded subset. Take \( \omega \) as in Exercise 9. Show that

\[
\int_{\partial \Omega} \omega = A_{n-1} \quad \text{if} \quad 0 \in \Omega,
\]

\[
0 \quad \text{if} \quad 0 \notin \overline{\Omega}.
\]
9. The classical Gauss, Green, and Stokes formulas

The case of (8.1) where \( S = \Omega \) is a region in \( \mathbb{R}^2 \) with smooth boundary yields the classical Green Theorem. In this case, we have

\[
\beta = f \, dx + g \, dy, \quad d\beta = \left( \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \wedge dy,
\]

and hence (8.1) becomes the following

**Proposition 9.1.** If \( \Omega \) is a region in \( \mathbb{R}^2 \) with smooth boundary, and \( f \) and \( g \) are smooth functions on \( \Omega \), which vanish outside some compact set in \( \Omega \), then

\[
\iint_{\Omega} \left( \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \, dy = \int_{\partial \Omega} (f \, dx + g \, dy).
\]

Note that, if we have a vector field \( X = X_1 \partial/\partial x + X_2 \partial/\partial y \) on \( \Omega \), then the integrand on the left side of (9.2) is

\[
\frac{\partial X_1}{\partial x} + \frac{\partial X_2}{\partial y} = \text{div} \, X,
\]

provided \( g = X_1 \), \( f = -X_2 \). We obtain

\[
\iint_{\Omega} \left( \text{div} \, X \right) dx \, dy = \int_{\partial \Omega} (X \cdot \nu) \, ds.
\]

If \( \partial \Omega \) is parametrized by arc-length, as \( \gamma(s) = (x(s), y(s)) \), with orientation as defined for Proposition 8.1, then the unit normal \( \nu \), to \( \partial \Omega \), pointing **out** of \( \Omega \), is given by \( \nu(s) = (y'(s), -x'(s)) \), and (9.4) is equivalent to

\[
\iint_{\Omega} \left( \text{div} \, X \right) dx \, dy = \int_{\partial \Omega} \langle X, \nu \rangle \, ds.
\]

This is a special case of Gauss’ Divergence Theorem. We now derive a more general form of the Divergence Theorem. We begin with a definition of the divergence of a vector field on a surface \( M \).

Let \( M \) be a region in \( \mathbb{R}^n \), or an \( n \)-dimensional surface in \( \mathbb{R}^{n+k} \), provided with a volume form

\[
\omega_M \in \Lambda^n M.
\]
Let $X$ be a vector field on $M$. Then the divergence of $X$, denoted $\text{div} \ X$, is a function on $M$ given by

$$\text{div} \ X \omega_M = d(\omega_M \vert X). \tag{9.7}$$

If $M = \mathbb{R}^n$, with the standard volume element

$$\omega = dx_1 \wedge \cdots \wedge dx_n, \tag{9.8}$$

and if

$$X = \sum X^j(x) \frac{\partial}{\partial x_j}, \tag{9.9}$$

then

$$\omega \vert X = \sum_{j=1}^{n} (-1)^{j-1} X^j(x) dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n. \tag{9.10}$$

Hence, in this case, (9.7) yields the familiar formula

$$\text{div} \ X = \sum_{j=1}^{n} \partial_j X^j, \tag{9.11}$$

where we use the notation

$$\partial_j f = \frac{\partial f}{\partial x_j}. \tag{9.12}$$

Suppose now that $M$ is endowed with both an orientation and a metric tensor $g_{jk}(x)$. Then $M$ carries a natural volume element $\omega_M$, determined by the condition that, if one has an orientation-preserving coordinate system in which $g_{jk}(p_0) = \delta_{jk}$, then $\omega_M(p_0) = dx_1 \wedge \cdots \wedge dx_n$. This condition produces the following formula, in any orientation-preserving coordinate system:

$$\omega_M = \sqrt{g} \; dx_1 \wedge \cdots \wedge dx_n, \quad g = \det(g_{jk}), \tag{9.13}$$

by the same sort of calculations as done in (5.10)--(5.15).

We now compute $\text{div} \ X$ when the volume element on $M$ is given by (9.13). We have

$$\omega_M \vert X = \sum_{j} (-1)^{j-1} X^j \sqrt{g} \; dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n \tag{9.14}$$

and hence

$$d(\omega_M \vert X) = \partial_j (\sqrt{g}X^j) \; dx_1 \wedge \cdots \wedge dx_n. \tag{9.15}$$
Here, as below, we use the summation convention. Hence the formula (9.7) gives

\begin{equation}
\text{div } X = g^{-1/2} \partial_j (g^{1/2} X^j).
\end{equation}

Compare (5.55).

We now derive the Divergence Theorem, as a consequence of Stokes’ formula, which we recall is

\begin{equation}
\int_M d\alpha = \int_{\partial M} \alpha,
\end{equation}

for an \((n-1)\)-form on \(\overline{M}\), assumed to be a smooth compact oriented surface with boundary. If \(\alpha = \omega_M \mid X\), formula (9.7) gives

\begin{equation}
\int_M (\text{div } X) \omega_M = \int_{\partial M} \langle X, \nu \rangle \omega_{\partial M}.
\end{equation}

This is one form of the Divergence Theorem. We will produce an alternative expression for the integrand on the right before stating the result formally.

Given that \(\omega_M\) is the volume form for \(M\) determined by a Riemannian metric, we can write the interior product \(\omega_M \mid X\) in terms of the volume element \(\omega_{\partial M}\) on \(\partial M\), with its induced orientation and Riemannian metric, as follows. Pick coordinates on \(M\), centered at \(p_0 \in \partial M\), such that \(\partial M\) is tangent to the hyperplane \(\{x_1 = 0\}\) at \(p_0 = 0\) (with \(M\) to the left of \(\partial M\)), and such that \(g_{jk}(p_0) = \delta_{jk}\), so \(\omega_M(p_0) = dx_1 \wedge \cdots \wedge dx_n\). Consequently, \(\omega_{\partial M}(p_0) = dx_2 \wedge \cdots \wedge dx_2\). It follows that, at \(p_0\),

\begin{equation}
\langle X, \nu \rangle \omega_{\partial M},
\end{equation}

where \(\nu\) is the unit vector normal to \(\partial M\), pointing out of \(M\) and \(j : \partial M \hookrightarrow M\) the natural inclusion. The two sides of (9.19), which are both defined in a coordinate independent fashion, are hence equal on \(\partial M\), and the identity (9.18) becomes

\begin{equation}
\int_M (\text{div } X) \omega_M = \int_{\partial M} \langle X, \nu \rangle \omega_{\partial M}.
\end{equation}

Finally, we adopt the notation of the sort used in \S\S 4–5. We denote the volume element on \(M\) by \(dV\) and that on \(\partial M\) by \(dS\), obtaining the Divergence Theorem:

**Theorem 9.2.** If \(\overline{M}\) is a compact surface with boundary, \(X\) a smooth vector field on \(\overline{M}\), then

\begin{equation}
\int_M (\text{div } X) dV = \int_{\partial M} \langle X, \nu \rangle dS,
\end{equation}
where \( \nu \) is the unit outward-pointing normal to \( \partial M \).

The only point left to mention here is that \( M \) need not be orientable. Indeed, we can treat the integrals in (9.21) as surface integrals, as in \( \S 5 \), and note that all objects in (9.21) are independent of a choice of orientation. To prove the general case, just use a partition of unity supported on orientable pieces.

We obtain some further integral identities. First, we apply (9.21) with \( X \) replaced by \( uX \).

We have the following “derivation” identity:

\[
\text{div } uX = u \text{ div } X + \langle du, X \rangle = u \text{ div } X + Xu,
\]

which follows easily from the formula (9.16). The Divergence Theorem immediately gives

\[
\int_M (\text{div } X) u \, dV + \int_M Xu \, dV = \int_{\partial M} \langle X, \nu \rangle u \, dS.
\]

Replacing \( u \) by \( uv \) and using the derivation identity \( X(uv) = (Xu)v + u(Xv) \), we have

\[
\int_M [(Xu)v + u(Xv)] \, dV = -\int_M (\text{div } X) uv \, dV + \int_{\partial M} \langle X, \nu \rangle uv \, dS.
\]

It is very useful to apply (9.23) to a gradient vector field \( X \): If \( v \) is a smooth function on \( M \), \( \text{grad } v \) is a vector field satisfying

\[
\langle \text{grad } v, Y \rangle = \langle Y, dv \rangle,
\]

for any vector field \( Y \), where the brackets on the left are given by the metric tensor on \( M \) and those on the right by the natural pairing of vector fields and 1-forms. Hence \( \text{grad } v \) has components \( X^j = g^{jk} \partial_k v \), where \( (g^{jk}) \) is the matrix inverse of \( (g_{jk}) \).

Applying \( \text{div } \) to \( \text{grad } v \), we have the Laplace operator:

\[
\Delta v = \text{div } \text{grad } v = g^{-1/2} \partial_j \left( g^{jk} g^{1/2} \partial_k v \right).
\]

When \( M \) is a region in \( \mathbb{R}^n \) and we use the standard Euclidean metric, so \( \text{div } X \) is given by (9.11), we have the Laplace operator on Euclidean space:

\[
\Delta v = \frac{\partial^2 v}{\partial x_1^2} + \cdots + \frac{\partial^2 v}{\partial x_n^2}.
\]

Now, setting \( X = \text{grad } v \) in (9.23), we have \( Xu = \langle \text{grad } u, \text{grad } v \rangle \), and \( \langle X, \nu \rangle = \langle \nu, \text{grad } v \rangle \), which we call the normal derivative of \( v \), and denote \( \partial v / \partial \nu \). Hence

\[
\int_M u(\Delta v) \, dV = -\int_M \langle \text{grad } u, \text{grad } v \rangle \, dV + \int_{\partial M} u \frac{\partial v}{\partial \nu} \, dS.
\]
If we interchange the roles of $u$ and $v$ and subtract, we have

\[ \int_M u(\Delta v) \, dV = \int_M (\Delta u)v \, dV + \int_{\partial M} \left[ u \frac{\partial v}{\partial n} - \frac{\partial u}{\partial n} v \right] \, dS. \]  

Formulas (9.28)–(9.29) are also called Green formulas. We will make further use of them in §10.

We return to the Green formula (9.2), and give it another formulation. Consider a vector field $Z = (f, g, h)$ on a region in $\mathbb{R}^3$ containing the planar surface $U = \{(x, y, 0) : (x, y) \in \Omega\}$. If we form

\[ \text{curl } Z = \det \begin{pmatrix} i & j & k \\ \partial_x & \partial_y & \partial_z \\ f & g & h \end{pmatrix} \]

we see that the integrand on the left side of (9.2) is the $k$-component of curl $Z$, so (9.2) can be written

\[ \int_U (\text{curl } Z) \cdot k \, dA = \int_{\partial U} (Z \cdot T) \, ds, \]

where $T$ is the unit tangent vector to $\partial U$. To see how to extend this result, note that $k$ is a unit normal field to the planar surface $U$.

To formulate and prove the extension of (9.31) to any compact oriented surface with boundary in $\mathbb{R}^3$, we use the relation between curl and exterior derivative discussed in Exercises 2–3 of §7. In particular, if we set

\[ F = \sum_{j=1}^{3} f_j(x) \frac{\partial}{\partial x_j}, \quad \varphi = \sum_{j=1}^{3} f_j(x) \, dx_j, \]

then $\text{curl } F = \sum_{j=1}^{3} g_j(x) \partial/\partial x_j$ where

\[ d\varphi = g_1(x) \, dx_2 \wedge dx_3 + g_2(x) \, dx_3 \wedge dx_1 + g_3(x) \, dx_1 \wedge dx_2. \]

Now Suppose $\overline{M}$ is a smooth oriented $(n - 1)$-dimensional surface with boundary in $\mathbb{R}^n$. Using the orientation of $M$, we pick a unit normal field $N$ to $M$ as follows. Take a smooth function $v$ which vanishes on $M$ but such that $\nabla v(x) \neq 0$ on $M$. Thus $\nabla v$ is normal to $M$. Let $\sigma \in \Lambda^{n-1}(M)$ define the orientation of $M$. Then $dv \wedge \sigma = a(x) \, dx_1 \wedge \cdots \wedge dx_n$, where $a(x)$ is nonvanishing on $M$. For $x \in M$, we take $N(x) = \nabla v(x)/|\nabla v(x)|$ if $a(x) > 0$, and $N(x) = -\nabla v(x)/|\nabla v(x)|$ if $a(x) < 0$. We call $N$ the “positive” unit normal field to the oriented surface $M$. Part of the motivation for this characterization of $N$ is that, if $\Omega \subset \mathbb{R}^n$ is an open set with smooth boundary $M = \partial \Omega$, and we give $M$ the induced orientation, as described in §8, then the positive normal field $N$ just defined coincides with the unit normal field pointing out of $\Omega$. Compare Exercises 2–3 of §8.
Now, if \( G = (g_1, \ldots, g_n) \) is a vector field defined near \( M \), then

\[
\int_\mathcal{M} (N \cdot G) \, dS = \int_\mathcal{M} \left( \sum_{j=1}^{n} (-1)^{j-1} g_j(x) \, dx_1 \wedge \cdots \wedge dx_j \wedge \cdots \wedge dx_n \right).
\]

This result follows from (9.19). When \( n = 3 \) and \( G = \text{curl} \, F \), we deduce from (9.32)–(9.33) that

\[
\int_\mathcal{M} d\varphi = \int_\mathcal{M} (N \cdot \text{curl} \, F) \, dS.
\]

Furthermore, in this case we have

\[
\int_{\partial \mathcal{M}} \varphi = \int_{\partial \mathcal{M}} (F \cdot T) \, ds,
\]

where \( T \) is the unit tangent vector to \( \partial \mathcal{M} \), specified as follows by the orientation of \( \partial \mathcal{M} \); if \( \tau \in \Lambda^1(\partial \mathcal{M}) \) defines the orientation of \( \partial \mathcal{M} \), then \( \langle T, \tau \rangle > 0 \) on \( \partial \mathcal{M} \). We call \( T \) the “forward” unit tangent vector field to the oriented curve \( \partial \mathcal{M} \). By the calculations above, we have the classical Stokes formula:

**Proposition 9.3.** If \( \mathcal{M} \) is a compact oriented surface with boundary in \( \mathbb{R}^3 \), and \( F \) is a \( C^1 \) vector field on a neighborhood of \( \mathcal{M} \), then

\[
\int_\mathcal{M} (N \cdot \text{curl} \, F) \, dS = \int_{\partial \mathcal{M}} (F \cdot T) \, ds,
\]

where \( N \) is the positive unit normal field on \( \mathcal{M} \) and \( T \) the forward unit tangent field to \( \partial \mathcal{M} \).

**Remark.** The right side of (9.37) is called the circulation of \( F \) about \( \partial \mathcal{M} \). Proposition 9.3 shows how curl \( F \) arises to measure this circulation.

**Direct proof of the Divergence Theorem**

Let \( \Omega \) be a bounded open subset of \( \mathbb{R}^n \), with a \( C^1 \) smooth boundary \( \partial \Omega \). Hence, for each \( p \in \partial \Omega \), there is a neighborhood \( U \) of \( p \) in \( \mathbb{R}^n \), a rotation of coordinate axes, and a \( C^1 \) function \( u : \mathcal{O} \to \mathbb{R} \), defined on an open set \( \mathcal{O} \subset \mathbb{R}^{n-1} \), such that

\[
\Omega \cap U = \{ x \in \mathbb{R}^n : x_n \leq u(x'), x' \in \mathcal{O} \} \cap U,
\]

where \( x = (x', x_n), \ x' = (x_1, \ldots, x_{n-1}) \).
We aim to prove that, given \( f \in C^1(\Omega) \), and any constant vector \( e \in \mathbb{R}^n \),

\[
\int_{\Omega} e \cdot \nabla f(x) \, dx = \int_{\partial \Omega} (e \cdot N) f \, dS,
\]

where \( dS \) is surface measure on \( \partial \Omega \) and \( N(x) \) is the unit normal to \( \partial \Omega \), pointing out of \( \Omega \). At \( x = (x', u(x')) \in \partial \Omega \), we have

\[
N = (1 + |\nabla u|^2)^{-1/2}(-\nabla u, 1).
\]

To prove (9.38), we may as well suppose \( f \) is supported in such a neighborhood \( U \). Then we have

\[
\int_{\Omega} \frac{\partial f}{\partial x_n} \, dx = \int_{\Omega} \left( \int_{x_n \leq u(x')} \partial_n f(x', x_n) \, dx_n \right) \, dx' \\
= \int_{\Omega} f(x', u(x')) \, dx' \\
= \int_{\partial \Omega} (e_n \cdot N) f \, dS.
\]

The first identity in (9.40) follows from Theorem 4.9, the second identity from the Fundamental Theorem of Calculus, and the third identity from the identification

\[
dS = \left(1 + |\nabla u|^2\right)^{1/2} \, dx',
\]

established in (5.21). We use the standard basis \( \{e_1, \ldots, e_n\} \) of \( \mathbb{R}^n \).

Such an argument works when \( e_n \) is replaced by any constant vector \( e \) with the property that we can represent \( \partial \Omega \cap U \) as the graph of a function \( y_n = \hat{u}(y') \), with the \( y_n \)-axis parallel to \( e \). In particular, it works for \( e = e_n + ae_j \), for \( 1 \leq j \leq n-1 \) and for \( |a| \) sufficiently small. Thus, we have

\[
\int_{\Omega} (e_n + ae_j) \cdot \nabla f(x) \, dx = \int_{\partial \Omega} (e_n + ae_j) \cdot N f \, dS.
\]

If we subtract (9.40) from this and divide the result by \( a \), we obtain (9.38) for \( e = e_j \), for all \( j \), and hence (9.38) holds in general.

Note that replacing \( e \) by \( e_j \) and \( f \) by \( f_j \) in (9.38), and summing over \( 1 \leq j \leq n \), yields

\[
\int_{\Omega} (\text{div } F) \, dx = \int_{\partial \Omega} N \cdot F \, dS.
\]
for the vector field $F = (f_1, \ldots, f_n)$. This is the usual statement of Gauss’ Divergence Theorem, as given in Theorem 9.2 (specialized to domains in $\mathbb{R}^n$).

Reversing the argument leading from (9.2) to (9.5), we also have another proof of Green’s Theorem, in the form (9.2).

Exercises

1. Newton’s equation $md^2x/dt^2 = -\nabla V(x)$ for the motion in $\mathbb{R}^n$ of a body of mass $m$, in a potential force field $F = -\nabla V$, can be converted to a first-order system for $(x, \xi)$, with $\xi = mx$. One gets

$$\frac{d}{dt}(x, \xi) = H_f(x, \xi),$$

where $H_f$ is a “Hamiltonian vector field” on $\mathbb{R}^{2n}$, given by

$$H_f = \sum_{j=1}^{n} \left[ \frac{\partial f}{\partial \xi_j} \frac{\partial}{\partial x_j} - \frac{\partial f}{\partial x_j} \frac{\partial}{\partial \xi_j} \right].$$

In the case described above,

$$f(x, \xi) = \frac{1}{2m} |\xi|^2 + V(x).$$

Calculate $\text{div} H_f$ from (9.11).

2. Let $X$ be a smooth vector field on a smooth surface $M$, generating a flow $\mathcal{F}_X^t$. Let $\mathcal{O} \subset M$ be a compact, smoothly bounded subset, and set $\mathcal{O}_t = \mathcal{F}_X^t(\mathcal{O})$. As seen in Proposition 5.7,

$$(9.43) \quad \frac{d}{dt} \text{Vol}(\mathcal{O}_t) = \int_{\mathcal{O}_t} (\text{div} X) dV.$$ 

Use the Divergence Theorem to deduce from this that

$$(9.44) \quad \frac{d}{dt} \text{Vol}(\mathcal{O}_t) = \int_{\partial \mathcal{O}_t} \langle X, \nu \rangle dS.$$ 

Remark. Conversely, a direct proof of (9.44), together with the Divergence Theorem, would lead to another proof of (9.43).

3. Show that, if $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a linear rotation, then, for a $C^1$ vector field $Z$ on $\mathbb{R}^3$,

$$(9.45) \quad F_\#(\text{curl} Z) = \text{curl}(F_\# Z).$$
4. Let $\overline{M}$ be the graph in $\mathbb{R}^3$ of a smooth function, $z = u(x, y)$, $(x, y) \in \mathcal{O} \subseteq \mathbb{R}^2$, a bounded region with smooth boundary (maybe with corners). Show that

$$\int_M (\text{curl } F \cdot N) \, dS = \int \int_{\mathcal{O}} \left[ \left( \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \left( -\frac{\partial u}{\partial x} \right) + \left( \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) \left( -\frac{\partial u}{\partial y} \right) \right. $$

$$\left. + \left( \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \right] \, dx \, dy,$$

where $\partial F_j / \partial x$ and $\partial F_j / \partial y$ are evaluated at $(x, y, u(x, y))$. Show that

$$\int_{\partial M} (F \cdot T) \, ds = \int_{\partial \mathcal{O}} \left( \widetilde{F}_1 + \widetilde{F}_3 \frac{\partial u}{\partial x} \right) \, dx + \left( \widetilde{F}_2 + \widetilde{F}_3 \frac{\partial u}{\partial y} \right) \, dy,$$

where $\widetilde{F}_j(x, y) = F_j(x, y, u(x, y))$. Apply Green’s Theorem, with $f = \widetilde{F}_1 + \widetilde{F}_3(\partial u / \partial x)$, $g = \widetilde{F}_2 + \widetilde{F}_3(\partial u / \partial y)$, to show that the right sides of (9.46) and (9.47) are equal, hence proving Stokes’ Theorem in this case.

5. Let $M \subseteq \mathbb{R}^n$ be the graph of a function $x_n = u(x')$, $x' = (x_1, \ldots, x_{n-1})$. If

$$\beta = \sum_{j=1}^{n} (-1)^{j-1} g_j(x) \, dx_1 \wedge \cdots \wedge d x_{j-1} \wedge d x_{j} \wedge \cdots \wedge d x_n,$$

as in (9.34), and $\varphi(x') = (x', u(x'))$, show that

$$\varphi^* \beta = (-1)^n \left[ \sum_{j=1}^{n-1} g_j(x', u(x')) \frac{\partial u}{\partial x_j} - g_n(x', u(x')) \right] \, dx_1 \wedge \cdots \wedge d x_{n-1}$$

$$= (-1)^{n-1} G \cdot (-\nabla u, 1) \, dx_1 \wedge \cdots \wedge d x_{n-1},$$

where $G = (g_1, \ldots, g_n)$, and verify the identity (9.34) in this case.

*Hint.* For the last part, recall Exercises 2–3 of §8, regarding the orientation of $M$.

6. Let $S$ be a smooth oriented 2-dimensional surface in $\mathbb{R}^3$, and $M$ an open subset of $S$, with smooth boundary; see Fig. 9.1. Let $N$ be the positive unit normal field to $S$, defined by its orientation. For $x \in \partial M$, let $\nu(x)$ be the unit vector, tangent to $M$, normal to $\partial M$, and pointing out of $M$, and let $T$ be the forward unit tangent vector field to $\partial M$. Show that, on $\partial M$,

$$N \times \nu = T, \quad \nu \times T = N.$$

7. If $M$ is an oriented $(n - 1)$-dimensional surface in $\mathbb{R}^n$, with positive unit normal field $N$, show that the volume element $\omega_M$ on $M$ is given by

$$\omega_M = \omega \wedge N,$$
where \( \omega = dx_1 \wedge \cdots \wedge dx_n \) is the standard volume form on \( \mathbb{R}^n \). Deduce that the volume element on the unit sphere \( S^{n-1} \subset \mathbb{R}^n \) is given by

\[
\omega_{S^{n-1}} = \sum_{j=1}^{n} (-1)^{j-1} x_j dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n,
\]

if \( S^{n-1} \) inherits the orientation as the boundary of the unit ball.

8. Let \( M \) be a \( C^k \) surface, \( k \geq 2 \). Suppose \( \varphi : M \to M \) is a \( C^k \) isometry, i.e., it preserves the metric tensor. Taking \( \varphi^* u(x) = u(\varphi(x)) \) for \( u \in C^2(M) \), show that

\[
\Delta \varphi^* u = \varphi^* \Delta u.
\]

**Hint.** The Laplace operator is uniquely specified by the metric tensor on \( M \), via (9.26).

9. Let \( X \) and \( Y \) be smooth vector fields on an open set \( \Omega \subset \mathbb{R}^3 \). Show that

\[
Y \cdot \text{curl} X - X \cdot \text{curl} Y = \text{div}(X \times Y).
\]

10. In the setting of Exercise 9, assume \( \overline{\Omega} \) is compact and smoothly bounded, and that \( X \) and \( Y \) are \( C^1 \) on \( \overline{\Omega} \). Show that

\[
\int_{\Omega} X \cdot \text{curl} Y \, dx = \int_{\Omega} Y \cdot \text{curl} X \, dx,
\]

provided either

(a) \( X \) is normal to \( \partial \Omega \),

or

(b) \( X \) is parallel to \( Y \) on \( \partial \Omega \).

11. Recall the formula (5.25) for the metric tensor of \( \mathbb{R}^n \) in spherical polar coordinates \( R : (0, \infty) \times S^{n-1} \to \mathbb{R}^n \), \( R(r, \omega) = r\omega \). Using (9.26), show that if \( u \in C^2(\mathbb{R}^n) \), then

\[
\Delta u(r\omega) = \frac{\partial^2}{\partial r^2} u(r\omega) + \frac{n-1}{r} \frac{\partial}{\partial r} u(r\omega) + \frac{1}{r^2} \Delta_S u(r\omega),
\]

where \( \Delta_S \) is the Laplace operator on \( S^{n-1} \). Deduce that

\[
u(x) = f(|x|) \implies \Delta u(r\omega) = f''(r) + \frac{n-1}{r} f'(r).
\]

12. Show that

\[|x|^{-(n-2)}\] is harmonic on \( \mathbb{R}^n \setminus 0 \).
In case \( n = 2 \), show that
\[ \log |x| \] is harmonic on \( \mathbb{R}^2 \setminus 0 \).

In Exercise 13, we take \( n \geq 3 \) and consider

\[
Gf(x) = \frac{1}{C_n} \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-2}} \, dy
= \frac{1}{C_n} \int_{\mathbb{R}^n} \frac{f(x-y)}{|y|^{n-2}} \, dy,
\]

with \( C_n = -(n-2)A_{n-1} \).

13. Assume \( f \in C^2_0(\mathbb{R}^n) \). Let \( \Omega_\varepsilon = \mathbb{R}^n \setminus B_\varepsilon \), where \( B_\varepsilon = \{ x \in \mathbb{R}^n : |x| < \varepsilon \} \). Verify that

\[
C_n \Delta Gf(0) = \lim_{\varepsilon \to 0} \int_{\Omega_\varepsilon} \Delta f(x) \cdot |x|^{2-n} \, dx
= \lim_{\varepsilon \to 0} \int_{\Omega_\varepsilon} [\Delta f(x) \cdot |x|^{2-n} - f(x)\Delta |x|^{2-n}] \, dx
= - \lim_{\varepsilon \to 0} \int_{\partial B_\varepsilon} \left[ \varepsilon^{2-n} \frac{\partial f}{\partial r} - (2-n)\varepsilon^{1-n} f \right] dS
= -(n-2)A_{n-1} f(0),
\]

using (9.29) for the third identity. Use this to show that

\[ \Delta Gf(x) = f(x). \]

14. Work out the analogue of Exercise 13 in case \( n = 2 \) and

\[
Gf(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} f(y) \log |x-y| \, dy.
\]
10. Holomorphic functions and harmonic functions

Let $f$ be a complex-valued $C^1$ function on a region $\Omega \subset \mathbb{R}^2$. We identify $\mathbb{R}^2$ with $\mathbb{C}$, via $z = x + iy$, and write $f(z) = f(x, y)$. We say $f$ is holomorphic on $\Omega$ provided it is complex differentiable, in the sense that

$$\lim_{h \to 0} \frac{1}{h} (f(z + h) - f(z)) \text{ exists},$$

for each $z \in \Omega$. When this limit exists, we denote it $f'(z)$, or $df/dz$. An equivalent condition (given $f \in C^1$) is that $f$ satisfies the Cauchy-Riemann equation:

$$(10.1A) \quad \frac{\partial f}{\partial x} = \frac{1}{i} \frac{\partial f}{\partial y}. \quad \frac{\partial f}{\partial y} = \frac{i}{i} \frac{\partial f}{\partial y}.$$

In such a case,

$$(10.1B) \quad f'(z) = \frac{\partial f}{\partial x}(z) = \frac{1}{i} \frac{\partial f}{\partial y}(z).$$

Note that $f(z) = z$ has this property, but $f(z) = \overline{z}$ does not. The following is a convenient tool for producing more holomorphic functions.

**Lemma 10.1.** If $f$ and $g$ are holomorphic on $\Omega$, so is $fg$.

*Proof.* We have

$$(10.2) \quad \frac{\partial}{\partial x} (fg) = \frac{\partial f}{\partial x} g + f \frac{\partial g}{\partial x}, \quad \frac{\partial}{\partial y} (fg) = \frac{\partial f}{\partial y} g + f \frac{\partial g}{\partial y},$$

so if $f$ and $g$ satisfy the Cauchy-Riemann equation, so does $fg$. Note that

$$(10.2A) \quad \frac{d}{dz} (fg)(z) = f'(z)g(z) + f(z)g'(z).$$

Using Lemma 10.1, one can show inductively that if $k \in \mathbb{N}$, $z^k$ is holomorphic on $\mathbb{C}$, and

$$(10.2B) \quad \frac{d}{dz} z^k = kz^{k-1}.$$  

Also, a direct analysis of (10.1) gives this for $k = -1$, on $\mathbb{C} \setminus 0$, and then an inductive argument gives it for each negative integer $k$, on $\mathbb{C} \setminus 0$. The exercises explore various other important examples of holomorphic functions.

Our goal in this section is to show how Green’s theorem can be used to establish basic results about holomorphic functions on domains in $\mathbb{C}$ (and also develop a study of harmonic
functions on domains in $\mathbb{R}^n$). In Theorems 10.2–10.4, $\Omega$ will be a bounded domain with piecewise smooth boundary, and we assume $\Omega$ can be partitioned into a finite number of $C^2$ domains with corners, as defined in §8.

To begin, we apply Green’s theorem to the line integral

$$\int_{\partial \Omega} f \, dz = \int_{\partial \Omega} f \, (dx + idy).$$

Clearly (9.2) applies to complex-valued functions, and if we set $g = if$, we get

$$\int_{\partial \Omega} f \, dz = \iint_{\Omega} \left( i \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} \right) \, dx \, dy. \tag{10.3}$$

Whenever $f$ is holomorphic, the integrand on the right side of (10.3) vanishes, so we have the following result, known as Cauchy’s Integral Theorem:

**Theorem 10.2.** If $f \in C^1(\overline{\Omega})$ is holomorphic, then

$$\int_{\partial \Omega} f(z) \, dz = 0. \tag{10.4}$$

Using (10.4), we can establish Cauchy’s Integral Formula:

**Theorem 10.3.** If $f \in C^1(\overline{\Omega})$ is holomorphic and $z_0 \in \Omega$, then

$$f(z_0) = \frac{1}{2\pi i} \int_{\partial \Omega} \frac{f(z)}{z - z_0} \, dz. \tag{10.5}$$

**Proof.** Note that $g(z) = f(z)/(z - z_0)$ is holomorphic on $\Omega \setminus \{z_0\}$. Let $D_r$ be the disk of radius $r$ centered at $z_0$. Pick $r$ so small that $D_r \subset \Omega$. Then (10.4) implies

$$\int_{\partial \Omega} \frac{f(z)}{z - z_0} \, dz = \int_{\partial D_r} \frac{f(z)}{z - z_0} \, dz. \tag{10.6}$$

To evaluate the integral on the right, parametrize the curve $\partial D_r$ by $\gamma(\theta) = z_0 + re^{i\theta}$. Hence $dz = ire^{i\theta} \, d\theta$, so the integral on the right is equal to

$$\int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{re^{i\theta}} ire^{i\theta} \, d\theta = i \int_0^{2\pi} f(z_0 + re^{i\theta}) \, d\theta. \tag{10.7}$$

As $r \to 0$, this tends in the limit to $2\pi i f(z_0)$, so (10.5) is established.
Suppose \( f \in C^1(\Omega) \) is holomorphic, \( z_0 \in D_r \subset \Omega \), where \( D_r \) is the disk of radius \( r \) centered at \( z_0 \), and suppose \( z \in D_r \). Then Theorem 10.3 implies

\[
(10.8) \quad f(z) = \frac{1}{2\pi i} \int_{\partial\Omega} \frac{f(\zeta)}{(\zeta - z_0) - (z - z_0)} \, d\zeta.
\]

We have the infinite series expansion

\[
(10.9) \quad \frac{1}{(\zeta - z_0) - (z - z_0)} = \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \left( \frac{z - z_0}{\zeta - z_0} \right)^n,
\]

valid as long as \(|z - z_0| < |\zeta - z_0|\). Hence, given \(|z - z_0| < r\), this series is uniformly convergent for \( \zeta \in \partial\Omega \), and we have

\[
(10.10) \quad f(z) = \frac{1}{2\pi i} \sum_{n=0}^{\infty} \int_{\partial\Omega} \frac{f(\zeta)}{\zeta - z_0} \left( \frac{z - z_0}{\zeta - z_0} \right)^n \, d\zeta.
\]

We summarize what has been established.

**Theorem 10.4.** Given \( f \in C^1(\Omega) \), holomorphic on \( \Omega \) and a disk \( D_r \subset \Omega \) as above, for \( z \in D_r \), \( f(z) \) has the convergent power series expansion

\[
(10.11) \quad f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n, \quad a_n = \frac{1}{2\pi i} \int_{\partial\Omega} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} \, d\zeta.
\]

Note that, when (10.5) is applied to \( \Omega = D_r \), the disk of radius \( r \) centered at \( z_0 \), the computation (10.7) yields

\[
(10.12) \quad f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) \, d\theta = \frac{1}{\ell(\partial D_r)} \int_{\partial D_r} f(z) \, ds(z),
\]

when \( f \) is holomorphic and \( C^1 \) on \( D_r \), and \( \ell(\partial D_r) = 2\pi r \) is the length of the circle \( \partial D_r \). This is a mean value property, which extends to harmonic functions on domains in \( \mathbb{R}^n \), as we will see below.

Note that we can write (10.1) as \((\partial_x + i\partial_y)f = 0\); applying the operator \( \partial_x - i\partial_y \) to this gives

\[
(10.13) \quad \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0
\]

for any holomorphic function. A general \( C^2 \) solution to (10.13) on a region \( \Omega \subset \mathbb{R}^2 \) is called a harmonic function. More generally, if \( \mathcal{O} \) is an open set in \( \mathbb{R}^n \), a function \( f \in C^2(\mathcal{O}) \) is called harmonic if \( \Delta f = 0 \) on \( \mathcal{O} \), where, as in (9.27),

\[
(10.14) \quad \Delta f = \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}.
\]

Generalizing (10.12), we have the following, known as the mean value property of harmonic functions:
Proposition 10.5. Let $\Omega \subset \mathbb{R}^n$ be open, $u \in C^2(\Omega)$ be harmonic, $p \in \Omega$, and $B_R(p) = \{x \in \Omega : |x - p| \leq R\} \subset \Omega$. Then

\begin{equation}
(10.15) \quad u(p) = \frac{1}{A(\partial B_R(p))} \int_{\partial B_R(p)} u(x) \, dS(x).
\end{equation}

For the proof, set

\begin{equation}
(10.16) \quad \psi(r) = \frac{1}{A(S^{n-1})} \int_{S^{n-1}} u(p + r\omega) \, dS(\omega),
\end{equation}

for $0 < r \leq R$. We have $\psi(R)$ equal to the right side of (10.15), while clearly $\psi(r) \to u(p)$ as $r \to 0$. Now

\begin{equation}
(10.17) \quad \psi'(r) = \frac{1}{A(S^{n-1})} \int_{S^{n-1}} \omega \cdot \nabla u(p + r\omega) \, dS(\omega) = \frac{1}{A(\partial B_r(p))} \int_{\partial B_r(p)} \frac{\partial u}{\partial \nu} \, dS(x).
\end{equation}

At this point, we establish:

Lemma 10.6. If $\mathcal{O} \subset \mathbb{R}^n$ is a bounded domain with smooth boundary and $u \in C^2(\mathcal{O})$ is harmonic in $\mathcal{O}$, then

\begin{equation}
(10.18) \quad \int_{\partial \mathcal{O}} \frac{\partial u}{\partial \nu}(x) \, dS(x) = 0.
\end{equation}

Proof. Apply the Green formula (9.29), with $M = \mathcal{O}$ and $v = 1$. If $\Delta u = 0$, every integrand in (9.29) vanishes, except the one appearing in (10.18), so this integrates to zero.

It follows from this lemma that (10.17) vanishes, so $\psi(r)$ is constant. This completes the proof of (10.15).

We can integrate the identity (10.15), to obtain

\begin{equation}
(10.19) \quad u(p) = \frac{1}{V(B_R(p))} \int_{B_R(p)} u(x) \, dV(x),
\end{equation}

where $u \in C^2(B_R(p))$ is harmonic. This is another expression of the mean value property.

The mean value property of harmonic functions has a number of important consequences. Here we mention one result, known as Liouville’s Theorem.
Proposition 10.7. If \( u \in C^2(\mathbb{R}^n) \) is harmonic on all of \( \mathbb{R}^n \) and bounded, then \( u \) is constant.

Proof. Pick any two points \( p, q \in \mathbb{R}^n \). We have, for any \( r > 0 \),

\[
(10.20) \quad u(p) - u(q) = \frac{1}{V(B_r(0))} \left[ \int_{B_r(p)} u(x) \, dx - \int_{B_r(q)} u(x) \, dx \right].
\]

Note that \( V(B_r(0)) = C_n r^n \), where \( C_n \) is evaluated in problem 2 of §5. Thus

\[
(10.21) \quad |u(p) - u(q)| \leq \frac{C_n}{r^n} \int_{\Delta(p,q,r)} |u(x)| \, dx,
\]

where

\[
(10.22) \quad \Delta(p,q,r) = B_r(p) \Delta B_r(q) = \left( B_r(p) \setminus B_r(q) \right) \cup \left( B_r(q) \setminus B_r(p) \right).
\]

Note that, if \( a = |p - q| \), then \( \Delta(p,q,r) \subset B_r + a(p) \setminus B_r - a(p) \); hence

\[
(10.23) \quad V(\Delta(p,q,r)) \leq C(p,q) r^{n-1}, \quad r \geq 1.
\]

It follows that, if \( |u(x)| \leq M \) for all \( x \in \mathbb{R}^n \), then

\[
(10.24) \quad |u(p) - u(q)| \leq M C_n C(p,q) r^{-1}, \quad \forall r \geq 1.
\]

Taking \( r \to \infty \), we obtain \( u(p) - u(q) = 0 \), so \( u \) is constant.

We will now use Liouville’s Theorem to prove the Fundamental Theorem of Algebra:

Theorem 10.8. If \( p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \) is a polynomial of degree \( n \geq 1 \) \( (a_n \neq 0) \), then \( p(z) \) must vanish somewhere in \( \mathbb{C} \).

Proof. Consider

\[
(10.25) \quad f(z) = \frac{1}{p(z)}.
\]

If \( p(z) \) does not vanish anywhere in \( \mathbb{C} \), then \( f(z) \) is holomorphic on all of \( \mathbb{C} \). (See Exercise 9 below.) On the other hand,

\[
(10.26) \quad f(z) = \frac{1}{z^n} \frac{1}{a_n + a_{n-1} z^{-1} + \cdots + a_0 z^{-n}},
\]

so

\[
(10.27) \quad |f(z)| \to 0, \quad \text{as} \quad |z| \to \infty.
\]
Thus \( f \) is bounded on \( \mathbb{C} \), if \( p(z) \) has no roots. By Proposition 10.7, \( f(z) \) must be constant, which is impossible, so \( p(z) \) must have a complex root.

From the fact that every holomorphic function \( f \) on \( \mathcal{O} \subset \mathbb{R}^2 \) is harmonic, it follows that its real and imaginary parts are harmonic. This result has a converse. Let \( u \in C^2(\mathcal{O}) \) be harmonic. Consider the 1-form

\[
\alpha = -\frac{\partial u}{\partial y} \, dx + \frac{\partial u}{\partial x} \, dy.
\]

We have \( d\alpha = -(\Delta u)\,dx \wedge dy \), so \( \alpha \) is closed if and only if \( u \) is harmonic. Now, if \( \mathcal{O} \) is diffeomorphic to a disk, it follows from Proposition 8.3 that \( \alpha \) is exact on \( \mathcal{O} \), whenever it is closed, so, in such a case,

\[
\Delta u = 0 \text{ on } \mathcal{O} \implies \exists \, v \in C^1(\mathcal{O}) \text{ s.t. } \alpha = dv.
\]

In other words,

\[
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.
\]

This is precisely the Cauchy-Riemann equation (10.1) for \( f = u + iv \), so we have:

**Proposition 10.9.** If \( \mathcal{O} \subset \mathbb{R}^2 \) is diffeomorphic to a disk and \( u \in C^2(\mathcal{O}) \) is harmonic, then \( u \) is the real part of a holomorphic function on \( \mathcal{O} \).

The function \( v \) (which is unique up to an additive constant) is called the *harmonic conjugate* of \( u \).

We close this section with a brief mention of holomorphic functions on a domain \( \mathcal{O} \subset \mathbb{C}^n \). We say \( f \in C^1(\mathcal{O}) \) is holomorphic provided it satisfies

\[
\frac{\partial f}{\partial x_j} = \frac{1}{i} \frac{\partial f}{\partial y_j}, \quad 1 \leq j \leq n.
\]

Suppose \( z \in \mathcal{O} \), \( z = (z_1, \ldots, z_n) \). Suppose \( \zeta \in \mathcal{O} \) whenever \( |z - \zeta| < r \). Then, by successively applying Cauchy’s integral formula (10.5) to each complex variable \( z_j \), we have that

\[
f(z) = (2\pi)^{-n} \int_{\gamma_n} \cdots \int_{\gamma_1} f(\zeta)(\zeta_1 - z_1)^{-1} \cdots (\zeta_n - z_n)^{-1} \, d\zeta_1 \cdots d\zeta_n,
\]

where \( \gamma_j \) is any simple counterclockwise curve about \( z_j \) in \( \mathbb{C} \) with the property that \( |\zeta_j - z_j| < r/\sqrt{n} \) for all \( \zeta_j \in \gamma_j \).

Consequently, if \( p \in \mathbb{C}^n \) and \( \mathcal{O} \) contains the “polydisc”

\[
\overline{D} = \{ z \in \mathbb{C}^n : |z_j - p_j| \leq \delta, \, \forall j \},
\]
then, for \( z \in D \), the interior of \( \overline{D} \), we have

\[
f(z) = (2\pi i)^{-n} \int_{C_n} \cdots \int_{C_1} f(\zeta) \frac{1}{[(\zeta_n - p_n) - (z_n - p_n)]^{-1}} \, d\zeta_1 \cdots d\zeta_n,
\]

where \( C_j = \{ \zeta \in \mathbb{C} : |\zeta - p_j| = \delta \} \). Then, parallel to (10.8)–(10.11), we have

\[
f(z) = \sum_{\alpha \geq 0} c_\alpha (z - p)^\alpha,
\]

for \( z \in D \), where \( \alpha = (\alpha_1, \ldots, \alpha_n) \) is a multi-index, \( (z - p)^\alpha = (z_1 - p_1)^{\alpha_1} \cdots (z_n - p_n)^{\alpha_n} \), as in (1.13), and

\[
c_\alpha = (2\pi i)^{-n} \int_{C_n} \cdots \int_{C_1} f(\zeta) (\zeta_1 - p_1)^{-\alpha_1 - 1} \cdots (\zeta_n - p_n)^{-\alpha_n - 1} \, d\zeta_1 \cdots d\zeta_n.
\]

Thus holomorphic functions on open domains in \( \mathbb{C}^n \) have convergent power series expansions.

We refer to [Ahl], [Hil], and [T6] for more material on holomorphic functions of one complex variable, and to [Kr] for material on holomorphic functions of several complex variables. A source of much information on harmonic functions is [Kel]. Also further material on these subjects can be found in [T].

**Exercises**

1. Let \( f_k : \Omega \to \mathbb{C} \) be holomorphic on an open set \( \Omega \subset \mathbb{C} \). Assume \( f_k \to f \) and \( \nabla f_k \to \nabla f \) locally uniformly in \( \Omega \). Show that \( f : \Omega \to \mathbb{C} \) is holomorphic.

2. Assume \n
\[
f(z) = \sum_{k=0}^{\infty} a_k z^k
\]

is absolutely convergent for \( |z| < R \). Deduce from Proposition 1.10 and Exercise 1 above that \( f \) is holomorphic on \( |z| < R \), and that

\[
f'(z) = \sum_{k=1}^{\infty} ka_k z^{k-1}, \quad \text{for } |z| < R.
\]
3. As in (3.89), the exponential function $e^z$ is defined by

\begin{equation}
(10.38) \quad e^z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k.
\end{equation}

Deduce from Exercise 2 that $e^z$ is holomorphic in $z$.

4. By (3.90), we have

\[ e^{z+h} = e^z e^h, \quad \forall z, h \in \mathbb{C}. \]

Use this to show directly from (10.1) that $e^z$ is complex differentiable and $(d/dz)e^z = e^z$ on $\mathbb{C}$, giving another proof that $e^z$ is holomorphic on $\mathbb{C}$.

*Hint.* Use the power series for $e^h$ to show that

\[ \lim_{h \to 0} \frac{e^h - 1}{h} = 1. \]

5. For another approach to the fact that $e^z$ is holomorphic, use

\[ e^z = e^x e^{iy} \]

and (3.89) to verify that $e^z$ satisfies the Cauchy-Riemann equation.

6. For $z \in \mathbb{C}$, set

\begin{equation}
(10.39) \quad \cos z = \frac{1}{2} (e^{iz} + e^{-iz}), \quad \sin z = \frac{1}{2i} (e^{iz} - e^{-iz}).
\end{equation}

Show that these functions agree with the definitions of $\cos t$ and $\sin t$ given in (3.91)--(3.92), for $z = t \in \mathbb{R}$. Show that $\cos z$ and $\sin z$ are holomorphic in $z \in \mathbb{C}$. Show that

\begin{equation}
(10.40) \quad \frac{d}{dz} \cos z = -\sin z, \quad \frac{d}{dz} \sin z = \cos z,
\end{equation}

and

\begin{equation}
(10.41) \quad \cos^2 z + \sin^2 z = 1,
\end{equation}

for all $z \in \mathbb{C}$.

7. Let $O, \Omega$ be open in $\mathbb{C}$. If $f$ is holomorphic on $O$, with range in $\Omega$, and $g$ is holomorphic on $\Omega$, show that $h = g \circ f$ is holomorphic on $O$, and $h'(z) = g'(f(z))f'(z)$.

*Hint.* See the proof of the chain rule in §1.
8. Let \( \Omega \subset \mathbb{C} \) be a connected open set and let \( f \) be holomorphic on \( \Omega \).

(a) Show that if \( f(z_j) = 0 \) for distinct \( z_j \in \Omega \) and \( z_j \to z_0 \in \Omega \), then \( f(z) = 0 \) for \( z \) in a neighborhood of \( z_0 \).

*Hint.* Use the power series expansion (10.1).

(b) Show that if \( f = 0 \) on a nonempty open set \( \mathcal{O} \subset \Omega \), then \( f \equiv 0 \) on \( \Omega \).

*Hint.* Let \( U \subset \Omega \) denote the interior of the set of points where \( f \) vanishes. Use part (a) to show that \( \overline{U} \cap \Omega \) is open.

9. Let \( \Omega = \mathbb{C} \setminus (-\infty, 0] \) and define \( \log : \Omega \to \mathbb{C} \) by

\[
(10.42) \quad \log z = \int_{\gamma_z} \frac{1}{\zeta} d\zeta,
\]

where \( \gamma_z \) is a path from 1 to \( z \) in \( \Omega \). Use Theorem 10.2 to show that this is independent of the choice of such path. Show that it yields a holomorphic function on \( \mathbb{C} \setminus (-\infty, 0] \), satisfying

\[
\frac{d}{dz} \log z = \frac{1}{z}, \quad z \in \mathbb{C} \setminus (-\infty, 0].
\]

10. Taking \( \log z \) as in Exercise 9, show that

\[
(10.43) \quad e^{\log z} = z, \quad \forall z \in \mathbb{C} \setminus (-\infty, 0].
\]

*Hint.* If \( \varphi(z) \) denotes the left side, show that \( \varphi(1) = 1 \) and \( \varphi'(z) = \varphi(z)/z \). Use uniqueness results from \( \S 3 \) to deduce that \( \varphi(x) = x \) for \( x \in (0, \infty) \), and from there deduce that \( \varphi(z) \equiv z \), using Exercise 8.

*Alternative.* Apply \( d/dz \) to show that

\[
\log e^z = z,
\]

for \( z \) in some neighborhood of 0. Deduce from this (and Exercise 3 of \( \S 2 \)) that (10.43) holds for \( z \) in some neighborhood of 1. Then get it for all \( z \in \mathbb{C} \setminus (-\infty, 0] \) using Exercise 8.

11. With \( \Omega = \mathbb{C} \setminus (-\infty, 0] \) as in Exercise 9, and \( a \in \mathbb{C} \), define \( z^a \) for \( z \in \Omega \) by

\[
(10.44) \quad z^a = e^{a \log z}.
\]

Show that this is holomorphic on \( \Omega \) and

\[
(10.45) \quad \frac{d}{dz} z^a = az^{a-1}, \quad z^a z^b = z^{a+b}, \quad \forall z \in \Omega.
\]
12. Let $\mathcal{O} = \mathbb{C} \setminus \{[1, \infty) \cup (-\infty, -1]\}$, and define $A_s : \mathcal{O} \rightarrow \mathbb{C}$ by

$$A_s(z) = \int_{\sigma_z} (1 - \zeta^2)^{-1/2} d\zeta,$$

where $\sigma_z$ is a path from 0 to $z$ in $\mathcal{O}$. Show that this is independent of the choice of such a path, and that it yields a holomorphic function on $\mathcal{O}$.

13. With $A_s$ as in Exercise 12, show that

$$A_s(\sin z) = z,$$

for $z$ in some neighborhood of 0. (Hint. Apply $d/dz$.) From here, show that

$$\sin(A_s(z)) = z, \quad \forall z \in \mathcal{O}.$$

Thus we write

$$(10.46) \quad \arcsin z = \int_0^z (1 - \zeta^2)^{-1/2} d\zeta.$$

Compare (3.97).

14. Look again at Exercise 4 in §1.

15. Look again at Exercises 3–5 in §2. Write the result as an inverse function theorem for holomorphic maps.

16. Differentiate (10.5) to show that, in the setting of Theorem 10.3, for $k \in \mathbb{N}$,

$$(10.47) \quad f^{(k)}(z_0) = \frac{k!}{2\pi i} \int_{\partial\Omega} \frac{f(z)}{(z - z_0)^{k+1}} dz.$$

Show that this also follows from (10.11).

17. Assume $f$ is holomorphic on $\mathbb{C}$, and set

$$M(z_0, R) = \sup_{|z - z_0| \leq R} |f(z)|.$$

Use the $k = 1$ case of Exercise 16 to show that

$$|f'(z_0)| \leq \frac{M(z_0, R)}{R}, \quad \forall R \in (0, \infty).$$
18. In the setting of Exercise 17, assume \( f \) is bounded, say \( |f(z)| \leq M \) for all \( z \in \mathbb{C} \). Deduce that \( f'(z_0) = 0 \) for all \( z_0 \in \mathbb{C} \), and in that way obtain another proof of Liouville’s theorem, in the setting of holomorphic functions on \( \mathbb{C} \). (Note that Proposition 10.7 is more general.)

The next four exercises deal with the function

\[
G(z) = \int_{-\infty}^{\infty} e^{-t^2 + tz} \, dt, \quad z \in \mathbb{C}.
\]

19. Show that \( G \) is continuous on \( \mathbb{C} \).

20. Show that \( G \) is holomorphic on \( \mathbb{C} \), with

\[
G'(z) = \int_{-\infty}^{\infty} te^{-t^2 + tz} \, dt.
\]

*Hint.* Write

\[
\frac{1}{h}[G(z + h) - G(z)] = \int_{-\infty}^{\infty} e^{-t^2 + tz} \frac{1}{h}(e^{th} - 1) \, dt,
\]

and

\[
\frac{1}{h}(e^{th} - 1) = t + \frac{1}{h}R(th),
\]

where

\[
e^w = 1 + w + R(w), \quad |R(w)| \leq C|w|^2 e^{|w|},
\]

so

\[
\left| \frac{1}{h}R(th) \right| \leq Ct^2 |h| e^{|th|}.
\]

21. Show that, for \( x \in \mathbb{R} \),

\[
G(x) = \sqrt{\pi}e^{x^2/4}.
\]

*Hint.* Write

\[
G(x) = e^{x^2/4} \int_{-\infty}^{\infty} e^{-(t-x/2)^2} \, dt,
\]

and make a change of variable in the integral.

22. Deduce from Exercises 21 and 8 that

\[
G(z) = \sqrt{\pi} e^{z^2/4}, \quad \forall z \in \mathbb{C}.
\]

\[
(10.49)
\]
The next exercises deal with the Gamma function,

(10.50) \[ \Gamma(z) = \int_0^\infty e^{-s} s^{z-1} ds, \]

defined for \( z > 0 \) in (5.32).

23. Show that the integral is absolutely convergent for \( \text{Re} \, z > 0 \) and defines \( \Gamma(z) \) as a holomorphic function on \( \{ z \in \mathbb{C} : \text{Re} \, z > 0 \} \).

24. Extend the identity (5.35), i.e.,

(10.51) \[ \Gamma(z + 1) = z \Gamma(z), \]

to \( \text{Re} \, z > 0 \).

25. Use (10.51) to extend \( \Gamma \) to be holomorphic on \( \mathbb{C} \setminus \{0, -1, -2, -3, \ldots \} \).

26. Use the result of Exercise 16 to show that if \( f_\nu \) are holomorphic on an open set \( \Omega \subset \mathbb{C} \) and \( f_\nu \to f \) uniformly on compact subsets of \( \Omega \), then \( f \) is holomorphic on \( \Omega \) and \( f_\nu' \to f' \) uniformly on compact subsets.

27. The Riemann zeta function \( \zeta(z) \) is defined for \( \text{Re} \, z > 1 \) by

(10.52) \[ \zeta(z) = \sum_{k=1}^\infty \frac{1}{k^z}. \]

Show that \( \zeta(z) \) is holomorphic on \( \{ z \in \mathbb{C} : \text{Re} \, z > 1 \} \).

The following exercises deal with harmonic functions on domains in \( \mathbb{R}^n \).

28. Using the formula (9.26) for the Laplace operator together with the formula (5.25) for the metric tensor on \( \mathbb{R}^n \) in spherical polar coordinates \( x = r \omega, \, x \in \mathbb{R}^n, \, r = |x|, \, \omega \in S^{n-1} \), show that if \( u \in C^2(\Omega), \, \Omega \subset \mathbb{R}^n \),

(10.53) \[ \Delta u(r\omega) = \frac{\partial^2}{\partial r^2} u(r\omega) + \frac{n-1}{r} \frac{\partial}{\partial r} u(r\omega) + \frac{1}{r^2} \Delta_S u(r\omega), \]

where \( \Delta_S \) is the Laplace operator on \( S^{n-1} \).

29. If \( f(x) = \varphi(|x|) \) on \( \mathbb{R}^n \), show that

(10.54) \[ \Delta f(x) = \varphi''(|x|) + \frac{n-1}{|x|} \varphi'(|x|). \]
In particular, show that

\[(10.55) \quad |x|^{-(n-2)} \text{ is harmonic on } \mathbb{R}^n \setminus 0,\]

if \( n \geq 3 \), and

\[(10.56) \quad \log |x| \text{ is harmonic on } \mathbb{R}^2 \setminus 0.\]

If \( \mathcal{O}, \Omega \) are open in \( \mathbb{R}^n \), a smooth map \( \varphi : \mathcal{O} \to \Omega \) is said to be \textit{conformal} provided the matrix function \( G(x) = D\varphi(x)^t D\varphi(x) \) is a multiple of the identity, \( G(x) = \gamma(x) I \). Recall formula (5.2).

30. Suppose \( n = 2 \) and \( \varphi \) preserves orientation. Show that \( \varphi \) (pictured as a function \( \varphi : \mathcal{O} \to \mathbb{C} \)) is conformal \textit{if and only if} it is holomorphic. If \( \varphi \) reverses orientation, \( \varphi \) is conformal \( \Leftrightarrow \bar{\varphi} \) is holomorphic (we say \( \varphi \) is anti-holomorphic).

31. If \( \mathcal{O} \) and \( \Omega \) are open in \( \mathbb{R}^2 \) and \( u \) is harmonic on \( \Omega \), show that \( u \circ \varphi \) is harmonic on \( \mathcal{O} \), whenever \( \varphi : \mathcal{O} \to \Omega \) is a smooth conformal map.
\textit{Hint.} Use Exercise 7 and Proposition 10.9.

The following exercises will present an alternative approach to the proof of Proposition 10.5 (the mean value property of harmonic functions). For this, let \( B_R = \{ x \in \mathbb{R}^n : |x| \leq R \} \). Assume \( u \) is continuous on \( B_R \) and \( C^2 \) and harmonic on the interior \( \overset{0}{B}_R \). We assume \( n \geq 2 \).

32. Given \( g \in SO(n) \), show that \( u_g(x) = u(gx) \) is harmonic on \( \overset{0}{B}_R \).
\textit{Hint.} See Exercise 7 of §9.

33. As in Exercise 24 of §5, define \( A \) by\n
\[ A u(x) = \int_{SO(n)} u(gx) \, dg. \]

Thus \( A u(x) \) is a radial function:

\[ A u(x) = S u(|x|), \quad S u(r) = \frac{1}{A_{n-1}} \int_{S^{n-1}} u(r \omega) \, dS(\omega). \]

Deduce from Exercise 32 above that \( A u \) is harmonic on \( \overset{0}{B}_R \).

34. Use Exercise 29 to show that \( \varphi'(r) = S u(r) \) satisfies

\[ \varphi''(r) + \frac{n-1}{r} \varphi'(r) = 0, \]
for \( r \in (0, R) \). Deduce from this differential equation that there exist constants \( C_0 \) and \( C_1 \) such that
\[
\varphi(r) = C_0 + C_1 r^{-(n-2)}, \quad \text{if } n \geq 3,
\]
\[
C_0 + C_1 \log r, \quad \text{if } n = 2.
\]
Then show that, since \( Au(x) \) does not blow up at \( x = 0 \), \( C_1 = 0 \). Hence
\[
Au(x) = C_0, \quad \forall x \in B_R.
\]

35. Note that \( Au(0) = u(0) \). Deduce that for each \( r \in (0, R] \),
\[
(10.57) \quad u(0) = Su(r) = \frac{1}{A_{n-1}} \int_{\mathbb{S}^{n-1}} u(r\omega) dS(\omega).
\]
11. Homotopies of maps and actions on forms

Let $X$ and $Y$ be smooth surfaces. Two smooth maps $f_0, f_1 : X \to Y$ are said to be smoothly homotopic provided there is a smooth $F : [0, 1] \times X \to Y$ such that $F(0, x) = f_0(x)$ and $F(1, x) = f_1(x)$. The following result illustrates the significance of maps being homotopic.

**Proposition 11.1.** Assume $X$ is a compact, oriented, $k$-dimensional surface and $\alpha \in \Lambda^k(Y)$ is closed, i.e., $d\alpha = 0$. If $f_0, f_1 : X \to Y$ are smoothly homotopic, then

\[
\int_X f_0^* \alpha = \int_X f_1^* \alpha.
\]

In fact, with $[0, 1] \times X = \overline{\Omega}$, this is a special case of the following.

**Proposition 11.2.** Assume $\overline{\Omega}$ is a smoothly bounded, compact, oriented $(k+1)$-dimensional surface, and $\alpha \in \Lambda^k(Y)$ is closed. If $F : \overline{\Omega} \to Y$ is a smooth map, then

\[
\int_{\partial \overline{\Omega}} F^* \alpha = 0.
\]

**Proof.** Stokes' theorem gives

\[
\int_{\partial \overline{\Omega}} F^* \alpha = \int_\Omega dF^* \alpha = 0,
\]

since $dF^* \alpha = F^* d\alpha$ and, by hypothesis, $d\alpha = 0$.

Proposition 11.2 is one generalization of Proposition 11.1. Here is another.

**Proposition 11.3.** Assume $X$ is a $k$-dimensional surface and $\alpha \in \Lambda^\ell(Y)$ is closed. If $f_0, f_1 : X \to Y$ are smoothly homotopic, then $f_0^* \alpha - f_1^* \alpha$ is exact, i.e.,

\[
f_0^* \alpha - f_1^* \alpha = d\beta,
\]

for some $\beta \in \Lambda^{\ell-1}(X)$.

**Proof.** Take a smooth $F : \mathbb{R} \times X \to Y$ such that $F(j, x) = f_j(x)$. Consider

\[
\tilde{\alpha} = F^* \alpha \in \Lambda^\ell(\mathbb{R} \times X).
\]

Note that $d\tilde{\alpha} = F^* d\alpha = 0$. Now consider

\[
\Phi_s : \mathbb{R} \times X \to \mathbb{R} \times X, \quad \Phi_s(t, x) = (s + t, x).
\]
We claim that

\[(11.7) \quad \tilde{\alpha} - \Phi_1^* \tilde{\alpha} = d\tilde{\beta}, \]

for some \( \tilde{\beta} \in \Lambda^{\ell-1}(\mathbb{R} \times X) \). Now take

\[(11.8) \quad \beta = j^* \tilde{\beta}, \quad j : X \to \mathbb{R} \times X, \quad j(x) = (0, x). \]

We have \( F \circ j = f_0, \ F \circ \Phi_1 \circ j = f_1 \), so it follows that

\[(11.9) \quad f_0^* \alpha - f_1^* \alpha = j^* \tilde{\alpha} - \Phi_1^* \tilde{\alpha} = j^* d\tilde{\beta}, \]

given (11.7), which yields (11.4) with \( s \) as in (11.8).

It remains to prove (11.7), under the hypothesis that \( d\tilde{\alpha} = 0 \). The following result gives this. The formula (11.10) uses the interior product, defined by (7.4)–(7.5).

**Lemma 11.4.** If \( \tilde{\alpha} \in \Lambda^\ell(\mathbb{R} \times X) \) and \( \Phi_s \) is as in (11.6), then

\[(11.10) \quad \frac{d}{ds} \Phi_s^* \tilde{\alpha} = \Phi_s^* \left( d(\tilde{\alpha}) | \partial_t \right) + (d\tilde{\alpha}) | \partial_t \).

Hence, if \( d\tilde{\alpha} = 0 \), (11.7) holds with

\[(11.11) \quad \tilde{\beta} = -\int_0^1 (\Phi_s^* \tilde{\alpha}) | \partial_t \ ds. \]

**Proof.** Since \( \Phi_{s+\sigma} = \Phi_s^* \Phi_\sigma^* = \Phi_\sigma^* \Phi_s^* \), it suffices to show that (11.10) holds at \( s = 0 \). It also suffices to work in local coordinates on \( X \). Say

\[(11.12) \quad \tilde{\alpha} = \sum_i \alpha_i^i(t, x) \ dx_{i_1} \wedge \cdots \wedge dx_{i_\ell} \]

\[+ \sum_j \alpha_j^b(t, x) \ dt \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_{\ell-1}}. \]

We have \( \Phi_s^* \tilde{\alpha} \) given by a similar formula, with coefficients replaced by \( \alpha_i^i(t+s, x) \) and \( \alpha_j^b(t+s, x) \), hence

\[(11.13) \quad \frac{d}{ds} \Phi_s^* \tilde{\alpha}|_{s=0} = \sum_i \partial_t \alpha_i^i(t, x) \ dx_{i_1} \wedge \cdots \wedge dx_{i_\ell} \]

\[+ \sum_j \partial_t \alpha_j^b(t, x) \ dt \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_{\ell-1}}. \]
Meanwhile

\[(11.14)\quad \tilde{\alpha} \partial_t = \sum_j \alpha_j^b(t, x) \, dx_j \wedge \cdots \wedge dx_{j-1}, \]

so

\[(11.15)\quad d(\tilde{\alpha} \partial_t) = \sum_j \partial_t \alpha_j^b(t, x) \, dt \wedge dx_j \wedge \cdots \wedge dx_{j-1} \]

+ \sum_{j, \nu} \partial_x \nu \alpha_j^b(t, x) \, dx_\nu \wedge dx_j \wedge \cdots \wedge dx_{j-1}.

A similar calculation yields

\[(11.16)\quad (d\tilde{\alpha}) \partial_t = \sum_i \partial_t \alpha_i^\#(t, x) \, dx_i \wedge \cdots \wedge dx_i \]

- \sum_{j, \nu} \partial_x \nu \alpha_j^b(t, x) \, dx_\nu \wedge dx_j \wedge \cdots \wedge dx_{j-1}.

Comparison of (11.15)–(11.16) with (11.13) yields (11.10) at \(s = 0\), proving Lemma 11.4.

The following consequence of Proposition 11.3 contains the Poincaré lemma.

**Proposition 11.5.** Let \(X\) be a smooth \(k\)-dimensional surface. Assume the identity map \(I : X \to X\) is smoothly homotopic to a constant map \(K : X \to X\), satisfying \(K(x) \equiv p\). Then, for all \(\ell \in \{1, \ldots, k\}\),

\[(11.17)\quad \alpha \in \Lambda^\ell(X), \; d\alpha = 0 \implies \alpha \text{ is exact.}\]

**Proof.** By Proposition 11.3, \(\alpha - K^*\alpha\) is exact. However, \(K^*\alpha = 0\).

Proposition 11.5 applies to any open \(X \subset \mathbb{R}^k\) that is star-shaped, so

\[(11.18)\quad D_s : X \to X \quad \text{for} \quad s \in [0, 1], \quad D_s(x) = sx.\]

Thus, for any open star-shaped \(X \subset \mathbb{R}^k\), each closed \(\alpha \in \Lambda^\ell(X)\) is exact.

We next present an important generalization of Lemma 11.4. Let \(\Omega\) be a smooth \(n\)-dimensional surface. If \(\alpha \in \Lambda^k(\Omega)\) and \(X\) is a vector field on \(\Omega\), generating a flow \(\mathcal{F}_X^t\), the **Lie derivative** \(L_X\alpha\) is defined to be

\[(11.19)\quad L_X\alpha = \left. \frac{d}{dt} (\mathcal{F}_X^t)^*\alpha \right|_{t=0}.\]

Note the similarity to the definition (3.77) of \(L_XY\) for a vector field \(Y\), for which there was the alternative formula (3.80). The following useful result is known as Cartan’s formula for the Lie derivative.
Proposition 11.6. We have

\begin{equation}
\mathcal{L}_X \alpha = d(\alpha \lvert X) + (d\alpha \lvert X).
\end{equation}

Proof. We can assume \( \Omega \) is an open subset of \( \mathbb{R}^n \). First we compare both sides in the special case \( X = \partial / \partial x_\ell = \partial_\ell \). Note that

\begin{equation}
(F^{t})^\ast \alpha = \sum_j a_j(x + t e_\ell) \, dx_{j_1} \wedge \cdots \wedge dx_{j_k},
\end{equation}

so

\begin{equation}
\mathcal{L}_{\partial_\ell} \alpha = \sum_j \partial_{x_\ell} a_j(x) \, dx_{j_1} \wedge \cdots \wedge dx_{j_k} = \partial_\ell \alpha.
\end{equation}

To evaluate the right side of (11.21), with \( X = \partial_\ell \), we could parallel the calculation (11.14)\textendash(11.16). Alternatively, we can use (7.12) to write this as

\begin{equation}
d(\iota_\ell \alpha) + \iota_\ell d\alpha = \sum_{j=1}^n (\partial_j \wedge \iota_\ell + \iota_\ell \partial_j \wedge j) \alpha.
\end{equation}

Using the commutativity of \( \partial_j \) with \( \wedge \) and with \( \iota_\ell \), and the anticommutativity relations (7.8), we see that the right side of (11.23) is \( \partial_\ell \alpha \), which coincides with (11.22). Thus the proposition holds for \( X = \partial / \partial x_\ell \).

Now we prove the proposition in general, for a smooth vector field \( X \) on \( \Omega \). It is to be verified at each point \( x_0 \in \Omega \). If \( X(x_0) \neq 0 \), we can apply Theorem 3.7 to choose a coordinate system about \( x_0 \) so \( X = \partial / \partial x_1 \) and use the calculation above. This shows that the desired identity holds on the set \( \{ x_0 \in \Omega : X(x_0) \neq 0 \} \), and by continuity it holds on the closure of this set. However, if \( x_0 \) has a neighborhood on which \( X \) vanishes, it is clear that \( \mathcal{L}_X \alpha = 0 \) near \( x_0 \) and also \( \alpha \lvert X \) and \( d\alpha \lvert X \) vanish near \( x_0 \). This completes the proof.

From (11.19) and the identity \( F^{s+t}_X = F^s_X F^t_X \), it follows that

\begin{equation}
\frac{d}{dt} (F^t_X)^\ast \alpha = \mathcal{L}_X (F^t_X)^\ast \alpha = (F^t_X)^\ast \mathcal{L}_X \alpha.
\end{equation}

It is useful to generalize this. Let \( F_t \) be a smooth family of diffeomorphisms of \( M \) into \( M \). Define vector fields \( X_t \) on \( F_t(M) \) by

\begin{equation}
\frac{d}{dt} F_t(x) = X_t(F_t(x)).
\end{equation}

Then, given \( \alpha \in \Lambda^k(M) \),

\begin{equation}
\frac{d}{dt} F_t^\ast \alpha = F_t^\ast \mathcal{L}_{X_t} \alpha
\end{equation}

\begin{equation}
= F_t^\ast \left[ d(\alpha \lvert X_t) + (d\alpha \lvert X_t) \right].
\end{equation}
In particular, if \( \alpha \) is closed, then, if \( F_t \) are diffeomorphisms for \( 0 \leq t \leq 1 \),

\[
F_t^* \alpha - F_0^* \alpha = d\beta, \quad \beta = \int_0^1 F_t^* (\alpha|X_t) \, dt.
\]

The fact that the left side of (11.27) is exact is a special case of Proposition 11.3, but the explicit formula given in (11.27) can be useful.

**More on the divergence of a vector field**

Let \( M \subset \mathbb{R}^n \) be an \( m \)-dimensional, oriented surface, with volume form \( \omega \). Then \( d\omega = 0 \) on \( M \), so, if \( X \) is a vector field on \( M \),

\[
L_X \omega = d(\omega|X).
\]

Comparison with (9.7) gives

\[
(\text{div} \, X) \omega = L_X \omega.
\]

This is sometimes taken as the definition of \( \text{div} \, X \). It readily leads to a formula for how the flow \( \mathcal{F}_X^t \) affects volumes.

To get this, we start with

\[
\frac{d}{dt} (\mathcal{F}_X^t)^* \omega = (\mathcal{F}_X^t)^* L_X \omega = (\mathcal{F}_X^t)^* ((\text{div} \, X) \omega).
\]

Hence, if \( \Omega \subset M \) is a smoothly bounded domain on which the flow \( \mathcal{F}_X^t \) is defined for \( t \in I \), then, for such \( t \),

\[
\frac{d}{dt} \text{Vol} \mathcal{F}_X^t (\Omega) = \frac{d}{dt} \int_{\Omega} (\mathcal{F}_X^t)^* \omega
\]

\[
= \int_{\Omega} (\mathcal{F}_X^t)^* ((\text{div} \, X) \omega)
\]

\[
= \int_{\mathcal{F}_X^t (\Omega)} (\text{div} \, X) \omega.
\]

In other words,

\[
\frac{d}{dt} \text{Vol} \mathcal{F}_X^t (\Omega) = \int_{\mathcal{F}_X^t (\Omega)} (\text{div} \, X) \, dV.
\]
This result is equivalent to Proposition 5.7, but the derivation here is substantially different. Compare also the discussion in Exercise 2 of §9.

**Exercises**

1. Show that if $\alpha$ is a $k$-form and $X, X_j$ are vector fields,

$$ (\mathcal{L}_X \alpha)(x_1, \ldots, x_k) = X \cdot \alpha(x_1, \ldots, x_k) - \sum_j \alpha(x_1, \ldots, \mathcal{L}_X X_j, \ldots, x_k). $$

Recall from (3.80) that $\mathcal{L}_X X_j = [X, X_j]$, and rewrite (11.33) accordingly.

2. Writing (11.20) as

$$ \iota_X d\alpha = \mathcal{L}_X \alpha - dt \iota_X \alpha, $$
deduce that

$$ (d\alpha)(X_0, X_1, \ldots, X_k) = (\mathcal{L}_{X_0} \alpha)(X_1, \ldots, x_k) - (\iota_{X_0} \alpha)(X_1, \ldots, x_k). $$

3. In case $\alpha$ is a one-form, deduce from (11.33)–(11.34) that

$$ (d\alpha)(X_0, X_1) = X_0 \cdot \alpha(X_1) - X_1 \cdot \alpha(X_0) - \alpha([X_0, X_1]). $$

4. Using (11.33)–(11.34) and induction on $k$, show that, if $\alpha$ is a $k$-form,

$$ (d\alpha)(X_0, \ldots, X_k) = \sum_{\ell=0}^k (-1)^\ell X_0 \cdot \alpha(X_0, \ldots, \hat{X}_\ell, \ldots, X_k) $$

$$ + \sum_{0 \leq \ell < j \leq k} (-1)^{j+\ell} \alpha([X_\ell, X_j], X_0, \ldots, \hat{X}_\ell, \ldots, \hat{X}_j, \ldots, X_k). $$

Here, $\hat{X}_\ell$ indicates that $X_\ell$ has been omitted.

5. Show that if $X$ is a vector field, $\beta$ a 1-form, and $\alpha$ a $k$-form, then

$$ (\bigwedge \beta \iota_X + \iota_X \bigwedge \beta) \alpha = \langle X, \beta \rangle \alpha. $$

Deduce that

$$ (df) \wedge (\alpha X) + (df \wedge \alpha) X = (Xf) \alpha. $$

6. Show that the definition (11.19) implies

$$ \mathcal{L}_X (f \alpha) = f \mathcal{L}_X \alpha + (Xf) \alpha. $$
7. Show that the definition (11.19) implies

\begin{equation}
(11.40) \quad d\mathcal{L}_X \alpha = \mathcal{L}_X (d\alpha).
\end{equation}

8. Denote the right side of (11.20) by $L_X \alpha$, i.e., set

\begin{equation}
(11.41) \quad L_X \alpha = d(\alpha | X) + (d\alpha | X).
\end{equation}

Show that this definition directly implies

\begin{equation}
(11.42) \quad L_X (d\alpha) = d(L_X \alpha).
\end{equation}

9. With $L_X$ defined by (11.41), show that

\begin{equation}
(11.43) \quad L_X (f\alpha) = fL_X \alpha + (Xf)\alpha.
\end{equation}

*Hint.* Use (11.38).

10. Use the results of Exercises 6–9 to give another proof of Proposition 11.6, i.e., $\mathcal{L}_X \alpha = L_X \alpha$.

*Hint.* Start with $\mathcal{L}_X f = Xf = \langle X, df \rangle = L_X f$.

In Exercises 11–12, let $X$ and $Y$ be smooth vector fields on $M$ and $\alpha \in \Lambda^k(M)$.

11. Show that $\mathcal{L}_{[X,Y]} \alpha = \mathcal{L}_X \mathcal{L}_Y \alpha - \mathcal{L}_Y \mathcal{L}_X \alpha$.

12. Using Exercise 11 and (11.29), show that

$$
div[X,Y] = X(div Y) - Y(div X).$$
12. Differential forms and degree theory

Degree theory assigns an integer, $\text{Deg}(f)$, to a smooth map $f : X \to Y$, when $X$ and $Y$ are smooth, compact, oriented surfaces of the same dimension, and $Y$ is connected. This has many uses, as we will see. Results of §11 provide tools for this study. A major ingredient is Stokes’ theorem.

As a prelude to our development of degree theory, we use the calculus of differential forms to provide simple proofs of some important topological results of Brouwer. The first two results concern retractions. If $Y$ is a subset of $X$, by definition a retraction of $X$ onto $Y$ is a map $\varphi : X \to Y$ such that $\varphi(x) = x$ for all $x \in Y$.

**Proposition 12.1.** There is no smooth retraction $\varphi : B \to S^{n-1}$ of the closed unit ball $B$ in $\mathbb{R}^n$ onto its boundary $S^{n-1}$.

In fact, it is just as easy to prove the following more general result. The approach we use is adapted from [Kan].

**Proposition 12.2.** If $M$ is a compact oriented $n$-dimensional surface with nonempty boundary $\partial M$, there is no smooth retraction $\varphi : M \to \partial M$.

**Proof.** Pick $\omega \in \Lambda^n(\partial M)$ to be the volume form on $\partial M$, so $\int_{\partial M} \omega > 0$. Now apply Stokes’ theorem to $\beta = \varphi^* \omega$. If $\varphi$ is a retraction, then $\varphi \circ j(x) = x$, where $j : \partial M \hookrightarrow M$ is the natural inclusion. Hence $j^* \varphi^* \omega = \omega$, so we have

$$\int_{\partial M} \omega = \int_M d(\varphi^* \omega).$$

But $d(\varphi^* \omega) = \varphi^* d\omega = 0$, so the integral (12.1) is zero. This is a contradiction, so there can be no retraction.

A simple consequence of this is the famous Brouwer Fixed-Point Theorem. We first present the smooth case.

**Theorem 12.3.** If $F : B \to B$ is a smooth map on the closed unit ball in $\mathbb{R}^n$, then $F$ has a fixed point.

**Proof.** We are claiming that $F(x) = x$ for some $x \in B$. If not, define $\varphi(x)$ to be the endpoint of the ray from $F(x)$ to $x$, continued until it hits $\partial B = S^{n-1}$. An explicit formula is

$$\varphi(x) = x + t(x - F(x)), \quad t = \frac{\sqrt{b^2 + 4ac} - b}{2a},$$

$$a = \|x - F(x)\|^2, \quad b = 2x \cdot (x - F(x)), \quad c = 1 - \|x\|^2.$$

Here $t$ is picked to solve the equation $\|x + t(x - F(x))\|^2 = 1$. Note that $ac \geq 0$, so $t \geq 0$. It is clear that $\varphi$ would be a smooth retraction, contradicting Proposition 11.1.

Now we give the general case, using the Stone-Weierstrass theorem (discussed in Appendix E) to reduce it to Theorem 12.3.
Theorem 12.4. If $G : B \to B$ is a continuous map on the closed unit ball in $\mathbb{R}^n$, then $G$ has a fixed point.

Proof. If not, then
\[
\inf_{x \in B} |G(x) - x| = \delta > 0.
\]
The Stone-Weierstrass theorem (Appendix E) implies there exists a polynomial $P$ such that $|P(x) - G(x)| < \delta/8$ for all $x \in B$. Set
\[
F(x) = \left(1 - \frac{\delta}{8}\right) P(x).
\]
Then $F : B \to B$ and $|F(x) - G(x)| < \delta/2$ for all $x \in B$, so
\[
\inf_{x \in B} |F(x) - x| > \frac{\delta}{2}.
\]
This contradicts Theorem 12.3.

As a second precursor to degree theory, we next show that an even dimensional sphere cannot have a smooth nonvanishing vector field.

Proposition 12.5. There is no smooth nonvanishing vector field on $S^n$ if $n = 2k$ is even.

Proof. If $X$ were such a vector field, we could arrange it to have unit length, so we would have $X : S^n \to S^n$ with $X(v) \perp v$ for $v \in S^n \subset \mathbb{R}^{n+1}$. Thus there would be a unique unit speed curve $\gamma_v$ along the great circle from $v$ to $X(v)$, of length $\pi/2$. Define a smooth family of maps $F_t : S^n \to S^n$ by $F_t(v) = \gamma_v(t)$. Thus $F_0(v) = v$, $F_{\pi/2}(v) = X(v)$, and $F_{\pi} = A$ would be the antipodal map, $A(v) = -v$. By Proposition 11.3, we deduce that $A^*\omega - \omega = d\beta$ is exact, where $\omega$ is the volume form on $S^n$. Hence, by Stokes’ theorem,
\[
\int_{S^n} A^*\omega = \int_{S^n}\omega.
\]
Alternatively, (12.2) follows directly from Proposition 11.1. On the other hand, it is straightforward that $A^*\omega = (-1)^{n+1}\omega$, so (12.2) is possible only when $n$ is odd.

Note that an important ingredient in the proof of both Proposition 12.2 and Proposition 12.5 is the existence of $n$-forms on a compact oriented $n$-dimensional surface $M$ that are not exact (though of course they are closed). We next establish the following counterpoint to the Poincaré lemma.

Proposition 12.6. If $M$ is a compact, connected, oriented surface of dimension $n$ and $\alpha \in \Lambda^n M$, then $\alpha = d\beta$ for some $\beta \in \Lambda^{n-1}(M)$ if and only if
\[
\int_M \alpha = 0.
\]
We have already discussed the necessity of (12.3). To prove the sufficiency, we first look at the case $M = S^n$.

In that case, any $n$-form $\alpha$ is of the form $a(x) \omega$, $a \in C^\infty(S^n)$, $\omega$ the volume form on $S^n$, with its standard metric. The group $G = SO(n + 1)$ of rotations of $\mathbb{R}^{n+1}$ acts as a transitive group of isometries on $S^n$. In §5 we constructed the integral of functions over $SO(n + 1)$, with respect to Haar measure.

As seen in §5, we have the surjective map

$$\text{Exp} : \text{Skew}(n + 1) \longrightarrow SO(n + 1),$$

giving a diffeomorphism from a ball $\mathcal{O}$ about $0$ in $\text{Skew}(n + 1)$ onto an open set $U \subset SO(n + 1) = G$, a neighborhood of the identity. Since $G$ is compact, we can pick a finite number of elements $\xi_j \in G$ such that the open sets $U_j = \{\xi_j g : g \in U\}$ cover $G$. Pick $\eta_j \in \text{Skew}(n + 1)$ such that $\text{Exp} \eta_j = \xi_j$. Define $\Phi_{jt} : U_j \rightarrow G$ for $0 \leq t \leq 1$ by

$$\Phi_{jt}(\xi_j \text{Exp}(A)) = (\text{Exp} \ t\eta_j)(\text{Exp} \ tA), \quad A \in \mathcal{O}. \quad (12.4)$$

Now partition $G$ into subsets $\Omega_j$, each of whose boundaries has content zero, such that $\Omega_j \subset U_j$. If $g \in \Omega_j$, set $g(t) = \Phi_{jt}(g)$. This family of elements of $SO(n + 1)$ defines a family of maps $F_{gt} : S^n \rightarrow S^n$. Now by (11.27) we have

$$\alpha = g^*\alpha - d\kappa_g(\alpha), \quad \kappa_g(\alpha) = \int_0^1 F_{gt}^*(\alpha) X_{gt} \, dt, \quad (12.5)$$

for each $g \in SO(n + 1)$, where $X_{gt}$ is the family of vector fields on $S^n$ associated to $F_{gt}$, as in (11.25). Therefore,

$$\alpha = \int_G g^*\alpha \, dg - d\int_G \kappa_g(\alpha) \, dg. \quad (12.6)$$

Now the first term on the right is equal to $\overline{\alpha} \omega$, where $\overline{\alpha} = \int a(g \cdot x) \, dg$ is a constant; in fact, the constant is

$$\overline{\alpha} = \frac{1}{\text{Vol} \ S^n} \int_{S^n} \alpha. \quad (12.7)$$

Thus in this case (12.3) is precisely what serves to make (12.6) a representation of $\alpha$ as an exact form. This takes care of the case $M = S^n$.

For a general compact, oriented, connected $M$, proceed as follows. Cover $M$ with open sets $\mathcal{O}_1, \ldots, \mathcal{O}_K$ such that each $\overline{\mathcal{O}}_j$ is diffeomorphic to the closed unit ball in $\mathbb{R}^n$. Set $U_1 = \mathcal{O}_1$, and inductively enlarge each $\mathcal{O}_j$ to $U_j$, so that $\overline{U}_j$ is also diffeomorphic to the closed ball, and such that $U_{j+1} \cap U_j \neq \emptyset$, $1 \leq j < K$. You can do this by drawing a simple curve from $\overline{\mathcal{O}}_{j+1}$ to a point in $U_j$ and thickening it. Pick a smooth partition of unity $\varphi_j$, subordinate to this cover. (See Appendix B.)
Given \( \alpha \in \Lambda^n M \), satisfying (12.3), take \( \tilde{\alpha}_j = \varphi_j \alpha \). Most likely \( \int \tilde{\alpha}_1 = c_1 \neq 0 \), so take \( \sigma_1 \in \Lambda^n M \), with compact support in \( U_1 \cap U_2 \), such that \( \int \sigma_1 = c_1 \). Set \( \alpha_1 = \tilde{\alpha}_1 - \sigma_1 \), and redefine \( \tilde{\alpha}_2 \) to be the old \( \tilde{\alpha}_2 \) plus \( \sigma_1 \). Make a similar construction using \( \int \tilde{\alpha}_2 = c_2 \), and continue. When you are done, you have

\[
\alpha = \alpha_1 + \cdots + \alpha_K,
\]

with \( \alpha_j \) compactly supported in \( U_j \). By construction,

\[
\int \alpha_j = 0
\]

for \( 1 \leq j < K \). But then (12.3) implies \( \int \alpha_K = 0 \) too.

Now pick \( p \in S^n \) and define smooth maps

\[
\psi_j : M \to S^n
\]

which map \( U_j \) diffeomorphically onto \( S^n \setminus p \), and map \( M \setminus U_j \) to \( p \). There is a unique \( v_j \in \Lambda^n S^n \), with compact support in \( S^n \setminus p \), such that \( \psi^* v_j = \alpha_j \). Clearly

\[
\int_{S^n} v_j = 0,
\]

so by the case \( M = S^n \) of Proposition 12.6 already established, we know that \( v_j = d w_j \) for some \( w_j \in \Lambda^{n-1} S^n \), and then

\[
\alpha_j = d \beta_j, \quad \beta_j = \psi_j^* w_j.
\]

This concludes the proof of Proposition 12.6.

We are now ready to introduce the notion of the degree of a map between compact oriented surfaces. Let \( X \) and \( Y \) be compact oriented \( n \)-dimensional surfaces. We want to define the degree of a smooth map \( F : X \to Y \). To do this, assume \( Y \) is connected. We pick \( \omega \in \Lambda^n Y \) such that

\[
\int_Y \omega = 1.
\]

We propose to define

\[
\text{Deg}(F) = \int_X F^* \omega.
\]

The following result shows that \( \text{Deg}(F) \) is indeed well defined by this formula. The key argument is an application of Proposition 12.6.
Lemma 12.7. The quantity (12.13) is independent of the choice of $\omega$, as long as (12.12) holds.

Proof. Pick $\omega_1 \in \Lambda^n Y$ satisfying $\int_Y \omega_1 = 1$, so $\int_Y (\omega - \omega_1) = 0$. By Proposition 12.6, this implies

$$\omega - \omega_1 = d\alpha, \text{ for some } \alpha \in \Lambda^{n-1} Y.$$  

(12.14)

Thus

$$\int_X F^*\omega - \int_X F^*\omega_1 = \int_X dF^*\alpha = 0,$$

(12.15)

and the lemma is proved.

The following is a most basic property.

Proposition 12.8. If $F_0$ and $F_1$ are smoothly homotopic, then $\text{Deg}(F_0) = \text{Deg}(F_1)$.

Proof. By Proposition 11.1, if $F_0$ and $F_1$ are smoothly homotopic, then $\int_X F_0^*\omega = \int_X F_1^*\omega$.

The following result is a simple but powerful extension of Proposition 12.8. Compare the relation between Propositions 11.1 and 11.2.

Proposition 12.9. Let $\overline{M}$ be a compact oriented surface with boundary, $\dim M = n + 1$. Take $Y$ as above, $n = \dim Y$. Given a smooth map $F : \overline{M} \to Y$, let $\alpha = F_{|\partial M} : \partial M \to Y$. Then

$$\text{Deg}(f) = 0.$$

Proof. Applying Stokes’ Theorem to $\alpha = F^*\omega$, we have

$$\int_{\partial M} f^*\omega = \int_{\overline{M}} dF^*\omega.$$  

But $dF^*\omega = F^*d\omega$, and $d\omega = 0$ if $\dim Y = n$, so we are done.

Brouwer’s no-retraction theorem is an easy corollary of Proposition 12.9. Compare the proof of Proposition 12.2.

Corollary 12.10. If $\overline{M}$ is a compact oriented surface with nonempty boundary $\partial M$, then there is no smooth retraction $\varphi : \overline{M} \to \partial M$.

Proof. Without loss of generality, we can assume $\overline{M}$ is connected. If there were a retraction, then $\partial M = \varphi(\overline{M})$ must also be connected, so Proposition 12.9 applies. But then we would have, for the map $\text{id.} = \varphi_{|\partial M}$, the contradiction that its degree is both zero and 1.

We next give an alternative formula for the degree of a map, which is very useful in many applications. In particular, it implies that the degree is always an integer.

A point $y_0 \in Y$ is called a regular value of $F$, provided that, for each $x \in X$ satisfying $F(x) = y_0$, $DF(x) : T_x X \to T_{y_0} Y$ is an isomorphism. The easy case of Sard’s Theorem, discussed in Appendix F, implies that most points in $Y$ are regular. Endow $X$ with a volume element $\omega_X$, and similarly endow $Y$ with $\omega_Y$. If $DF(x)$ is invertible, define $JF(x) \in \mathbb{R} \setminus 0$ by $F^*(\omega_Y) = JF(x)\omega_X$. Clearly the sign of $JF(x)$, i.e., $\text{sgn } JF(x) = \pm 1$, is independent of choices of $\omega_X$ and $\omega_Y$, as long as they determine the given orientations of $X$ and $Y$. 


Proposition 12.11. If $y_0$ is a regular value of $F$, then

$$
\text{Deg}(F) = \sum \{\text{sgn} JF(x_j) : F(x_j) = y_0\}.
$$

Proof. Pick $\omega \in \Lambda^n Y$, satisfying (12.12), with support in a small neighborhood of $y_0$. Then $F^*\omega$ will be a sum $\sum \omega_j$, with $\omega_j$ supported in a small neighborhood of $x_j$, and $\int \omega_j = \pm 1$ as $\text{sgn} JF(x_j) = \pm 1$.

For an application of Proposition 12.11, let $X$ be a compact smooth oriented hypersurface in $\mathbb{R}^{n+1}$, and set $\Omega = \mathbb{R}^{n+1} \setminus X$. Given $p \in \Omega$, define

$$
F_p : X \to S^n, \quad F_p(x) = \frac{x - p}{|x - p|}.
$$

It is clear that $\text{Deg}(F_p)$ is constant on each connected component of $\Omega$. It is also easy to see that, when $p$ crosses $X$, $\text{Deg}(F_p)$ jumps by $\pm 1$. Thus $\Omega$ has at least two connected components. This is most of the smooth case of the Jordan-Brouwer separation theorem:

**Theorem 12.12.** If $X$ is a smooth compact oriented hypersurface of $\mathbb{R}^{n+1}$, which is connected, then $\Omega = \mathbb{R}^{n+1} \setminus X$ has exactly 2 connected components.

Proof. $X$ being oriented, it has a smooth global normal vector field. Use this to separate a small collar neighborhood $C$ of $X$ into 2 pieces; $C \setminus X = C_0 \cup C_1$. The collar $C$ is diffeomorphic to $[-1, 1] \times X$, and each $C_j$ is clearly connected. It suffices to show that any connected component $O$ of $\Omega$ intersects either $C_0$ or $C_1$. Take $p \in \partial O$. If $p \notin X$, then $p \in \Omega$, which is open, so $p$ cannot be a boundary point of any component of $\Omega$. Thus $\partial O \subset X$, so $O$ must intersect a $C_j$. This completes the proof.

Let us note that, of the two components of $\Omega$, exactly one is unbounded, say $\Omega_0$, and the other is bounded; call it $\Omega_1$. Then we claim

$$
p \in \Omega_j \implies \text{Deg}(F_p) = j.
$$

Indeed, for $p$ very far from $X$, $F_p : X \to S^n$ is not onto, so its degree is 0. And when $p$ crosses $X$, from $\Omega_0$ to $\Omega_1$, the degree jumps by $+1$.

For a simple closed curve in $\mathbb{R}^2$, Theorem 12.12 is the smooth case of the Jordan curve theorem. That special case of the argument given above can be found in [Sto]. The existence of a smooth normal field simplifies the use of basic degree theory to prove such a result. For a general continuous, simple closed curve in $\mathbb{R}^2$, such a normal field is not available, and the proof of the Jordan curve theorem in this more general context requires a different argument, which can be found in [GrH].

We apply results just established on degree theory to properties of vector fields, particularly of their critical points. A critical point of a vector field $V$ is a point where $V$
vanishes. Let \( V \) be a vector field defined on a neighborhood \( \mathcal{O} \) of \( p \in \mathbb{R}^n \), with a single critical point, at \( p \). Then, for any small ball \( B_r \) about \( p \), \( B_r \subset \mathcal{O} \), we have a map
\[
(12.19) \quad V_r : \partial B_r \to S^{n-1}, \quad V_r(x) = \frac{V(x)}{|V(x)|}
\]
The degree of this map is called the index of \( V \) at \( p \), denoted \( \text{ind}_p(V) \); it is clearly independent of \( r \). If \( V \) has a finite number of critical points, then the index of \( V \) is defined to be
\[
(12.20) \quad \text{Index}(V) = \sum \text{ind}_p(V).
\]
If \( \psi : \mathcal{O} \to \mathcal{O}' \) is an orientation preserving diffeomorphism, taking \( p \) to \( p' \) and \( V \) to \( W \), then we claim
\[
(12.21) \quad \text{ind}_p(V) = \text{ind}_{p'}(W).
\]
In fact, \( D\psi(p) \) is an element of \( GL(n, \mathbb{R}) \) with positive determinant, so it is homotopic to the identity, and from this it readily follows that \( V_r \) and \( W_r \) are homotopic maps of \( \partial B_r \to S^{n-1} \). Thus one has a well defined notion of the index of a vector field with a finite number of critical points on any oriented surface \( M \).

There is one more wrinkle. Suppose \( X \) is a smooth vector field on \( M \) and \( p \) an isolated critical point. If you change the orientation of a small coordinate neighborhood \( \mathcal{O} \) of \( p \), then the orientations of both \( \partial B_r \) and \( S^{n-1} \) in (12.19) get changed, so the associated degree is not changed. Hence one has a well defined notion of the index of a vector field with a finite number of critical points on any smooth surface \( M \), oriented or not.

A vector field \( V \) on \( \mathcal{O} \subset \mathbb{R}^n \) is said to have a non-degenerate critical point at \( p \) provided \( DV(p) \) is a nonsingular \( n \times n \) matrix. The following formula is convenient.

**Proposition 12.13.** If \( V \) has a non-degenerate critical point at \( p \), then
\[
(12.22) \quad \text{ind}_p(V) = \text{sgn det} DV(p).
\]

**Proof.** If \( p \) is a nondegenerate critical point, and we set \( \psi(x) = DV(p)x, \ \psi_r(x) = \psi(x)/|\psi(x)| \), for \( x \in \partial B_r \), it is readily verified that \( \psi_r \) and \( V_r \) are homotopic, for \( r \) small. The fact that \( \text{Deg}(\psi_r) \) is given by the right side of (12.22) is an easy consequence of Proposition 12.11.

The following is an important global relation between index and degree.

**Proposition 12.14.** Let \( \Omega \) be a smooth bounded region in \( \mathbb{R}^{n+1} \). Let \( V \) be a vector field on \( \Omega \), with a finite number of critical points \( p_j \), all in the interior \( \Omega \). Define \( F : \partial \Omega \to S^n \) by \( F(x) = V(x)/|V(x)| \). Then
\[
(12.23) \quad \text{Index}(V) = \text{Deg}(F).
\]

**Proof.** If we apply Proposition 12.9 to \( \overline{M} = \Omega \setminus \bigcup_j B_\epsilon(p_j) \), we see that \( \text{Deg}(F) \) is equal to the sum of degrees of the maps of \( \partial B_\epsilon(p_j) \) to \( S^n \), which gives (12.23).

Next we look at a process of producing vector fields in higher dimensional spaces from vector fields in lower dimensional spaces.
Proposition 12.15. Let $W$ be a vector field on $\mathbb{R}^n$, vanishing only at 0. Define a vector field $V$ on $\mathbb{R}^{n+k}$ by $V(x, y) = (W(x), y)$. Then $V$ vanishes only at $(0, 0)$. Then we have

\[(12.24) \quad \text{ind}_0 W = \text{ind}_{(0, 0)} V.\]

**Proof.** If we use Proposition 12.11 to compute degrees of maps, and choose $y_0 \in S^{n-1} \subset S^{n+k-1}$, a regular value of $W$, and hence also for $V$, this identity follows.

We turn to a more sophisticated variation. Let $X$ be a compact $n$-dimensional surface in $\mathbb{R}^{n+k}$, $W$ a (tangent) vector field on $X$ with a finite number of critical points $p_j$. Let $\overline{\Omega}$ be a small tubular neighborhood of $X$, $\pi : \overline{\Omega} \to X$ mapping $z \in \overline{\Omega}$ to the nearest point in $X$. Let $\varphi(z) = \text{dist}(z, X)^2$. Now define a vector field $V$ on $\overline{\Omega}$ by

\[(12.25) \quad V(z) = W(\pi(z)) + \nabla \varphi(z).\]

**Proposition 12.16.** If $F : \partial \Omega \to S^{n+k-1}$ is given by $F(z) = V(z)/|V(z)|$, then

\[(12.26) \quad \text{Deg}(F) = \text{Index}(W).\]

**Proof.** We see that all the critical points of $V$ are points in $X$ that are critical for $W$, and, as in Proposition 12.15, $\text{Index}(W) = \text{Index}(V)$. Then Proposition 12.14 implies $\text{Index}(V) = \text{Deg}(F)$.

Since $\varphi(z)$ is increasing as one goes away from $X$, it is clear that, for $z \in \partial \Omega$, $V(z)$ points out of $\overline{\Omega}$, provided it is a sufficiently small tubular neighborhood of $X$. Thus $F : \partial \Omega \to S^{n+k-1}$ is homotopic to the *Gauss map*

\[(12.27) \quad N : \partial \Omega \to S^{n+k-1},\]

given by the outward pointing normal. This immediately gives:

**Corollary 12.17.** Let $X$ be a compact $n$-dimensional surface in $\mathbb{R}^{n+k}$, $\overline{\Omega}$ a small tubular neighborhood of $X$, and $N : \partial \Omega \to S^{n+k-1}$ the Gauss map. If $W$ is a vector field on $X$ with a finite number of critical points, then

\[(12.28) \quad \text{Index}(W) = \text{Deg}(N).\]

Clearly the right side of (12.28) is independent of the choice of $W$. Thus any two vector fields on $X$ with a finite number of critical points have the same index, i.e., $\text{Index}(W)$ is an invariant of $X$. This invariant is denoted

\[(12.29) \quad \text{Index}(W) = \chi(X),\]

and is called the Euler characteristic of $X$. 


Remark. The existence of smooth vector fields with only nondegenerate critical points (hence only finitely many critical points) on a given compact surface $X$ follows from results presented in Appendix G.

Exercises

1. Let $X$ be a compact, oriented, connected surface. Show that the identity map $I : X \to X$ has degree 1.

2. Suppose $Y$ is also a compact, oriented, connected surface. Show that if $F : X \to Y$ is not onto, then $\text{Deg}(F) = 0$.

3. If $A : S^n \to S^n$ is the antipodal map, show that $\text{Deg}(A) = (-1)^{n-1}$.

4. Show that the homotopy invariance property given in Proposition 12.8 can be deduced as a corollary of Proposition 12.9.
   
   Hint. Take $\tilde{M} = X \times [0,1]$.

5. Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ be a polynomial of degree $n \geq 1$. The fundamental theorem of algebra, proved in §10, states that $p(z_0) = 0$ for some $z_0 \in \mathbb{C}$. We aim for another proof, using degree theory. To get this, by contradiction, assume $p : \mathbb{C} \to \mathbb{C} \setminus 0$. For $r \geq 0$, define

   $$F_r : S^1 \to S^1, \quad F_r(e^{i\theta}) = \frac{p(re^{i\theta})}{|p(re^{i\theta})|}.$$ 

   Show that each $F_r$ is smoothly homotopic to $F_0$, and note that $\text{Deg}(F_0) = 0$. Then show that there exists $r_0$ such that

   $$r \geq r_0 \Rightarrow F_r \text{ is homotopic to } \Phi,$$

   where $\Phi(e^{i\theta}) = e^{in\theta}$. Show that $\text{Deg}(\Phi) = n$, and obtain a contradiction.

   Note. Regarding the use of degree theory here, show how the proof of Proposition 12.6 vastly simplifies when $M = S^1$.

6. Show that each odd-dimensional sphere $S^{2k-1}$ has a smooth, nowhere vanishing tangent vector field.
   
   Hint. Regard $S^{2k-1} \subset \mathbb{C}^k$, and multiply the unit normal by $i$.

7. Let $V$ be a planar vector field. Assume it has a nondegenerate critical point at $p$. Show
that

\[ p \text{ saddle } \implies \text{ind}_p(V) = -1 \]
\[ p \text{ source } \implies \text{ind}_p(V) = 1 \]
\[ p \text{ sink } \implies \text{ind}_p(V) = 1 \]
\[ p \text{ center } \implies \text{ind}_p(V) = 1 \]

8. Let \( M \) be a compact oriented 2-dimensional surface. Given a triangulation of \( M \), within each triangle construct a vector field, vanishing at 7 points as illustrated in Fig. 12.1, with the vertices as attractors, the center as a repeller, and the midpoints of each edge as saddle points. Fit these together to produce a smooth vector field \( X \) on \( M \). Show directly that

\[ \text{Index}(X) = V - E + F, \]

where

\[ V = \# \text{ vertices}, \quad E = \# \text{ edges}, \quad F = \# \text{ faces}, \]

in the triangulation.

9. With \( X = S^n \subset \mathbb{R}^{n+1} \), note that the manifold \( \partial \Omega \) in (12.27) consists of two copies of \( S^n \), with opposite orientations. Compute the degree of the map \( N \) in (12.27)–(12.28), and use this to show that

\[ \chi(S^n) = 2 \text{ if } n \text{ even, } \quad 0 \text{ if } n \text{ odd,} \]

granted (12.28)–(12.29).

10. Consider the vector field \( R \) on \( S^2 \) generating rotation about an axis. Show that \( R \) has two critical points, at the “poles.” Classify the critical points, compute \( \text{Index}(R) \), and compare the \( n = 2 \) case of (12.30).

11. Generalizing Exercise 9, Let \( X \subset \mathbb{R}^{n+1} \) be a smooth, compact, oriented, \( n \)-dimensional surface, so again the neighborhood \( \Omega \) of \( X \) as in (12.27) has boundary \( \partial \Omega \) consisting essentially of two copies of \( X \), with opposite orientations. Let \( \vec{N} : X \rightarrow S^n \) be the outward pointing unit normal. Show that

\[ \text{Deg} \vec{N} = \frac{1}{2} \chi(X), \quad \text{if } n \text{ is even.} \]

Remark. If \( \omega_S \) is the volume form on \( S^n \) and \( \omega_X \) that on \( X \), then \( \vec{N}^* \omega_S = K \omega_X \), and \( K : X \rightarrow \mathbb{R} \) is called the Gauss curvature of \( X \). Then (12.31) implies

\[ \int_X K(x) dS(x) = \frac{1}{2} A_n \chi(X), \]

if \( n \) is even. This is a basic case of the Gauss-Bonnet formula.
12. In the setting of Exercise 11, assume $n$ is odd. Show that
\begin{equation}
\chi(X) = 0.
\end{equation}
Give examples where the identity in (12.31) fails.

13. Retain the setting of Exercise 12, especially that $n$ is odd. Let $X = \partial \mathcal{O}$, with $\mathcal{O} \subset \mathbb{R}^{n+1}$ bounded and open. Take a smooth function
\[ \varphi : \overline{\mathcal{O}} \to [0, \infty), \quad \varphi(x) = \text{dist}(x, X) \text{ near } X, \quad \varphi(x) > 0 \text{ on } \mathcal{O}. \]
Let $\Sigma \subset \mathbb{R}^{n+2}$ be the surface
\[ \Sigma = \{(x, y) : x \in \overline{\mathcal{O}}, \ y^2 = \varphi(x)\}, \]
and let $\nu : \Sigma \to S^{n+1}$ be the outward pointing unit normal. Show that
\begin{equation}
\text{Deg } \nu = \text{Deg } \tilde{N},
\end{equation}
and deduce that
\begin{equation}
\text{Deg } \tilde{N} = \frac{1}{2} \chi(\Sigma).
\end{equation}
\textit{Hint.} Taking $\tilde{N} : X \to S^n \subset S^{n+1}$, show that each regular value of $\tilde{N}$ is also a regular value of $\nu$, with the same preimage in $X \subset \Sigma$. Then show that Proposition 12.11 applies.

14. Actually, (12.33) holds whether $n$ is odd or even. Can you get anything else from this?

15. In the setting of Exercise 12 ($n$ is odd), generalize the construction of Exercise 6 to show directly that there is a smooth, nowhere vanishing vector field tangent to $X$.

16. Let $\mathcal{O} \subset \mathbb{R}^n$ be open and $f : \mathcal{O} \to \mathbb{R}$ smooth of class $C^2$. Let $V = \nabla f$. Assume $p \in \mathcal{O}$ is a nondegenerate critical point of $f$, so its Hessian $D^2 f(p)$ is a nondegenerate $n \times n$ symmetric matrix. Say
\begin{equation}
D^2 f(p) \text{ has } \ell \text{ positive eigenvalues and } n - \ell \text{ negative eigenvalues.}
\end{equation}
Show that Proposition 12.13 implies
\begin{equation}
\text{ind}_p(V) = (-1)^{n-\ell}.
\end{equation}

17. Let $X \subset \mathbb{R}^{n+k}$ be a smooth, compact, $n$-dimensional surface. Assume there exists
f ∈ C^2(X), with just two critical points, a max and a min, both nondegenerate. Use Exercise 16 to show that

\[ \chi(X) = 2 \text{ if } n \text{ is even, } 0 \text{ if } n \text{ is odd.} \]

Considering \( S^n \subset \mathbb{R}^{n+1} \), use this to give another demonstration of (12.30).

18. Let \( \mathcal{T} \subset \mathbb{R}^3 \) be the “inner tube” surface described in Exercise 15 of §5.

(a) Show that rotation about the z-axis is generated by a vector field that is tangent to \( \mathcal{T} \) and nowhere vanishing of \( \mathcal{T} \).

(b) Define \( f : \mathcal{T} \rightarrow \mathbb{R} \) by \( f(x, y, z) = x \), \( (x, y, z) \in \mathcal{T} \). Show that \( f \) has four critical points, a max, a min, and two saddles. Deduce from Exercise 16 that \( \nabla f \) is a vector field on \( \mathcal{T} \) of index 0.

(c) Show that both part (a) and part (b) imply \( \chi(\mathcal{T}) = 0 \).

Let \( X \subset \mathbb{R}^n \) be an \( m \)-dimensional surface, and let \( Y \subset \mathbb{R}^\nu \) be a \( \mu \)-dimensional surface, both smooth of class \( C^k \). Then

\[ X \times Y = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^\nu : x \in X, y \in Y\} \subset \mathbb{R}^n \times \mathbb{R}^\nu \]

has a natural structure of an \((m + \mu)\)-dimensional \( C^k \) surface.

19. Let \( X \) and \( Y \) be as above, and assume they are both compact. Let \( V_1 \) be a smooth vector field tangent to \( X \) and \( V_2 \) a smooth vector field tangent to \( Y \), both with only nondegenerate critical points. Say \( \{p_i\} \) are the critical points of \( V_1 \) and \( \{q_j\} \) those of \( V_2 \).

(a) Show that \( W(x, y) = V_1(x) + V_2(y) \) is a smooth vector field tangent to \( X \times Y \). Show that its critical points are precisely the points \( \{(p_i, q_j)\} \), each nondegenerate. Show that Proposition 12.13 gives

\[ \text{ind}_{(p_i, q_j)} W = (\text{ind}_{p_i} V_1)(\text{ind}_{q_j} V_2). \]

(b) Show that

\[ \text{Index } W = (\text{Index } V_1)(\text{Index } V_2). \]

(c) Deduce that

\[ \chi(X \times Y) = \chi(X)\chi(Y). \]

Let \( X \) and \( Y \) be smooth, compact, oriented surfaces in \( \mathbb{R}^n \). Assume \( k = \dim X, \ell = \dim Y \), and \( k + \ell = n - 1 \). Assume \( X \cap Y = \emptyset \). Set

\[ \varphi : X \times Y \rightarrow S^{n-1}, \quad \varphi(x, y) = \frac{x - y}{|x - y|}. \]
We define the linking number

\[ \lambda(X, Y, \mathbb{R}^n) = \text{Deg } \varphi. \]

20. Let \( \gamma \) and \( \sigma \subset \mathbb{R}^3 \) be the following simple closed curves, parametrized by \( s, t \in \mathbb{R}/(2\pi\mathbb{Z}) \):

\[ \gamma(s) = (\cos s, \sin s, 0), \quad \sigma(t) = (0, 1 + \cos t, \sin t). \]

Thus \( \gamma \) is a circle in the \((x, y)\)-plane centered at \((0, 0, 0)\) and \( \sigma \) is a circle in the \((y, z)\)-plane, centered at \((0, 1, 0)\), both of unit radius. Show that

\[ \lambda(\gamma, \sigma, \mathbb{R}^3) = 1. \]

**Hint.** With \( \varphi \) as above, show that \((0, 1, 0) \in S^2\) has exactly one preimage point, under \( \varphi : \gamma \times \sigma \to S^2 \).

21. Let \( M \) be a smooth, compact, oriented, \((n - 1)\)-dimensional surface, and assume \( \varphi : M \to \mathbb{R}^n \setminus 0 \) is a smooth map. Set 

\[ F(x) = \frac{\varphi(x)}{|\varphi(x)|}, \quad F : M \to S^{n-1}. \]

Take \( \omega \in A^{n-1}(\mathbb{R}^n \setminus 0) \) to be the form considered in Exercises 9–10 of §8, i.e.,

\[ \omega = |x|^{-n} \sum_{j=1}^{n} x_j dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n. \]

Show that

\[ \text{Deg}(F) = \frac{1}{A_{n-1}} \int_M \varphi^* \omega. \]

**Hint.** Use Proposition 11.1 (with \( Y = \mathbb{R}^n \setminus 0 \)), plus Exercise 9 of §8, to show that

\[ \int_M \varphi^* \omega = \int_M F^* \omega, \]

and show that, under \( S^{n-1} \hookrightarrow \mathbb{R}^n \), \( j^* \omega \) is the area form on \( S^{n-1} \).

22. In Exercise 21, take \( n = 3 \) and \( M = T^2 = \mathbb{R}^2/(2\pi\mathbb{Z}^2) \), parametrized by \((s, t) \in \mathbb{R}^2\). Show that

\[ \varphi^* \omega = |\varphi|^{-3} \det \begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 \\ \partial_s \varphi_1 & \partial_s \varphi_2 & \partial_s \varphi_3 \\ \partial_t \varphi_1 & \partial_t \varphi_2 & \partial_t \varphi_3 \end{pmatrix} (ds \wedge dt) \]

\[ = |\varphi|^{-3} \varphi \cdot \left( \frac{\partial \varphi}{\partial s} \times \frac{\partial \varphi}{\partial t} \right) (ds \wedge dt). \]

In case \( \varphi(s, t) = \gamma(s) - \sigma(t) \), deduce the Gauss linking number formula:

\[ \lambda(\gamma, \sigma, \mathbb{R}^3) = -\frac{1}{4\pi} \int_{\mathbb{R}^2} \frac{\gamma(s) - \sigma(t)}{|\gamma(s) - \sigma(t)|^3} \cdot (\gamma'(s) \times \sigma'(t)) \, ds \, dt. \]
13. Fourier series

We consider Fourier series of functions on the $n$-dimensional torus $\mathbb{T}^n = \mathbb{R}^n/(2\pi \mathbb{Z}^n)$. Given $f \in C(\mathbb{T}^n)$ (or more generally $f \in R^\#(\mathbb{T}^n)$) we set, for $k \in \mathbb{Z}^n$,

\begin{equation}
\hat{f}(k) = (2\pi)^{-n} \int_{\mathbb{T}^n} f(\theta)e^{-ik \cdot \theta} \, d\theta,
\end{equation}

i.e.,

\begin{equation}
\hat{f}(k) = (2\pi)^{-n} \int_0^{2\pi} \cdots \int_0^{2\pi} f(\theta_1, \ldots, \theta_n)e^{-i(k_1\theta_1 + \cdots + k_n\theta_n)} \, d\theta_1 \cdots d\theta_n.
\end{equation}

We call $\hat{f}(k)$ the Fourier coefficients of $f$. The first major problem is to recover $f$ from its Fourier coefficients. We first accomplish this when $f$ belongs to the space

\begin{equation}
A(\mathbb{T}^n) = \left\{ f \in C(\mathbb{T}^n) : \sum_{k \in \mathbb{Z}^n} |\hat{f}(k)| < \infty \right\}.
\end{equation}

**Proposition 13.1.** If $f \in A(\mathbb{T}^n)$, then the following Fourier inversion formula holds:

\begin{equation}
f(\theta) = \sum_{k \in \mathbb{Z}^n} \hat{f}(k)e^{ik \cdot \theta}.
\end{equation}

**Proof.** Given $\sum |\hat{f}(k)| < \infty$, the right side of (13.4) is absolutely and uniformly convergent, defining

\begin{equation}
g(\theta) = \sum_{k \in \mathbb{Z}^n} \hat{f}(k)e^{ik \cdot \theta}, \quad g \in C(\mathbb{T}^n),
\end{equation}

and our task is to show that $f \equiv g$. Making use of the identities

\begin{equation}
(2\pi)^{-n} \int_{\mathbb{T}^n} e^{i\ell \cdot \theta} \, d\theta = 0, \quad \text{if} \ \ell \neq 0,
\end{equation}

\begin{equation}
1, \quad \text{if} \ \ell = 0,
\end{equation}

for $\ell \in \mathbb{Z}^n$, we get $\hat{g}(k) = \hat{f}(k)$ for all $k \in \mathbb{Z}^n$. Let us set $u = f - g$. We have

\begin{equation}
u \in C(\mathbb{T}^n), \quad \hat{u}(k) = 0, \quad \forall k \in \mathbb{Z}^n.
\end{equation}
It remains to show that this implies $u \equiv 0$. To prove this, we use Corollary E.4 (a consequence of the Stone-Weierstrass theorem, treated in Appendix E), which implies that for each $v \in C(\mathbb{T}^n)$, there exist trigonometric polynomials, i.e., finite linear combinations $v_N$ of $\{e^{ik\cdot \theta} : k \in \mathbb{Z}^n\}$, such that

$$v_N \longrightarrow v \text{ uniformly on } \mathbb{T}^n.$$

Now (13.7) implies

$$\int_{\mathbb{T}^n} u(\theta)v_N(\theta) \, d\theta = 0, \quad \forall N,$$

and passing to the limit, using (13.8), gives

$$\int_{\mathbb{T}^n} u(\theta)v(\theta) \, d\theta = 0, \quad \forall v \in C(\mathbb{T}^n).$$

Taking $v = u$ gives

$$\int_{\mathbb{T}^n} |u(\theta)|^2 \, d\theta = 0,$$

forcing $u \equiv 0$ and completing the proof.

We seek conditions on $f$ implying that $f \in \mathcal{A}(\mathbb{T}^n)$. Assume $f \in C^\ell(\mathbb{T}^n)$. Then integration by parts gives

$$\int_{\mathbb{T}^n} f^{(\alpha)}(\theta)e^{-ik\cdot \theta} \, d\theta = (ik)^\alpha \hat{f}(k), \quad \text{for } |\alpha| \leq \ell.$$

Hence

$$f \in C^\ell(\mathbb{T}^n) \Rightarrow |\hat{f}(k)| \leq C(1 + |k|)^{-\ell} \sum_{|\alpha| \leq \ell} \|f^{(\alpha)}\|_{L^1},$$

where we set

$$\|u\|_{L^1} = (2\pi)^{-n} \int_{\mathbb{T}^n} |u(\theta)| \, d\theta.$$

Since $\sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-(n+1)} < \infty$, we deduce that

$$C^{n+1}(\mathbb{T}^n) \subset \mathcal{A}(\mathbb{T}^n).$$

We will sharpen this implication below.

We next make use of (13.6) to produce results on $\int_{\mathbb{T}^n} |f(\theta)|^2 \, d\theta$, starting with the following.
Proposition 13.2. Given \( f \in \mathcal{A}(\mathbb{T}^n) \),

\[
\sum_{k \in \mathbb{Z}^n} |\hat{f}(k)|^2 = (2\pi)^{-n} \int_{\mathbb{T}^n} |f(\theta)|^2 \, d\theta.
\]

More generally, if also \( g \in \mathcal{A}(\mathbb{T}^n) \),

\[
\sum_{k \in \mathbb{Z}^n} \hat{f}(k)\overline{g(k)} = (2\pi)^{-n} \int_{\mathbb{T}^n} f(\theta)g(\theta) \, d\theta.
\]

Proof. Switching order of summation and integration and using (13.6), we have

\[
(2\pi)^{-n} \int_{\mathbb{T}^n} f(\theta)g(\theta) \, d\theta = (2\pi)^{-n} \int_{\mathbb{T}^n} \sum_{j,k \in \mathbb{Z}^n} \hat{f}(j)\overline{g(k)}e^{-i(j-k)\cdot \theta} \, d\theta
\]

\[
= \sum_k \hat{f}(k)\overline{g(k)},
\]

giving (13.17). Taking \( g = f \) gives (13.16).

We will extend the scope of Proposition 13.2 below. A related issue is the convergence of \( S_N f \) to \( f \) as \( N \to \infty \), where

\[
S_N f(\theta) = \sum_{|k| \leq N} \hat{f}(k)e^{i k \cdot \theta}.
\]

Here we take \( |k| = (k_1^2 + \cdots + k_n^2)^{1/2} \). Clearly \( f \in \mathcal{A}(\mathbb{T}^n) \Rightarrow S_n f \to f \) uniformly on \( \mathbb{T}^n \) as \( N \to \infty \). Here, we are interested in convergence in \( L^2 \)-norm, where

\[
\|f\|_{L^2}^2 = (2\pi)^{-n} \int_{\mathbb{T}^n} |f(\theta)|^2 \, d\theta.
\]

Given \( f \in \mathcal{R}(\mathbb{T}^n) \), this defines a “norm,” satisfying the following result, called the triangle inequality:

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

See Appendix H for details on this. Behind this result is the fact that

\[
\|f\|_{L^2}^2 = (f,f)_{L^2},
\]

where, when \( f, g \in L^2(\mathbb{T}^n) \), we set

\[
(f,g)_{L^2} = (2\pi)^{-n} \int_{\mathbb{T}^n} f(\theta)\overline{g(\theta)} \, d\theta.
\]
Thus the content of (13.16) is that

\[(13.24) \quad \sum_{k \in \mathbb{Z}^n} |\hat{f}(k)|^2 = \|f\|_{L^2}^2,\]

and that of (13.17) is that

\[(13.25) \quad \sum_{k \in \mathbb{Z}^n} \hat{f}(k)\bar{g}(k) = (f, g)_{L^2}.\]

The left side of (13.24) can be regarded as the square norm of an element of $\ell^2(\mathbb{Z}^n)$. Generally, an element $(a_k)_{k \in \mathbb{Z}^n}$ belongs to $\ell^2(\mathbb{Z}^n)$ if and only if

\[(13.26) \quad \|(a_k)\|_{\ell^2}^2 = \sum_{k \in \mathbb{Z}^n} |a_k|^2 < \infty.\]

There is an associated inner product

\[(13.27) \quad ((a_n), (b_n))_{\ell^2} = \sum_{k \in \mathbb{Z}^n} a_k \overline{b}_k.\]

As in (13.21), one has

\[(13.28) \quad \|(a_k) + (b_k)\|_{\ell^2}^2 \leq \|(a_k)\|_{\ell^2}^2 + \|(b_k)\|_{\ell^2}^2.\]

As for the notion of $L^2$-convergence, we say

\[(13.29) \quad f_\nu \to f \quad \text{in} \quad L^2 \iff \|f - f_\nu\|_{L^2} \to 0.\]

There is a similar notion of convergence in $\ell^2$. Clearly

\[(13.30) \quad \|f - f_\nu\|_{L^2} \leq \sup_{\theta} |f(\theta) - f_\nu(\theta)|.\]

In view of the uniform convergence $S_N f \to f$ for $f \in \mathcal{A}(\mathbb{T}^n)$, noted above, we have

\[(13.31) \quad f \in \mathcal{A}(\mathbb{T}^n) \implies S_N f \to f \quad \text{in} \quad L^2, \quad \text{as} \quad N \to \infty.\]

The triangle inequality implies

\[(13.32) \quad \|f\|_{L^2} - \|S_N f\|_{L^2} \leq \|f - S_N f\|_{L^2},\]

and clearly (by Proposition 13.2)

\[(13.33) \quad \|S_N f\|_{L^2}^2 = \sum_{|k| \leq N} |\hat{f}(k)|^2,\]
so
\[
\| f - S_N f \|_{L^2}^2 \to 0 \quad \text{as} \quad N \to \infty \implies \| f \|_{L^2}^2 = \sum_{k \in \mathbb{Z}^n} |\hat{f}(k)|^2.
\]

We now consider more general functions \( f \in \mathcal{R}(\mathbb{T}^n) \). With \( \hat{f}(k) \) and \( S_N f \) defined by (13.1) and (13.19), we define \( R_N f \) by
\[
(13.35) \quad f = S_N f + R_N f.
\]
Note that 
\[
\int_{\mathbb{T}^n} f(\theta)e^{-ik\cdot\theta} d\theta = \int_{\mathbb{T}^n} S_N f(\theta)e^{-ik\cdot\theta} d\theta \quad \text{for} \quad |k| \leq N.
\]
Hence
\[
(13.36) \quad (f, S_N f)_{L^2} = (S_N f, S_N f)_{L^2},
\]
and hence
\[
(13.37) \quad (S_N f, R_N f)_{L^2} = 0.
\]
Consequently,
\[
(13.38) \quad \| f \|_{L^2}^2 = (S_N f + R_N f, S_N f + R_N f)_{L^2} = \| S_N f \|_{L^2}^2 + \| R_N f \|_{L^2}^2.
\]
In particular,
\[
(13.39) \quad \| S_N f \|_{L^2} \leq \| f \|_{L^2}.
\]
We are now in a position to prove the following.

**Lemma 13.3.** Let \( f, f_\nu \in \mathcal{R}(\mathbb{T}^n) \). Assume
\[
(13.40) \quad \lim_{\nu \to \infty} \| f - f_\nu \|_{L^2} = 0,
\]
and, for each \( \nu \),
\[
(13.41) \quad \lim_{N \to \infty} \| f_\nu - S_N f_\nu \|_{L^2} = 0.
\]
Then
\[
(13.42) \quad \lim_{N \to \infty} \| f - S_N f \|_{L^2} = 0.
\]

**Proof.** Writing \( f - S_N f = (f - f_\nu) + (f_\nu - S_N f_\nu) + S_N (f_\nu - f) \) and using the triangle inequality, we have, for each \( \nu \),
\[
(13.43) \quad \| f - S_N f \|_{L^2} \leq \| f - f_\nu \|_{L^2} + \| f_\nu - S_N f_\nu \|_{L^2} + \| S_N (f_\nu - f) \|_{L^2}.
\]
Taking \( N \to \infty \) and using (13.39), we have
\[
(13.44) \quad \limsup_{N \to \infty} \| f - S_N f \|_{L^2} \leq 2\| f - f_\nu \|_{L^2},
\]
for each \( \nu \). Then (13.40) yields the desired conclusion (13.42).

Given \( f \in C(\mathbb{T}^n) \), we have trigonometric polynomials \( f_\nu \to f \) uniformly on \( \mathbb{T}^n \) (by Corollary E.4), and clearly (13.41) holds for each such \( f_\nu \). Thus Lemma 13.3 yields the following.
\[
f \in C(\mathbb{T}^n) \implies S_N f \to f \quad \text{in} \quad L^2,
\]
and
\[
(13.45) \quad \sum_{k \in \mathbb{Z}^n} |\hat{f}(k)|^2 = \| f \|_{L^2}^2.
\]
To go further, we bring in the following lemma.
Lemma 13.4. Given $f \in \mathcal{R}(\mathbb{T}^n)$, there exist $f_\nu \in C(\mathbb{T}^n)$ such that $f_\nu \to f$ in $L^2$.

Proof. One easily reduces this to treating the case $f \geq 0$. Then Proposition 4.11 implies that there exist $f_\nu \in C(\mathbb{T}^n)$ such that $0 \leq f_\nu \leq f$ and $\int_{\mathbb{T}^n} (f - f_\nu) \, d\theta \to 0$. Hence

$$\int_{\mathbb{T}^n} |f - f_\nu|^2 \, d\theta \leq (\sup f) \int_{\mathbb{T}^n} (f - f_\nu) \, d\theta \to 0.$$ 

The last two lemmas and (13.45) yield the following result.

Proposition 13.5. We have

$$f \in \mathcal{R}(\mathbb{T}^n) \implies S_N f \to f \quad \text{in} \quad L^2, \quad \text{and}$$

$$\sum_{k \in \mathbb{Z}^n} |\hat{f}(k)|^2 = \|f\|_{L^2}^2.$$ 

Having some general results guaranteeing the conclusion in (13.16), we now improve (13.15).

Proposition 13.6. We have

$$\ell > \frac{n}{2} \implies C^\ell(\mathbb{T}^n) \subset A(\mathbb{T}^n).$$ 

Proof. As before, we have (13.12). We deduce from this that

$$\sum_{k \in \mathbb{Z}^n} |k^\alpha \hat{f}(k)|^2 = \|f^{(\alpha)}\|_{L^2}^2, \quad \text{for} \quad |\alpha| \leq \ell.$$ 

Summing over $|\alpha| \leq \ell$ yields

$$\sum_{k \in \mathbb{Z}^n} (1 + |k|)^{2\ell} |\hat{f}(k)|^2 \leq C \sum_{|\alpha| \leq \ell} \|f^{(\alpha)}\|_{L^2}^2.$$ 

Then Cauchy’s inequality gives

$$\sum_{k \in \mathbb{Z}^n} |\hat{f}(k)| = \sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-\ell} (1 + |k|)^\ell |\hat{f}(k)|$$

$$\leq A_{n,\ell}^{1/2} \left\{ \sum_{k \in \mathbb{Z}^n} (1 + |k|)^{2\ell} |\hat{f}(k)|^2 \right\}^{1/2},$$

where

$$A_{n,\ell} = \sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-2\ell} < \infty, \quad \text{if} \quad 2\ell > n.$$
This establishes the desired finiteness of \( \sum |\hat{f}(k)| \).

It is illuminating to restate some of the results established above, bringing in the spaces \( \ell^p(\mathbb{Z}^n) \) of functions \( a : \mathbb{Z}^n \to \mathbb{C} \), with norm \( \| \cdot \|_{\ell^p} \), defined by

\[
\|a\|_{\ell^p}^p = \sum_{k \in \mathbb{Z}^n} |a(k)|^p < \infty,
\]

for \( 1 \leq p < \infty \), and

\[
\|a\|_{\ell^\infty} = \sup_k |a(k)| < \infty.
\]

The case \( p = 2 \) arose in (13.26)–(13.27). Also, we denote by \( \mathcal{F} \) the transformation that assigns to an integrable function \( f \) on \( \mathbb{T}^n \) its Fourier coefficients \( (\hat{f}(k)) \). The definition (13.1) readily gives

\[
\mathcal{F} : \mathcal{R}^\#(\mathbb{T}^n) \to \ell^\infty(\mathbb{Z}^n), \quad \|\mathcal{F}f\|_{\ell^\infty} \leq \|f\|_{L^1},
\]

with \( \|f\|_{L^1} \) defined as in (13.14). Proposition 13.5 gives

\[
\mathcal{F} : \mathcal{R}(\mathbb{T}^n) \to \ell^2(\mathbb{Z}^n), \quad \|\mathcal{F}f\|_{\ell^2} = \|f\|_{L^2}.
\]

Meanwhile, the definition (13.3) gives

\[
\mathcal{F} : \mathcal{A}(\mathbb{T}^n) \to \ell^1(\mathbb{Z}^n).
\]

We can restate Proposition 13.1 by bringing in the transformation \( \mathcal{F}^* \), defined on \( \ell^1(\mathbb{Z}^n) \) by

\[
\mathcal{F}^*(a)(\theta) = \sum_{k \in \mathbb{Z}^n} a(k)e^{ik\cdot\theta}.
\]

The content of Proposition 13.1 is that

\[
\mathcal{F}^* : \ell^1(\mathbb{Z}^n) \to \mathcal{A}(\mathbb{T}^n) \text{ is the 2-sided inverse of } \mathcal{F} \text{ in (13.56).}
\]

The step (13.7) in the proof of that result is equivalent to \( \mathcal{F}\mathcal{F}^* = I \) on \( \ell^1(\mathbb{Z}^n) \), and the reverse result, \( \mathcal{F}^*\mathcal{F} = I \) on \( \mathcal{A}(\mathbb{T}^n) \), made use of the Stone-Weierstrass theorem.

The analytical apparatus for an extension of (13.55) to a result involving \( \mathcal{F} \) and \( \mathcal{F}^* \) as inverses of each other was produced by H. Lebesgue in what is now known as the theory of Lebesgue measure and integration. We give a brief description of this, referring the reader to other sources, such as [Fol] or [T2], for a detailed presentation.

To start, we say a set \( S \subseteq \mathbb{T}^n \) is measurable provided that

\[
m^*(S) + m^*(\mathbb{T}^n \setminus S) = V(\mathbb{T}^n),
\]
where \( m^* \) is the outer measure defined by (4.118). We say a function \( f : \mathbb{T}^n \to \mathbb{C} \) is measurable provided that, for each open \( O \subset \mathbb{C} \), \( f^{-1}(O) \subset \mathbb{T}^n \) is measurable. The Lebesgue integral associates a value in \([0, +\infty]\) to

\[
\int_{\mathbb{T}^n} f(\theta) \, d\theta
\]

for each measurable \( f \) satisfying \( f(\theta) \geq 0 \) for all \( \theta \). The space \( L^1(\mathbb{T}^n) \) consists of all measurable functions \( f \) such that

\[
\|f\|_{L^1} = (2\pi)^{-n} \int_{\mathbb{T}^n} |f(\theta)| \, d\theta < \infty.
\]

In such a case, one can write \( f = f_{0+} - f_{0-} + i(f_{1+} - f_{1-}) \), with all \( f_{j\pm} \) measurable and \( \geq 0 \), and with finite integral, and the process alluded to above applies to evaluate these integrals, and hence to evaluate \( \int_{\mathbb{T}^n} f(\theta) \, d\theta \). There is one further wrinkle. The space \( L^1(\mathbb{T}^n) \) actually consists of equivalence classes of measurable functions satisfying (13.61), where one says \( f_1 \sim f_2 \) provided \( \{x \in \mathbb{T}^n : f_1(x) \neq f_2(x)\} \) has outer measure zero. This makes \( L^1(\mathbb{T}^n) \) a normed space.

More generally, for \( p \in [1, \infty) \), \( L^p(\mathbb{T}^n) \) consists of equivalence classes of measurable functions for which

\[
\|f\|_{L^p} = (2\pi)^{-n} \int_{\mathbb{T}^n} |f(\theta)|^p \, d\theta < \infty.
\]

The only case of \( p > 1 \) we work with here is \( p = 2 \). One has

\[
f, g \in L^2(\mathbb{T}^n) \implies fg \in L^1(\mathbb{T}^n),
\]

and \( L^2(\mathbb{T}^n) \) is an inner product space, via (13.23), extended to this setting. These \( L^p \) norms satisfy the triangle inequality

\[
\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.
\]

For \( p = 1 \), this inequality is a simple consequence of the pointwise inequality \( |f(\theta) + g(\theta)| \leq |f(\theta)| + |g(\theta)| \). For \( p = 2 \), the proof, as indicated in (13.21)–(13.23), follows from material in Appendix H. For other \( p \), which we do not need here, the reader can consult the references mentioned above. Thanks to (13.63), \( L^p(\mathbb{T}^n) \) has the structure of a metric space, with \( d(f, g) = \|f - g\|_{L^p} \). We say \( f_\nu \rightarrow f \) in \( L^p \) if \( \|f - f_\nu\|_{L^p} \to 0 \). For all \( p \in [1, \infty) \), these spaces have the following important metric properties. The first is a denseness property.

**Proposition A.** Given \( f \in L^p(\mathbb{T}^n) \) and \( \ell \in \mathbb{N} \), there exist \( f_\nu \in C^\ell(\mathbb{T}^n) \) such that \( f_\nu \rightarrow f \) in \( L^p \).

The second is a completeness property.
Proposition B. If \((f_\nu)\) is a Cauchy sequence in \(L^p(\mathbb{T}^n)\), then there exists \(f \in L^p(\mathbb{T}^n)\) such that \(f_\nu \to f\) in \(L^p\).

We refer to [Fol] or [T2] for proofs of these results.

The following two neat \(L^2\) results illustrate the usefulness of the Lebesgue theory of integration in Fourier analysis.

**Proposition 13.7.** The maps \(\mathcal{F}\) and \(\mathcal{F}^*\) have unique continuous linear extensions from

\[(13.65) \quad \mathcal{F} : \mathcal{A}(\mathbb{T}^n) \to \ell^1(\mathbb{Z}^n), \quad \mathcal{F}^* : \ell^1(\mathbb{Z}^n) \to \mathcal{A}(\mathbb{T}^n)\]

to

\[(13.66) \quad \mathcal{F} : L^2(\mathbb{T}^n) \to \ell^2(\mathbb{Z}^n), \quad \mathcal{F}^* : \ell^2(\mathbb{Z}^n) \to L^2(\mathbb{T}^n),\]

and the identities

\[(13.67) \quad \mathcal{F}^* \mathcal{F} f = f, \quad \mathcal{F} \mathcal{F}^* a = a\]

and

\[(13.68) \quad \|\mathcal{F} f\|_{\ell^2} = \|f\|_{L^2}, \quad \|\mathcal{F}^* a\|_{L^2} = \|a\|_{\ell^2}\]

hold for all \(f \in L^2(\mathbb{T}^n)\) and \(a \in \ell^2(\mathbb{Z}^n)\).

**Proposition 13.8.** Define \(S_N\) as in (13.19). Then \(S_N : L^2(\mathbb{T}^n) \to L^2(\mathbb{T}^n)\) and

\[(13.69) \quad f \in L^2(\mathbb{T}^n) \implies S_N f \to f \text{ in } L^2(\mathbb{T}^n).\]

Rather than presenting arguments here, we refer the reader to §14, and the analogous results, Propositions 14.10 and 14.11.

As a complement to Proposition 13.7, we mention that, using the inner products (13.23) and (13.27), we have

\[(13.70) \quad (a, \mathcal{F} f)_{\ell^2} = (\mathcal{F}^* a, f)_{L^2},\]

for all \(f \in L^2(\mathbb{T}^n), \ a \in \ell^2(\mathbb{Z}^n)\). In case \(f \in \mathcal{A}(\mathbb{T}^n)\) and \(a \in \ell^1(\mathbb{Z}^n)\), such an identity is elementary.

We now discuss another way of taking Fourier analysis beyond the setting of \(\mathcal{A}(\mathbb{T}^n)\), which is a normed space, with norm

\[(13.71) \quad \|f\|_A = \|\mathcal{F} f\|_{\ell^1},\]

for which we have

\[(13.72) \quad \|\mathcal{F}^* a\|_A = \|a\|_{\ell^1},\]
for all $a \in \ell^1(\mathbb{Z}^n)$. We define the dual space $\mathcal{A}'(\mathbb{T}^n)$ to consist of continuous linear functionals

$$w : \mathcal{A}(\mathbb{T}^n) \to \mathbb{C},$$

(13.73)
i.e., those linear maps satisfying

$$|w(f)| \leq C\|f\|_{\mathcal{A}}, \quad \forall f \in \mathcal{A}(\mathbb{T}^n).$$

(13.74)The optimal constant in (13.74) is denoted $\|w\|_{\mathcal{A}'}$. Other useful notations are

$$\langle f, w \rangle = w(f), \quad \text{and} \quad (f, w) = \overline{w(f)},$$

(13.75)for $f \in \mathcal{A}(\mathbb{T}^n)$, $w \in \mathcal{A}'(\mathbb{T}^n)$. We want to pass from $\mathcal{F}$ and $\mathcal{F}^*$ in (13.69) to

$$\mathcal{F} : \mathcal{A}'(\mathbb{T}^n) \to \ell^\infty(\mathbb{Z}^n), \quad \mathcal{F}^* : \ell^\infty(\mathbb{Z}^n) \to \mathcal{A}'(\mathbb{T}^n).$$

(13.76)To explain the use of $\ell^\infty(\mathbb{Z}^n)$, we establish the following.

**Lemma 13.9.** The dual space to $\ell^1(\mathbb{Z}^n)$ is $\ell^\infty(\mathbb{Z}^n)$.

**Proof.** What is to be established is a natural identification of the set of continuous linear functionals on $\ell^1(\mathbb{Z}^n)$,

$$\beta : \ell^1(\mathbb{Z}^n) \to \mathbb{C},$$

(13.77)with the set of bounded functions $b : \mathbb{Z}^n \to \mathbb{C},$

$$b \in \ell^\infty(\mathbb{Z}^n).$$

(13.78)The map $\ell^\infty(\mathbb{Z}^n) \to \ell^1(\mathbb{Z}^n)'$ is given by

$$\langle a, b \rangle = \sum_{k \in \mathbb{Z}^n} a_k b_k, \quad a \in \ell^1(\mathbb{Z}^n), \quad b \in \ell^\infty(\mathbb{Z}^n).$$

(13.79)Note that

$$|\langle a, b \rangle| \leq \left( \sup_k |b_k| \right) \sum_k |a_k| = \|a\|_{\ell^1} \|b\|_{\ell^\infty}.$$  (13.80)

Conversely, given $\beta$ as in (13.77), satisfying

$$|\beta(a)| \leq B\|a\|_{\ell^1},$$

(13.81)let us define

$$b_k = \beta(\varepsilon_k), \quad \text{where} \quad \varepsilon_k(\ell) = 1 \text{ if } \ell = k, \quad 0 \text{ if } \ell \neq k.$$  (13.82)
We have $|b_k| \leq B$ for all $k$, so $b = (b_k) \in \ell^\infty(Z^n)$, and $\|b\|_{\ell^\infty} \leq B$. Furthermore,

\[(13.83) \quad \langle a, b \rangle = \beta(a),\]

for $a = \varepsilon_k$, for each $k \in Z^n$, and hence (13.83) holds for all $a \in \ell^1(Z^n)$. This proves the lemma.

**Remark.** An inspection of the proof shows that the correspondence $\beta \leftrightarrow b$ of $\ell^1(Z^n)'$ with $\ell^\infty(Z^n)$ is norm preserving, i.e., the optional constant $\|\beta\|$ in (13.81) is equal to $\|b\|_{\ell^\infty}$.

We are now ready to define $\mathcal{F}$ and $\mathcal{F}^*$ in (13.76). Taking off from (13.70), we define $\mathcal{F}w \in \ell^\infty(Z^n)$ for $w \in A'(T^n)$ by

\[(13.84) \quad (a, \mathcal{F}w) = (\mathcal{F}^*a, w), \quad a \in \ell^1(Z^n),\]

and we define $\mathcal{F}^*b \in A'(T^n)$ for $b \in \ell^\infty(Z^n)$ by

\[(13.85) \quad (f, \mathcal{F}^*b) = (\mathcal{F}f, b), \quad f \in A(T^n).\]

Note that (13.84) yields

\[(13.86) \quad |(a, \mathcal{F}w)| \leq \|w(\mathcal{F}a)\| \leq \|w\|_{A'} \|\mathcal{F}^*a\|_A = \|w\|_{A'} \|a\|_{\ell^1},\]

the last identity by (13.72). Hence, by Lemma 13.9 and the remark following it,

\[(13.87) \quad \|\mathcal{F}w\|_{\ell^\infty} \leq \|w\|_{A'}.\]

Next, (13.85) yields

\[(13.88) \quad |(f, \mathcal{F}^*b)| \leq \|\mathcal{F}f\|_{\ell^1} \|b\|_{\ell^\infty} = \|f\|_A \|b\|_{\ell^\infty},\]

the last identity by (13.71), hence

\[(13.89) \quad \|\mathcal{F}^*b\|_{A'} \leq \|b\|_{\ell^\infty}.\]

Thus $\mathcal{F}$ and $\mathcal{F}^*$ in (13.76) are well defined.

The following Fourier inversion formula complements those in Propositions 13.1 and 13.7.

**Proposition 13.10.** The maps $\mathcal{F}$ and $\mathcal{F}^*$ in (13.76) are two-sided inverses of each other, i.e.,

\[(13.90) \quad \mathcal{F}^*\mathcal{F}w = w \quad \text{and} \quad \mathcal{F}\mathcal{F}^*b = b, \quad \forall w \in A'(T^n), \quad b \in \ell^\infty(Z^n).\]

**Proof.** Given $f \in A(T^n)$, $a \in \ell^1(Z^n)$, we have

\[(13.91) \quad (f, \mathcal{F}^*\mathcal{F}w) = (\mathcal{F}f, \mathcal{F}w) = (\mathcal{F}^*\mathcal{F}f, w) = (f, w),\]

the first identity by (13.85), the second by (13.84), and the third by (13.58). This implies $\mathcal{F}^*\mathcal{F}w = w$. Similarly,

\[(13.92) \quad (a, \mathcal{F}\mathcal{F}^*b) = (\mathcal{F}^*a, \mathcal{F}^*b) = (\mathcal{F}\mathcal{F}^*a, b) = (a, b),\]

yielding $\mathcal{F}\mathcal{F}^*b = b$.

We can now sharpen (13.87) and (13.89).
\textbf{Corollary 13.11.} For \( w \in \mathcal{A}'(\mathbb{T}^n), \ b \in \ell^\infty(\mathbb{Z}^n), \)

\begin{equation}
\|Fw\|_{\ell^\infty} = \|w\|_{\mathcal{A}'}, \quad \text{and} \quad \|F^*b\|_{\mathcal{A}'} = \|b\|_{\ell^\infty}.
\end{equation}

\textit{Proof.} First, \( w = F^*Fw \Rightarrow \|w\|_{\mathcal{A}'} \leq \|Fw\|_{\ell^\infty}, \) by (13.89) with \( b = Fw \). This together with (13.87) yields the first identity in (13.93). The second identity has a similar proof.

Let us note that (13.84), applied to \( a = \varepsilon_k \), defined as in (13.82), gives, for \( w \in \mathcal{A}'(\mathbb{T}^n), \)

\begin{equation}
Fw(k) = (2\pi)^{-n}\langle e_k, w \rangle, \quad e_k \in \mathcal{A}(\mathbb{T}^n), \quad e_k(\theta) = e^{-i\theta k}.
\end{equation}

The space \( \mathcal{A}'(\mathbb{T}^n) \) contains an interesting variety of functions and other objects. For example, there is a natural map

\begin{equation}
\iota : \mathcal{R}^\#(\mathbb{T}^n) \rightarrow \mathcal{A}'(\mathbb{T}^n),
\end{equation}

given by

\begin{equation}
\langle f, \iota(u) \rangle = (2\pi)^{-n} \int_{\mathbb{T}^n} f(\theta)u(\theta) \, d\theta.
\end{equation}

We have

\begin{equation}
|\langle f, \iota(u) \rangle| \leq (2\pi)^{-n}\left(\sup |f|\right)\|u\|_{L^1} \leq (2\pi)^{-n}\|f\|_{\mathcal{A}'}\|u\|_{L^1}.
\end{equation}

Given the results on the Lebesgue integral mentioned above, this extends to

\begin{equation}
\iota : L^1(\mathbb{T}^n) \rightarrow \mathcal{A}'(\mathbb{T}^n).
\end{equation}

We have from (13.94) that

\begin{equation}
F\iota(u)(k) = (2\pi)^{-n} \int_{\mathbb{T}^n} u(\theta)e^{-i\theta k} \, d\theta, \quad k \in \mathbb{Z}^n.
\end{equation}

The space \( \mathcal{A}'(\mathbb{T}^n) \) also contains some objects more singular than functions. For example, given \( p \in \mathbb{T}^n \), we define \( \delta_p \in \mathcal{A}'(\mathbb{T}^n) \) by

\begin{equation}
\langle f, \delta_p \rangle = f(p).
\end{equation}

More generally, let \( M \subset \mathbb{T}^n \) be a compact, \( m \)-dimensional, \( C^1 \) surface, and \( u \in C(M). \) Define \( u\delta_M \in \mathcal{A}'(\mathbb{T}^n) \) by

\begin{equation}
\langle f, u\delta_M \rangle = \int_M f(x)u(x) \, dS(x).
\end{equation}
We have

\[(13.102) \quad |\langle f, u\delta_M \rangle| \leq C_M (\sup |u|)(\sup |f|) \leq C_M (\sup |u|)\|f\|_A. \]

Again (13.94) applies, to give

\[(13.103) \quad \mathcal{F}(u\delta_M)(k) = (2\pi)^{-n} \int_M u(\theta)e^{-ik\cdot\theta} dS(\theta), \quad k \in \mathbb{Z}^n. \]

Objects such as \(\delta_p\) and \(\delta_M\) are examples of “distributions.” A beautiful theory of Fourier analysis on the space of distributions on \(\mathbb{T}^n\), which includes some more singular objects, was constructed by L. Schwartz. We present some of his results here. We start with the space \(C^\infty(\mathbb{T}^n)\). By (13.12)–(13.13),

\[(13.104) \quad f \in C^\infty(\mathbb{T}^n) \implies \mathcal{F}f \in s(\mathbb{Z}^n), \]

where

\[(13.105) \quad s(\mathbb{Z}^n) = \{a \in \ell^\infty(\mathbb{Z}^n) : |a(k)| \leq C_N (1 + |k|)^{-N}, \forall N\}. \]

It is also easy to see that

\[(13.106) \quad \mathcal{F}^* : s(\mathbb{Z}^n) \rightarrow C^\infty(\mathbb{T}^n), \]

and (13.58) specializes, to yield

\[(13.107) \quad \mathcal{F}^* \mathcal{F}f = f, \quad \mathcal{F} \mathcal{F}^* a = a, \quad \forall f \in C^\infty(\mathbb{T}^n), \quad a \in s(\mathbb{Z}^n). \]

The space \(C^\infty(\mathbb{T}^n)\) carries the following sequence of norms:

\[(13.108) \quad p_k(f) = \max_{|\alpha| \leq k} \sup_{\theta \in \mathbb{T}^n} |f^{(\alpha)}(\theta)|, \]

and the space \(s(\mathbb{Z}^n)\) carries the norms

\[(13.105) \quad q_\nu(a) = \sup_{k \in \mathbb{Z}^n} (1 + |k|)^\nu |a(k)|. \]

We see from (13.13) that

\[(13.110) \quad q_\nu(\mathcal{F}f) \leq C_\nu p_\nu(f). \]

Also, since \(\sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-(n+1)} < \infty\), we have

\[(13.111) \quad p_0(\mathcal{F}^* a) \leq C q_{n+1}(a), \]
and from this we get

\[(13.112) \quad p_k(F^*a) \leq Cq_{k+n+1}(a).\]

Now a distribution \( w \in \mathcal{D}'(\mathbb{T}^n) \) is a continuous linear functional

\[(13.113) \quad w : C^\infty(\mathbb{T}^n) \to \mathbb{C},\]

that is to say, \( w \) is a linear map from \( C^\infty(\mathbb{T}^n) \) to \( \mathbb{C} \) with the property that there exist \( k \in \mathbb{Z}^+ \) and \( C < \infty \) such that

\[(13.114) \quad |w(f)| \leq Cp_k(f), \quad \forall f \in C^\infty(\mathbb{T}^n).\]

As in (13.75), we use the notation

\[(13.115) \quad \langle f, w \rangle = w(f), \quad (f, w) = \overline{w(f)}.\]

It follows from (13.15) that

\[(13.116) \quad \|f\|_A \leq Cp_{n+1}(f),\]

so each \( w \in A'(\mathbb{T}^n) \) also defines an element of \( \mathcal{D}'(\mathbb{T}^n) \). Thus \( \delta_p \) in (13.100) and \( \delta_M \) in (13.101) are examples of distributions on \( \mathbb{T}^n \). To produce more singular distributions, we can define

\[(13.117) \quad \partial^\alpha : \mathcal{D}'(\mathbb{T}^n) \to \mathcal{D}'(\mathbb{T}^n),\]

by

\[(13.118) \quad \langle f, \partial^\alpha w \rangle = (-1)^{|\alpha|}(f^{(\alpha)}, w).\]

We seek to define

\[(13.119) \quad \mathcal{F} : \mathcal{D}'(\mathbb{T}^n) \to s'(\mathbb{Z}^n), \quad \mathcal{F}^* : s'(\mathbb{Z}^n) \to \mathcal{D}'(\mathbb{T}^n),\]

where

\[(13.120) \quad s'(\mathbb{Z}^n) = \{ b : (1 + |k|)^{-N}b \in \ell^\infty(\mathbb{Z}^n), \text{ for some } N \in \mathbb{Z}^+ \} \].

The significance of \( s'(\mathbb{Z}^n) \) is explained by the following.
Lemma 13.12. The dual space to \( s(\mathbb{Z}^n) \) is \( s'(\mathbb{Z}^n) \).

Proof. What we claim is that there is a natural identification of the set \( s(\mathbb{Z}^n)' \) of continuous linear functionals on \( s(\mathbb{Z}^n) \),

\[
\beta : s(\mathbb{Z}^n) \rightarrow \mathbb{C},
\]

with the set of functions \( s'(\mathbb{Z}^n) \),

\[
b \in s'(\mathbb{Z}^n).
\]

The condition of continuity on \( \beta \) in (13.121) is that there exist \( \nu \in \mathbb{N} \) and \( C < \infty \) such that

\[
|\beta(a)| \leq C q_\nu(a).
\]

The map \( s'(\mathbb{Z}^n) \rightarrow s(\mathbb{Z}^n)' \) is given by

\[
\langle a, b \rangle = \sum_{k \in \mathbb{Z}^n} a_k b_k, \quad a \in s(\mathbb{Z}^n), \quad b \in s'(\mathbb{Z}^n).
\]

Note that

\[
|\langle a, b \rangle| \leq C \left( \sup_k (1 + |k|)^{N+n+1}|a(k)| \right) \left( \sup_k (1 + |k|)^{-N}|b(k)| \right).
\]

Conversely, given \( \beta \) as in (13.121), satisfying (13.123), we define

\[
b_k = \beta(\varepsilon_k),
\]

with \( \varepsilon_k \) as in (13.82), and verify that \( b \in s'(\mathbb{Z}^n) \) and that

\[
\langle a, b \rangle = \beta(a),
\]

first for \( a = \varepsilon_k \), for all \( k \in \mathbb{Z}^n \), and then for all \( a \in s(\mathbb{Z}^n) \).

We are now ready to define \( \mathcal{F} \) and \( \mathcal{F}^* \) in (13.119). Parallel to (13.84), we define \( \mathcal{F}w \in s'(\mathbb{Z}^n) \) for \( w \in \mathcal{D}'(\mathbb{T}^n) \) by

\[
(a, \mathcal{F}w) = (\mathcal{F}^*a, w), \quad a \in s(\mathbb{Z}^n),
\]

and we define \( \mathcal{F}^*b \in \mathcal{D}'(\mathbb{T}^n) \) for \( b \in s'(\mathbb{Z}^n) \) by

\[
(f, \mathcal{F}^*b) = (\mathcal{F}f, b), \quad f \in C^\infty(\mathbb{T}^n).
\]

As we have seen, \( a \in s(\mathbb{Z}^n) \Rightarrow \mathcal{F}^*a \in s(\mathbb{Z}^n) \), with estimates (13.112), and \( f \in C^\infty(\mathbb{T}^n) \Rightarrow \mathcal{F}f \in s(\mathbb{Z}^n) \), with estimates (13.110). These results enable one to deduce that (13.128) defines \( \mathcal{F}w \in s'(\mathbb{Z}^n) \) and (13.129) defines \( \mathcal{F}^*b \in \mathcal{D}'(\mathbb{T}^n) \).

Here is a further extension of the Fourier inversion formula.
**Proposition 13.13.** The maps $\mathcal{F}$ and $\mathcal{F}^*$ in (13.119) are two-sided inverses of each other, i.e.,

\[(13.130) \quad \mathcal{F}^* \mathcal{F} w = w, \quad \text{and} \quad \mathcal{F} \mathcal{F}^* b = b, \quad \forall w \in \mathcal{D}'(\mathbb{T}^n), \ b \in s'(\mathbb{Z}^n).\]

The proof is parallel to that of Proposition 13.10. The result (13.12) implies

\[(13.131) \quad \mathcal{F} f^{(\alpha)}(k) = (ik)^{\alpha} \mathcal{F} f(k), \quad \forall f \in C^\infty(\mathbb{T}^n).\]

We can extend this to $\mathcal{D}'(\mathbb{T}^n)$, using (13.117)–(13.118) and (13.128)–(13.129), to obtain:

**Proposition 13.14.** Given $w \in \mathcal{D}'(\mathbb{T}^n)$,

\[(13.132) \quad \mathcal{F} (\partial^\alpha w)(k) = (ik)^{\alpha} \mathcal{F} w(k).\]

**Exercises**

1. Consider $f(\theta) = |\theta|$ for $-\pi \leq \theta \leq \pi$, extended periodically to define an element of $C(\mathbb{T}^1)$.

   (a) Compute $\hat{f}(k)$.
   
   (b) Show that $f \in \mathcal{A}(\mathbb{T}^1)$.
   
   (c) Use the Fourier inversion formula (13.4) at $\theta = 0$ to show that

   \[(13.133) \quad \sum_{\ell=1}^{\infty} \frac{1}{(2\ell + 1)^2} = \frac{\pi^2}{8}.\]

   (d) Deduce from (c) that

   \[(13.134) \quad \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.\]

   *Hint.* Decompose this sum into the sum over $k$ odd and the sum over $k$ even.

   *Remark.* $\sum_{k=1}^{\infty} k^{-2}$ is $\zeta(2)$.

2. Consider $g(\theta) = 1$ for $0 < \theta < \pi$, $0$ for $-\pi < \theta < 0$, defining a bounded integrable function on $\mathbb{T}^1$.

   (a) Compute $\hat{g}(k)$.
   
   (b) Use the Plancherel identity (13.46) to obtain another derivation of (13.133).

3. Apply the Plancherel identity to $f$ and $\hat{f}$ in Exercise 1. Use this to show that

   \[(13.135) \quad \sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}.\]
Note. This sum is $\zeta(4)$.

4. Given $f \in \mathcal{R}(\mathbb{T}^1)$, set

$$P_r f(\theta) = \sum_{k \in \mathbb{Z}} r^{|k|} \hat{f}(k)e^{ik\theta}, \quad 0 \leq r < 1.$$ 

Show that

$$\|P_r f - f\|_{L^2} \to 0, \quad \text{as} \quad r \to 1.$$ 

5. In the setting of Exercise 4, show that

$$P_r f(\theta) = \sum_{k=0}^{\infty} \hat{f}(k)z^k + \sum_{k=1}^{\infty} \hat{f}(-k)z^{-k}, \quad z = re^{i\theta}.$$ 

Deduce that $u(re^{i\theta}) = P_r f(\theta)$ is harmonic on the unit disk $D = \{z \in \mathbb{C} : |z| < 1\}$.

6. Interpret the results of Exercises 4–5 as providing a solution to the Dirichlet boundary problem

$$\Delta u = 0 \quad \text{on} \quad D, \quad u\big|_{\partial D} = f.$$ 

7. In the setting of Exercises 4–6, show that

$$P_r f(\theta) = \int_0^{2\pi} f(\varphi)p(r, \theta - \varphi) \, d\varphi,$$

where

$$p(r, \theta) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r^{|k|}e^{ik\theta}.$$ 

Decompose this sum into a sum of two geometric series and show that

$$p(r, \theta) = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r \cos \theta + r^2}.$$ 

8. Show that

$$\int_{\mathbb{T}^1} p(r, \theta) \, d\theta = 1, \quad \forall r \in [0, 1),$$

and, for all $\delta > 0$,

$$p(r, \theta) \to 0 \quad \text{uniformly for} \quad \theta \in [-\pi, \pi] \setminus [-\delta, \delta], \quad \text{as} \quad r \to 1.$$ 

*Hint.* For the first part, integrate the series for $p(r, \theta)$ term by term.

9. Show that if $f \in C(\mathbb{T}^1)$, then $P_r f \to f$ uniformly on $\mathbb{T}^1$, as $r \to 1$.

*Hint.* Look forward, to Lemma 14.3.

10. Adapt the proof of Lemma 14.3 to establish the following.
**Proposition X.** If $f \in \mathcal{R}^\#(T^1)$, $I \subset T^1$ is an open set on which $f$ is continuous, and $K \subset I$ is compact, then $P_r f \to f$ uniformly on $K$, as $r \nearrow 1$.

For comparison, we mention the following result, given in Proposition 4.12 in Chapter 5 of [T4].

**Proposition Y.** If $f \in \mathcal{R}^\#(T^1)$ and $I \subset T^1$ is an open set on which $f$ is Hölder continuous, with some positive exponent, then $S_N f(\theta) \to f(\theta)$ as $N \to \infty$, for each $\theta \in I$.

Here $S_N f$ is as in (13.19), which for $n = 1$ is

$$S_N f(\theta) = \sum_{k=-N}^{N} \hat{f}(k)e^{ik\theta}.$$  

11. From the geometric series $\sum_{k=0}^{\infty} z^k = 1/(1 - z)$, deduce that

$$\sum_{k=1}^{\infty} \frac{z^k}{k} = \log \frac{1}{1 - z}, \quad \text{for } |z| < 1.$$  

12. Show that

$$f(\theta) = \log \frac{1}{1 - e^{i\theta}} \implies f \in \mathcal{R}^\#(T^1).$$  

Set

$$f_r(\theta) = \log \frac{1}{1 - re^{i\theta}}, \quad \text{for } 0 \leq r < 1.$$  

Show that $f_r \in C(T^1)$ for each $r \in [0,1)$ and

$$\|f - f_r\|_{L^1} \to 0, \quad \text{as } r \nearrow 1.$$  

Deduce that

$$\hat{f}_r(k) \to \hat{f}(k) \quad \text{as } r \nearrow 1, \quad \text{for each } k \in \mathbb{Z}.$$  

13. In the setting of Exercise 12, note that, for $0 \leq r < 1$,

$$\hat{f}_r(k) = \frac{1}{2\pi} \int_{0}^{2\pi} \left( \log \frac{1}{1 - re^{i\theta}} \right) e^{-ik\theta} d\theta.$$  

Input the power series from Exercise 11 and deduce (via Exercise 12) that

$$\hat{f}(k) = \begin{cases} \frac{1}{k}, & \text{if } k \geq 1, \\ 0, & \text{if } k \leq 0. \end{cases}$$
14. Deduce from Exercise 13 and Proposition Y that there is pointwise convergence (though not absolute convergence)

\[ \sum_{k=1}^{\infty} \frac{z^k}{k} = \log \frac{1}{1-z}, \quad \text{for } |z| = 1, \ z \neq 1. \]

Note that taking \( z = -1 \) yields

\[ 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2. \]

For a more elementary approach to this special case, see Exercise 30 for Chapter 4, §5, of [T4].

15. Let \( f \in C(\mathbb{T}^n) \) and assume that there exist \( f_\nu \in \mathcal{A}(\mathbb{T}^n) \) such that \( f_\nu \to f \) uniformly, and, for some \( K < \infty \), independent of \( \nu \),

\[ \| f_\nu \|_{\ell^1} \leq K. \]

Show that \( f \in \mathcal{A}(\mathbb{T}^n) \). Show that this holds if there exists \( m > n/2 \) and \( K_1 < \infty \) such that each \( f_\nu \in C^m(\mathbb{T}^n) \) and

\[ \sum_{|\alpha| \leq m} \| f_\nu^{(\alpha)} \|_{L^2} \leq K_1. \]

16. Deduce from Exercise 15 that

\[ \text{Lip}(\mathbb{T}^1) \subseteq \mathcal{A}(\mathbb{T}^1). \]

Note the improvement over (13.47) (in case \( n = 1 \)). Note that this result applies to the function in Exercise 1.
14. The Fourier transform

Given \( f \in \mathcal{R}(\mathbb{R}^n) \) (or more generally \( f \in \mathcal{R}^\#(\mathbb{R}^n) \)), we define its Fourier transform to be

\[
\hat{f}(\xi) = \mathcal{F}f(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-ix \cdot \xi} \, dx, \quad x \in \mathbb{R}^n.
\]

Similarly, we set

\[
\mathcal{F}^* f(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{ix \cdot \xi} \, dx, \quad \xi \in \mathbb{R}^n,
\]

and ultimately plan to identify \( \mathcal{F}^* \) as the inverse Fourier transform. Here, of course we take \( \mathcal{R}(\mathbb{R}^n) \) to consist of complex-valued functions. Recall from §4 that this means their real and imaginary parts are separately Riemann integrable.

Clearly

\[
|\hat{f}(\xi)| \leq (2\pi)^{-n/2} \|f\|_{L^1},
\]

where we use the notation

\[
\|f\|_{L^1} = \int_{\mathbb{R}^n} |f(x)| \, dx.
\]

We also have continuity.

**Proposition 14.1.** If \( f \in \mathcal{R}(\mathbb{R}^n) \), then \( \hat{f} \) is continuous on \( \mathbb{R}^n \).

**Proof.** Given \( \varepsilon > 0 \), pick \( N < \infty \) such that \( \int_{|x|>N} |f(x)| \, dx < \varepsilon \). Write \( f = f_N + g_N \) where \( f_N(x) = f(x) \) for \( |x| \leq N \), 0 for \( |x| > N \). Then

\[
\hat{f}(\xi) = \hat{f}_N(\xi) + \hat{g}_N(\xi),
\]

and

\[
|\hat{g}_N(\xi)| < \varepsilon, \quad \forall \xi.
\]

Meanwhile, for \( \xi, \zeta \in \mathbb{R}^n \),

\[
\hat{f}_N(\xi) - \hat{f}_N(\zeta) = (2\pi)^{-n/2} \int_{|x| \leq N} f(x) \left(e^{-ix \cdot \xi} - e^{-ix \cdot \zeta}\right) \, dx,
\]
and

\[ |e^{-ix\cdot \xi} - e^{-ix\cdot \zeta}| \leq |\xi - \zeta| \max_{\eta} |\nabla_\eta e^{-ix\cdot \eta}| \]

(14.8)

\[ \leq |x| \cdot |\xi - \zeta| \]

\[ \leq N|\xi - \zeta|, \]

for \(|x| \leq N\), so

(14.9)

\[ |\hat{f}_N(\xi) - \hat{f}_N(\zeta)| \leq \frac{N}{(2\pi)^{n/2}} \|f\|_{L^1} |\xi - \zeta|. \]

Hence each \(\hat{f}_N\) is continuous, and, by (14.6), \(\hat{f}\) is a uniform limit of continuous functions, so it is continuous.

**Remark.** Proposition 14.1 also holds for \(f \in \mathcal{R}^\#(\mathbb{R}^n)\).

We compute some Fourier transforms, making use of the result (10.49) that

(14.10)

\[ \int_{-\infty}^{\infty} e^{-x^2+ixz} \, dx = \sqrt{\pi} e^{z^2/4}, \quad \forall z \in \mathbb{C}. \]

Taking \(z = i\xi, \xi \in \mathbb{R}\), gives

(14.11)

\[ \int_{-\infty}^{\infty} e^{x^2+ix\xi} \, dx = \sqrt{\pi} e^{-\xi^2/4}, \quad \forall \xi \in \mathbb{R}. \]

Writing

(14.12)

\[ e^{-|x|^2+ix\cdot \xi} = e^{-x_1^2+ix_1\xi_1} \ldots e^{-x_n^2+ix_n\xi_n}, \]

we get

(14.13)

\[ \int_{\mathbb{R}^n} e^{-|x|^2} e^{ix\cdot \xi} \, dx = \pi^{n/2} e^{-|\xi|^2/4}, \quad \forall \xi \in \mathbb{R}^n. \]

Thus

(14.14)

\[ g(x) = e^{-|x|^2} \text{ on } \mathbb{R}^n \implies \mathcal{F}g(\xi) = \mathcal{F}^* g(\xi) = 2^{-n/2} e^{-|\xi|^2/4}. \]

A change of variable gives generally, for \(f \in \mathcal{R}(\mathbb{R}^n), a > 0\),

(14.15)

\[ f_a(x) = f(ax) \implies \mathcal{F}f_a(\xi) = a^{-n} \hat{f}(a^{-1}\xi), \]

and consequently, for \(b > 0\)

(14.16)

\[ g_b(x) = e^{-b|x|^2} \text{ on } \mathbb{R}^n \implies \mathcal{F}g_b(\xi) = \mathcal{F}^* g_b(\xi) = (2b)^{-n/2} e^{-|\xi|^2/4b}. \]
From (13.16) we see that $Fg_{1/2} = g_{1/2}$ and also that $F^*Fg_0 = F^*Fg_0 = g_0$.

The Fourier inversion formula asserts that

$$(14.17) \quad f(x) = (2\pi)^{-1/2} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{ix \cdot \xi} \, d\xi,$$

in appropriate senses, depending on the nature of $f$. We will approach this by examining

$$(14.18) \quad J_\varepsilon f(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{-\varepsilon|\xi|^2} e^{ix \cdot \xi} \, d\xi,$$

with $\varepsilon > 0$. By (14.3), $\hat{f}(\xi)e^{-\varepsilon|\xi|^2}$ is Riemann integrable over $\mathbb{R}^n$ whenever $f \in \mathcal{R}(\mathbb{R}^n)$. We can plug in (14.1) for $\hat{f}(\xi)$ and switch order of integration, getting

$$(14.19) \quad J_\varepsilon f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \int f(y) e^{i(x-y) \cdot \xi} e^{-\varepsilon|\xi|^2} \, dy \, d\xi$$

$$= \int f(y) H_\varepsilon(x-y) \, dy,$$

where

$$(14.20) \quad H_\varepsilon(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-\varepsilon|\xi|^2 + ix \cdot \xi} \, d\xi.$$

Using (14.16), we have

$$(14.21) \quad H_\varepsilon(x) = (4\pi\varepsilon)^{-n/2} e^{-|x|^2/4\varepsilon}.$$

A change of variable and use of $\int_{\mathbb{R}^n} e^{-|x|^2} \, dx = \pi^{n/2}$ gives

$$(14.22) \quad \int_{\mathbb{R}^n} H_\varepsilon(x) \, dx = 1, \quad \forall \varepsilon > 0.$$

Using this information, we will be able to prove the following.

**Proposition 14.2.** Assume $f$ is bounded and continuous on $\mathbb{R}^n$, and take $J_\varepsilon f(x) = \int f(y) H_\varepsilon(x-y) \, dy$, with $H_\varepsilon$ as in (14.21)-(14.22). Then, as $\varepsilon \searrow 0$,

$$(14.23) \quad J_\varepsilon f(x) \rightarrow f(x), \quad \forall x \in \mathbb{R}^n.$$

It is convenient to put this result in a more general context. If $f$ is bounded and continuous on $\mathbb{R}^n$ and $h \in \mathcal{R}(\mathbb{R}^n)$, we define the convolution $h \ast f$ by

$$(14.24) \quad h \ast f(x) = \int_{\mathbb{R}^n} h(y) f(x-y) \, dy.$$
Clearly

(14.25) \[ \int |h| \, dx = A, \quad |f| \leq M \text{ on } \mathbb{R}^n \implies |h \ast f| \leq AM \text{ on } \mathbb{R}^n. \]

Also, a change of variables gives

(14.26) \[ h \ast f(x) = \int_{\mathbb{R}^n} h(x - y)f(y) \, dy. \]

We want to analyze the convolution action of a family of integrable functions \( h_\nu \) on \( \mathbb{R}^n \) that satisfy the following conditions:

(14.27) \[ h_\nu \geq 0, \quad \int h_\nu \, dx = 1, \quad \int_{\mathbb{R}^n \setminus S_\nu} h_\nu \, dx \leq \varepsilon_\nu \to 0, \]

where

(14.28) \[ S_\nu = \{ x \in \mathbb{R}^n : |x| \leq \delta_\nu \}, \quad \delta_\nu \to 0. \]

Assume

(14.29) \[ f \in C(\mathbb{R}^n), \quad |f| \leq M \text{ on } \mathbb{R}^n. \]

We aim to prove the following.

**Lemma 14.3.** If \( h_\nu \in \mathcal{R}(\mathbb{R}^n) \) satisfy (14.27)–(14.28) and if \( f \in C(\mathbb{R}^n) \) satisfies (14.29), then

(14.30) \[ f_\nu(x) = h_\nu \ast f(x) \rightarrow f(x), \quad \forall x \in \mathbb{R}^n, \text{ locally uniformly in } x. \]

**Proof.** Given that \( f \) is continuous, it is uniformly continuous on compact sets, so we can supplement (14.29) with

(14.32) \[ |x - x'| \leq \delta_\nu, \quad |x| \leq R \implies |f(x) - f(x')| \leq \bar{\varepsilon}_\nu(R) \to 0, \]

for each \( R < \infty \). To proceed, write

(14.32) \[ f_\nu(x) = \int_{S_\nu} h_\nu(y)f(x - y) \, dy + \int_{\mathbb{R}^n \setminus S_\nu} h_\nu(y)f(x - y) \, dy \]

\[ = g_\nu(x) + r_\nu(x). \]

Clearly

(14.33) \[ |r_\nu(x)| \leq M\varepsilon, \quad \forall x \in \mathbb{R}^n. \]
Next,
\[(14.34) \quad g_\nu(x) - f(x) = \int_{S_\nu} h_\nu(y) [f(x - y) - f(x)] \, dy - \varepsilon_\nu f(x),\]
so
\[(14.35) \quad |g_\nu(x) - f(x)| \leq \varepsilon_\nu(R) + M\varepsilon_\nu, \quad \text{for } |x| \leq R,\]
hence
\[(14.36) \quad |f_\nu(x) - f(x)| \leq \varepsilon_\nu(R) + 2M\varepsilon_\nu, \quad \text{for } |x| \leq R,\]
yielding (14.30).

In view of (14.21)–(14.22), Proposition 14.2 follows from Lemma 14.3. From here, we obtain the following.

**Proposition 14.4.** Assume \(f\) is bounded and continuous on \(\mathbb{R}^n\), and \(\hat{f} \in \mathcal{R}(\mathbb{R}^n)\). Then the Fourier inversion formula (14.17) holds for all \(x \in \mathbb{R}^n\).

**Proof.** If \(f \in \mathcal{R}(\mathbb{R}^n)\), then \(\hat{f}\) is bounded and continuous. If also \(\hat{f} \in \mathcal{R}(\mathbb{R}^n)\), then the right side of (14.18) converges to the right side of (14.17), i.e., to \(\mathcal{F}^*\hat{f}(x)\), for each \(x \in \mathbb{R}^n\), as \(\varepsilon \searrow 0\). That is to say,

\[(14.37) \quad \lim_{\varepsilon \searrow 0} J_\varepsilon f(x) = \mathcal{F}^*\hat{f}(x), \quad \forall x \in \mathbb{R}^n.\]

In concert with (14.23), this proves the proposition.

**Remark.** With some more work, one can omit the hypothesis in Proposition 14.4 that \(f\) be bounded and continuous, and use (14.37) to deduce these properties as a conclusion. This sort of reasoning is best carried out in a course on measure theory and integration.

In light of the arguments given above, we see that the following class of functions arises as one that is significant for Fourier analysis.

\[(14.38) \quad \mathcal{A}(\mathbb{R}^n) = \{f \in \mathcal{R}(\mathbb{R}^n) : f \text{ bounded and continuous}, \hat{f} \in \mathcal{R}(\mathbb{R}^n)\}.

By Proposition 14.4, the Fourier inversion formula (14.17) holds for all \(f \in \mathcal{A}(\mathbb{R}^n)\). It also follows that \(f \in \mathcal{A}(\mathbb{R}^n) \Rightarrow \hat{f} \in \mathcal{A}(\mathbb{R}^n)\).

It is of interest to know when \(f \in \mathcal{A}(\mathbb{R}^n)\). We begin with the following simple result. Suppose \(f \in C^k(\mathbb{R}^n)\) has compact support (we write \(f \in C^k_c(\mathbb{R}^n)\)). Then integration by parts yields

\[(14.39) \quad (2\pi)^{-n/2} \int_{\mathbb{R}^n} f^{(\alpha)}(x)e^{-ix \cdot \xi} \, dx = (i\xi)^{\alpha} \hat{f}(\xi), \quad |\alpha| \leq k.\]
Hence

\begin{equation}
 f \in C^k_c(\mathbb{R}^n) \implies |\hat{f}(\xi)| \leq C(1 + |\xi|)^{-k} \implies \hat{f} \in \mathcal{R}(\mathbb{R}^n), \text{ if } k > n,
\end{equation}

so

\begin{equation}
 C^{{n+1}}_c(\mathbb{R}^n) \subset A(\mathbb{R}^n).
\end{equation}

We will obtain some substantially sharper results below.

To proceed, it is useful to bring in the quantities $\|f\|_{L^2}$ and $(f,g)_{L^2}$, defined by

\begin{equation}
 \|f\|_{L^2}^2 = \int_{\mathbb{R}^n} |f(x)|^2 \, dx,
\end{equation}

and

\begin{equation}
 (f,g)_{L^2} = \int_{\mathbb{R}^n} f(x)g(x) \, dx.
\end{equation}

Note that $\|f\|_{L^2}^2 = (f,f)_{L^2}$. Since elements of $\mathcal{R}(\mathbb{R}^n)$ are both bounded and integrable, we have

\begin{equation}
 f \in \mathcal{R}(\mathbb{R}^n) \implies \int_{\mathbb{R}^n} |f(x)|^2 \, dx \leq \left( \sup_{\mathbb{R}^n} |f| \right) \|f\|_{L^1},
\end{equation}

where $\|f\|_{L^1}$ is defined by (14.4). Use of (14.43) makes $\mathcal{R}(\mathbb{R}^n)$ an inner product space and use of (14.42) makes it a normed space (if we identify $f$ and $g$ whenever $\int |f-g| \, dx = 0$). In particular, the triangle inequality holds:

\begin{equation}
 \|f + g\|_{L^2} \leq \|f\|_{L^2} + \|g\|_{L^2}.
\end{equation}

A parallel result holds for $L^1$:

\begin{equation}
 \|f + g\|_{L^1} \leq \|f\|_{L^1} + \|g\|_{L^1}.
\end{equation}

The inequality (14.46) is immediate from the definition (14.4) and the pointwise estimate $|f(x) + g(x)| \leq |f(x)| + |g(x)|$. The proof of (14.45) takes a longer argument, involving along the way Cauchy’s inequality,

\begin{equation}
 |(f,g)_{L^2}| \leq \|f\|_{L^2} \|g\|_{L^2}.
\end{equation}

See Appendix H, on inner product spaces, for a derivation of (14.47) and (14.45).
An important property of the Fourier transform $\mathcal{F}$ and of $\mathcal{F}^*$ is that they preserve the $L^2$-norm and inner product. We first derive this for elements of $A(\mathbb{R}^n)$. Since $\mathcal{F}$ and $\mathcal{F}^*$ differ only in replacing $e^{-ix\cdot \xi}$ by its complex conjugate $e^{ix\cdot \xi}$, we have, for $f, g \in A(\mathbb{R}^n)$,

\begin{equation}
(\mathcal{F} f, g)_{L^2} = (f, \mathcal{F}^* g)_{L^2}.
\end{equation}

Combining this with Proposition 14.4, we have

\begin{equation}
f, g \in A(\mathbb{R}^n) \implies (\mathcal{F} f, \mathcal{F} g)_{L^2} = (f, \mathcal{F}^* \mathcal{F} g)_{L^2} = (f, g)_{L^2}.
\end{equation}

One readily obtains a similar result with $\mathcal{F}$ replaced by $\mathcal{F}^*$. Hence

\begin{equation}
\|\mathcal{F} f\|_{L^2} = \|\mathcal{F}^* f\|_{L^2} = \|f\|_{L^2},
\end{equation}

for all $f \in A(\mathbb{R}^n)$.

The result (14.50) is called the Plancherel identity. It extends beyond $f \in A(\mathbb{R}^n)$. We aim to prove the following.

**Proposition 14.5.** If $f \in \mathcal{R}(\mathbb{R})$, then $|\mathcal{F} f|^2$ and $|\mathcal{F}^* f|^2$ belong to $\mathcal{R}(\mathbb{R})$, and (14.50) holds.

Note that we do not assert that $f \in \mathcal{R}(\mathbb{R})$ implies $\hat{f} \in \mathcal{R}(\mathbb{R})$. Indeed, that can fail, as the following simple example shows, in case $n = 1$. Namely, if

\begin{equation}
\chi_R(x) = 1 \text{ for } |x| \leq R, \quad 0 \text{ for } |x| > R,
\end{equation}

then

\begin{equation}
\hat{\chi}_R(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{-ix\xi} dx
= \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} \cos x\xi dx
= \frac{\sqrt{2}}{\pi} \sin R\xi
= \sqrt{\frac{2}{\pi}} \sin \frac{R\xi}{\xi},
\end{equation}

a function that is square integrable but not integrable.

We use an approximation argument to prove Proposition 14.5, making use of the following lemma. Pick $k > n$.

**Lemma 14.6.** Given $f \in \mathcal{R}(\mathbb{R})$, there exist $f_\nu \in C^k_c(\mathbb{R})$ such that $\sup |f_\nu| \leq \sup |f|$,

\begin{equation}
\|f - f_\nu\|_{L^1} \to 0, \quad \text{and} \quad \|f - f_\nu\|_{L^2} \to 0.
\end{equation}

**Proof.** Take $f_R(x) = f(x)$ for $|x| < R$, 0 for $|x| \geq R$. Then $f \in \mathcal{R}(\mathbb{R}) \to \int |f - f_R| dx \to 0$ as $R \to \infty$, so we specialize to the case $f \in \mathcal{R}_c(\mathbb{R})$. It suffices to treat the case $f \geq 0$. }
Then the existence of $f_\nu \in C_c(\mathbb{R}^n)$ such that $0 \leq f_\nu \leq f$ and $\|f - f_\nu\|_{L^1} \to 0$ follows from Proposition 4.11. Then also $\|f - f_\nu\|_{L^2} \to 0$. Further approximation by elements of $C^k_c(\mathbb{R}^n)$ is left to the reader. One might use the Stone-Weierstrass theorem, or perhaps a convolution argument.

We now move to the proof of Proposition 14.5. By (14.41), each $f_\nu$ belongs to $A(\mathbb{R}^n)$, so, by (14.50),

\[(14.54) \quad \|\hat{f}_\nu\| = \|f_\nu\|_{L^2}, \quad \forall \nu.\]

By (14.3) we have

\[(14.55) \quad \sup_{\xi} |\hat{f}(\xi) - \hat{f}_\nu(\xi)| \leq (2\pi)^{-n/2}\|f - f_\nu\|_{L^1},\]

so $\hat{f}_\nu \to \hat{f}$ uniformly on $\mathbb{R}^n$. Consequently, for each $R < \infty$,

\[(14.56) \quad \int_{|\xi| \leq R} |\hat{f}(\xi)|^2 d\xi = \lim_{\nu \to \infty} \int_{|\xi| \leq R} |\hat{f}_\nu(\xi)|^2 d\xi.\]

Meanwhile, the right side of (14.56) is dominated by $\|\hat{f}_\nu\|_{L^2}^2$, so

\[(14.57) \quad \int_{|\xi| \leq R} |\hat{f}(\xi)|^2 d\xi \leq \liminf_{\nu \to \infty} \|\hat{f}_\nu\|_{L^2}^2 = \liminf_{\nu \to \infty} \|f_\nu\|_{L^2}^2 = \|f\|_{L^2}^2,\]

the second line by (14.54) and the third by (14.53). Since (14.57) holds for all $R < \infty$, we have at this point that

\[(14.58) \quad f \in \mathcal{R}(\mathbb{R}^n) \implies \|\hat{f}\|_{L^2} \leq \|f\|_{L^2}.\]

To proceed, we apply this implication to $g_\nu = f - f_\nu$, obtaining $\|\hat{g}_\nu\|_{L^2} \leq \|g_\nu\|_{L^2}$, i.e.,

\[(14.59) \quad \|\hat{f} - \hat{f}_\nu\|_{L^2} \leq \|f - f_\nu\|_{L^2}.\]

By (14.53), $\|f - f_\nu\|_{L^2} \to 0$ as $\nu \to \infty$, so

\[(14.60) \quad \|\hat{f} - \hat{f}_\nu\|_{L^2} \to 0.\]

Hence

\[(14.61) \quad \|\hat{f}\|_{L^2} = \lim_{\nu \to \infty} \|\hat{f}_\nu\|_{L^2}.\]
Applying (14.54) and (14.53) again, we get

\begin{equation}
(14.62) \quad f \in \mathcal{R}(\mathbb{R}^n) \implies \|\hat{f}\|_{L^2} = \|f\|_{L^2},
\end{equation}

as desired. The same sort of argument applies to \( F^* f \). Proposition 14.5 is proved.

This \( L^2 \) material sets us up for another type of Fourier inversion formula. To state it, we bring in the following family of linear transformations. For \( R \in (0, \infty) \) and \( f \in \mathcal{R}(\mathbb{R}^n) \), set

\begin{equation}
(14.63) \quad S_R f(x) = (2\pi)^{-n/2} \int_{|\xi| \leq R} \hat{f}(\xi) e^{i x \cdot \xi} d\xi.
\end{equation}

Equivalently,

\begin{equation}
(14.64) \quad S_R f = F^*(\chi_R \hat{f}),
\end{equation}

with \( \chi_R \) as in (14.51), but this time \( x \in \mathbb{R}^n \). Note that

\begin{equation}
(14.65) \quad f \in \mathcal{R}(\mathbb{R}^n) \implies \hat{f} \text{ bounded and continuous (and square integrable)}
\implies \chi_R f \in \mathcal{R}(\mathbb{R}^n)
\implies F^*(\chi_R \hat{f}) \text{ bounded and continuous, and square integrable}.
\end{equation}

We also have

\begin{equation}
(14.66) \quad \|\chi_R \hat{f} - \hat{f}\|_{L^2} \to 0 \quad \text{as} \quad R \to \infty.
\end{equation}

This suggests the following result, our second Fourier inversion formula.

**Proposition 14.7.** Given \( f \in \mathcal{R}(\mathbb{R}^n) \),

\begin{equation}
(14.67) \quad \|S_R f - f\|_{L^2} \to 0 \quad \text{as} \quad R \to \infty.
\end{equation}

It is convenient to approach this via a lemma.

**Lemma 14.8.** If \( f \in \mathcal{A}(\mathbb{R}^n) \), then (14.67) holds.

**Proof.** We already know that \( \chi_R \hat{f} \in \mathcal{R}(\mathbb{R}^n) \). If \( f \in \mathcal{A}(\mathbb{R}^n) \), then also \( \hat{f} \in \mathcal{R}(\mathbb{R}^n) \), so Proposition 14.5 applies to \( \chi_R \hat{f} - \hat{f} \), yielding

\begin{equation}
(14.68) \quad \|S_R f - f\|_{L^2} = \|F^*(\chi_R \hat{f} - \hat{f})\|_{L^2} = \|\chi_R \hat{f} - \hat{f}\|_{L^2} \to 0.
\end{equation}

To prove Proposition 14.7, we use Lemma 14.6 to obtain \( f_\nu \in \mathcal{A}(\mathbb{R}^n) \) such that \( \|f - f_\nu\|_{L^2} \to 0 \). Note also that, by (14.65),

\begin{equation}
(14.69) \quad f \in \mathcal{R}(\mathbb{R}^n) \implies \|S_R f\|_{L^2} = \|\chi_R \hat{f}\|_{L^2} \leq \|\hat{f}\|_{L^2} = \|f\|_{L^2},
\end{equation}

as desired.
for all $R \in (0, \infty)$. Now we can write

\begin{equation}
S_R f - f = (S_R f - S_R f_\nu) + (S_R f_\nu - f_\nu) + (f_\nu - f),
\end{equation}

and hence

\begin{equation}
\|S_R f - f\|_{L^2} \leq \|S_R (f - f_\nu)\|_{L^2} + \|S_R f_\nu - f_\nu\|_{L^2} + \|f - f_\nu\|_{L^2}
\leq \|S_R f_\nu - f_\nu\|_{L^2} + 2\|f - f_\nu\|_{L^2},
\end{equation}

the last inequality by (14.69), with $f$ replaced by $f - f_\nu$. Consequently, by Lemma 14.8, for each $f \in \mathcal{R}(\mathbb{R}^n)$,

\begin{equation}
\limsup_{R \to \infty} \|S_R f - f\|_{L^2} \leq 2\|f - f_\nu\|_{L^2}, \quad \forall \nu,
\end{equation}

and taking $\nu \to \infty$ gives (14.67). Proposition 14.7 is proved.

We next use the Plancherel theorem to establish the following sharpening of (14.41). See Propositions 14.17–14.18 for a further improvement.

**Proposition 14.9.** We have

\begin{equation}
k > \frac{n}{2} \implies C^k_c(\mathbb{R}^n) \subset \mathcal{A}(\mathbb{R}^n).
\end{equation}

**Proof.** Given $f \in C^k_c(\mathbb{R}^n)$, we again use the identity (14.39). The Plancherel identity then gives

\begin{equation}
\int_{\mathbb{R}^n} |\xi^\alpha \hat{f}(\xi)|^2 d\xi = \|f^{(\alpha)}\|_{L^2}^2, \quad \forall |\alpha| \leq k.
\end{equation}

Summing over $|\alpha| \leq k$ gives

\begin{equation}
\int_{\mathbb{R}^n} (1 + |\xi|)^{2k} |\hat{f}(\xi)|^2 d\xi \leq C \sum_{|\alpha| \leq k} \|f^{(\alpha)}\|_{L^2}^2.
\end{equation}

Now Cauchy’s inequality gives

\begin{align*}
\int_{\mathbb{R}^n} |\hat{f}(\xi)| d\xi &= \int_{\mathbb{R}^n} (1 + |\xi|)^{-k} |\hat{f}(\xi)| |(1 + |\xi|)^k d\xi| \\
&\leq A_{n,k}^{1/2} \left\{ \int_{\mathbb{R}^n} (1 + |\xi|)^{2k} |\hat{f}(\xi)|^2 d\xi \right\}^{1/2},
\end{align*}

\begin{equation}
(14.76)
\end{equation}
where

\[(14.77) \quad A_{n,k} = \int_{\mathbb{R}^n} (1 + |\xi|)^{-2k} d\xi < \infty, \quad \text{if } 2k > n.\]

This establishes the desired finiteness of \(\int_{\mathbb{R}^n} |\hat{f}(\xi)| d\xi\).

Having seen important roles played by \(L^1\) and \(L^2\) norms and \(L^2\) inner products, we are motivated to advertise how Fourier analysis is a natural setting in which to work with spaces of functions larger than \(\mathcal{R}(\mathbb{R}^n)\) or \(\mathcal{R}^\#(\mathbb{R}^n)\), spaces that are labeled \(L^1(\mathbb{R}^n)\) and \(L^2(\mathbb{R}^n)\). These are defined using the Lebesgue theory of integration. We give a brief description of this, referring the reader to other sources, such as [Fol] or [T2], for a detailed presentation.

To start, we say a set \(S \subset \mathbb{R}^n\) is measurable provided that, for each cell \(R \subset \mathbb{R}^n\),

\[(14.78) \quad m^*(S \cap R) + m^*(R \setminus S) = V(R),\]

where \(m^*\) is the outer measure, defined in (4.118). We say a function \(f : \mathbb{R}^n \to \mathbb{C}\) is measurable provided that, for each open \(\mathcal{O} \subset \mathbb{C}\), \(f^{-1}(\mathcal{O}) \subset \mathbb{R}^n\) is measurable. The Lebesgue integral associates a value in \([0, +\infty]\) to

\[(14.79) \quad \int_{\mathbb{R}^n} f(x) \, dx\]

for each measurable \(f\) satisfying \(f(x) \geq 0\) for all \(x\). The space \(L^1(\mathbb{R}^n)\) consists of all measurable functions \(f\) such that

\[(14.80) \quad \|f\|_{L^1} = \int_{\mathbb{R}^n} |f(x)| \, dx < \infty.\]

In such a case, one can write \(f = f_{0^+} - f_{0^-} + i(f_{1^+} - f_{1^-})\) with all \(f_{j\pm}\) measurable and \(\geq 0\), and with finite integral, and the process alluded to above applies to evaluate these integrals, and hence to evaluate \(\int_{\mathbb{R}^n} f(x) \, dx\). There is one further wrinkle. The space \(L^1(\mathbb{R}^n)\) actually consists of equivalence classes of measurable functions satisfying (14.80), where the equivalence is \(f_1 \sim f_2\) if and only if \(\{x \in \mathbb{R}^n : f_1(x) \neq f_2(x)\}\) has outer measure 0. This makes \(L^1(\mathbb{R}^n)\) a normed space.

More generally, for \(p \in [1, \infty)\), \(L^p(\mathbb{R}^n)\) consists of equivalence classes of measurable functions for which

\[(14.81) \quad \|f\|_{L^p} = \left( \int_{\mathbb{R}^n} |f(x)|^p \, dx \right)^{1/p} < \infty.\]

The only case of \(p > 1\) that we work with here is \(p = 2\). One has

\[(14.82) \quad f, g \in L^2(\mathbb{R}^n) \implies fg \in L^1(\mathbb{R}^n),\]
and \( L^2(\mathbb{R}^n) \) is an inner product space, via (14.43), extended to this setting. These \( L^p \) norms satisfy the triangle inequality:

\[
\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.
\]

For \( p = 1 \) or \( 2 \), the proofs of (14.83) are as described before. For other \( p \in (1, \infty) \), which we do not deal with here, the reader can consult the references mentioned above. Thanks to (14.83), \( L^p(\mathbb{R}^n) \) has the structure of a metric space, with \( d(f, g) = \|f - g\|_{L^p} \). We say \( f_\nu \to f \) in \( L^p \) if \( \|f_\nu - f\|_{L^p} \to 0 \). For all \( p \in [1, \infty) \), these spaces have the following important metric properties. The first is a denseness property.

**Proposition A.** Given \( f \in L^p(\mathbb{R}^n) \) and \( k \in \mathbb{N} \), there exist \( f_\nu \in C_c^k(\mathbb{R}^n) \) such that \( f_\nu \to f \) in \( L^p \).

The next is a completeness property.

**Proposition B.** If \( (f_\nu) \) is a Cauchy sequence in \( L^p(\mathbb{R}^n) \), then there exists \( f \in L^p(\mathbb{R}^n) \) such that \( f_\nu \to f \) in \( L^p \).

We refer to \([\text{Fol}]\) or \([\text{T2}]\) for proofs of these results.

We mention that if \( f \) is bounded and continuous on \( \mathbb{R}^n \), then \( f \in L^1(\mathbb{R}^n) \) if and only if \( f \in \mathcal{R}(\mathbb{R}^n) \). Hence (14.38) is equivalent to

\[
\mathcal{A}(\mathbb{R}^n) = \{f \in L^1(\mathbb{R}^n) : f \text{ is bounded and continuous, and } \hat{f} \in L^1(\mathbb{R}^n)\}.
\]

We also have

\[
\mathcal{A}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n),
\]

either by (13.44) or by

\[
\|f\|_{L^2} \leq \left(\sup |f|\right)\|f\|_{L^1} \leq (2\pi)^{-n/2}\|\hat{f}\|_{L^1}\|f\|_{L^1}.
\]

The following neat extensions of Propositions 14.5 and 14.7 illustrate the usefulness of the Lebesgue theory of integration in Fourier analysis.

**Proposition 14.10.** The maps \( \mathcal{F} \) and \( \mathcal{F}^* \) have unique continuous linear extensions from

\[
\mathcal{F}, \mathcal{F}^* : \mathcal{A}(\mathbb{R}^n) \longrightarrow \mathcal{A}(\mathbb{R}^n)
\]

to

\[
\mathcal{F}, \mathcal{F}^* : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n),
\]

and the identities

\[
\mathcal{F}^* \mathcal{F} f = f, \quad \mathcal{F} \mathcal{F}^* f = f,
\]

and

\[
\|\mathcal{F} f\|_{L^2} = \|\mathcal{F}^* f\|_{L^2} = \|f\|_{L^2}
\]

hold for all \( f \in L^2(\mathbb{R}^n) \).
Proposition 14.11. Define $S_R$ by

\begin{equation}
S_R f(x) = (2\pi)^{-n/2} \int_{|\xi| \leq R} \hat{f}(\xi) e^{ix \cdot \xi} \, d\xi.
\end{equation}

Then $S_R : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, and

\begin{equation}
f \in L^2(\mathbb{R}^n) \implies S_R f \to f \quad \text{in} \quad L^2(\mathbb{R}^n),
\end{equation}

as $R \to \infty$.

Proposition 14.10 can be proven using Propositions A and B (with $p = 2$) and the inclusion (14.41), which, together with Proposition A, implies that

\begin{equation}
given f \in L^2(\mathbb{R}^n), \text{ there exist } f_\nu \in \mathcal{A}(\mathbb{R}^n) \text{ such that } f_\nu \to f \text{ in } L^2.
\end{equation}

The argument goes like this. Given $f \in L^2(\mathbb{R}^n)$, take $f_\nu$ as in (14.93). Then $\|f_\mu - f_\nu\|_{L^2} \to 0$ as $\mu, \nu \to \infty$. Now (14.50), applied to $f_\mu - f_\nu \in \mathcal{A}(\mathbb{R}^n)$, gives

\begin{equation}
\|\mathcal{F}f_\mu - \mathcal{F}f_\nu\|_{L^2} = \|f_\mu - f_\nu\|_{L^2} \to 0,
\end{equation}

as $\mu, \nu \to \infty$. Hence $(\mathcal{F}f_\mu)$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. By Proposition B, there exits a limit $h \in L^2(\mathbb{R}^n)$, that is, $\mathcal{F}f_\nu \to h$ in $L^2$. One gets the same element $h$ regardless of the choice of $(f_\nu)$ such that (14.93) holds, and so we set $\mathcal{F}f = h$. The same argument applies to $\mathcal{F}^* f_\nu$, which hence converges to $\mathcal{F}^* f$. We have

\begin{equation}
\|\mathcal{F}f_\nu - \mathcal{F}f\|_{L^2}, \quad \|\mathcal{F}^* f_\nu - \mathcal{F}^* f\|_{L^2} \to 0.
\end{equation}

From here, the results (14.89)–(14.90) follow.

As for Proposition 14.11, we have

\begin{equation}
S_R f = \mathcal{F}^*(\chi_R \mathcal{F}f),
\end{equation}

and, thanks to Proposition 14.10,

\begin{equation}
f \in L^2(\mathbb{R}^n) \implies \mathcal{F}f \in L^2(\mathbb{R}^n)
\implies \chi_R \mathcal{F}f \to \mathcal{F}f \quad \text{in} \quad L^2(\mathbb{R}^n)
\implies \mathcal{F}^*(\chi_R \mathcal{F}f) \to \mathcal{F}^* \mathcal{F}f \quad \text{in} \quad L^2(\mathbb{R}^n),
\end{equation}

and finally, again by Proposition 14.10, $\mathcal{F}^* \mathcal{F}f = f$.

We now discuss another way of taking Fourier analysis beyond the setting of $\mathcal{A}(\mathbb{R}^n)$, which is a normed linear space, with norm

\begin{equation}
\|f\|_\mathcal{A} = \|f\|_{L^1} + \|\hat{f}\|_{L^1},
\end{equation}
for which we have

\[(14.99) \quad \|\mathcal{F}f\|_A = \|\mathcal{F}^*f\|_A = \|f\|_A,\]

for all \(f \in \mathcal{A}(\mathbb{R}^n)\), thanks to Proposition 14.4 and the fact that \(\mathcal{F}^*f(\xi) = \mathcal{F}f(-\xi)\). We define the dual space \(\mathcal{A}'(\mathbb{R}^n)\) to consist of continuous linear functionals

\[(14.100) \quad w : \mathcal{A}(\mathbb{R}^n) \rightarrow \mathbb{C},\]

i.e., those linear maps \(w\) satisfying

\[(14.101) \quad |w(f)| \leq C\|f\|_A, \quad \forall f \in \mathcal{A}(\mathbb{R}^n).\]

The optimal constant in (14.101) is denoted \(\|w\|_{\mathcal{A}'}\). We also use the notation

\[(14.102) \quad \langle f, w \rangle = w(f), \quad f \in \mathcal{A}(\mathbb{R}^n), \quad w \in \mathcal{A}'(\mathbb{R}^n).\]

Then, we define

\[(14.103) \quad \mathcal{F}, \mathcal{F}^* : \mathcal{A}'(\mathbb{R}^n) \rightarrow \mathcal{A}'(\mathbb{R}^n)\]

by

\[(14.103) \quad \langle f, \mathcal{F}w \rangle = \langle \mathcal{F}f, w \rangle, \quad \langle f, \mathcal{F}^*w \rangle = \langle \mathcal{F}^*f, w \rangle.\]

Note that

\[(14.104) \quad |\langle f, \mathcal{F}w \rangle| = |\langle \mathcal{F}f, w \rangle| \leq \|w\|_{\mathcal{A}'} \|\mathcal{F}f\|_A = \|w\|_{\mathcal{A}'} \|f\|_A,\]

so

\[(14.105) \quad \|\mathcal{F}w\|_{\mathcal{A}'} \leq \|w\|_{\mathcal{A}'}.\]

We also have the Fourier inversion formula on \(\mathcal{A}'(\mathbb{R}^n)\):

\[(14.106) \quad \mathcal{F}^* \mathcal{F}w = \mathcal{F} \mathcal{F}^*w = w, \quad \forall w \in \mathcal{A}'(\mathbb{R}^n).\]

Indeed, given \(f \in \mathcal{A}(\mathbb{R}^n), \quad w \in \mathcal{A}'(\mathbb{R}^n),\)

\[(14.107) \quad \langle f, \mathcal{F}^*w \rangle = \langle \mathcal{F}^*f, w \rangle = \langle \mathcal{F}^*f, w \rangle = \langle f, w \rangle,\]

the last identity by Proposition 14.4. The proof that \(\mathcal{F}^* \mathcal{F}w = w\) is similar. Incidentally, this Fourier inversion formula combines with (14.105) to yield

\[(14.107A) \quad \|\mathcal{F}w\|_{\mathcal{A}'} = \|w\|_{\mathcal{A}'}.\]
The space $A'(\mathbb{R}^n)$ contains an interesting variety of functions and other objects. For example, there is a natural map

$$\iota : BC(\mathbb{R}^n) \rightarrow A'(\mathbb{R}^n),$$

where $BC(\mathbb{R}^n)$ consists of the set of bounded continuous functions on $\mathbb{R}^n$, with norm

$$||w||_{BC} = \sup |w|.$$ 

The map is given by

$$\langle f, \iota(w) \rangle = \int_{\mathbb{R}^n} f(x)w(x) \, dx.$$ 

Clearly

$$|\langle f, \iota(w) \rangle| \leq (\sup |w|) ||f||_{L^1}.$$ 

We also claim that the map $\iota$ in (14.108) is injective, i.e., $w \in BC(\mathbb{R}^n), \langle f, \iota(w) \rangle = 0$ for all $f \in A(\mathbb{R}^n)$ implies $w = 0$. This is a consequence of the following more general result, whose proof we leave to the reader.

**Lemma 14.12.** If $w \in BC(\mathbb{R}^n)$ and $\int f(x)w(x) \, dx = 0$ for all $f \in C^\infty_c(\mathbb{R}^n)$, then $w \equiv 0$.

The space $A'(\mathbb{R}^n)$ also contains some objects more singular than functions. For example, given $p \in \mathbb{R}^n$, we define $\delta_p \in A'(\mathbb{R}^n)$ by

$$\langle f, \delta_p \rangle = f(p).$$

For $p = 0$, we simply set $\delta = \delta_0$.

Let us compute some Fourier transforms, and observe the Fourier inversion formula in action. We have

$$\langle f, F\delta \rangle = \langle Ff, \delta \rangle = \hat{f}(0) = (2\pi)^{-n/2} \int f(x) \, dx,$$

so

$$F\delta(\xi) = (2\pi)^{-n/2},$$

a constant function. By comparison,

$$\langle f, F^*1 \rangle = \langle F^*f, 1 \rangle = \int F^* f(\xi) \, d\xi$$

$$= (2\pi)^{n/2} F F^* f(0)$$

$$= (2\pi)^{n/2} f(0),$$
the last identity by Proposition 14.4. Hence

\[(14.116)\quad \mathcal{F}^* 1 = (2\pi)^{n/2} \delta.\]

Thus (14.114) and (14.116) illustrate (14.106).

Generalizing the construction of \(\delta_p\), let \(M \subset \mathbb{R}^n\) be a compact, \(m\)-dimensional, \(C^1\) surface, and \(u \in C(M)\). Define \(u\delta_M \in \mathcal{A}'(\mathbb{R}^n)\) by

\[(14.117)\quad \langle f, u\delta_M \rangle = \int_M f(x)u(x) \, dS(x).\]

Then

\[(14.118)\quad \langle f, \mathcal{F}(u\delta_M) \rangle = \langle \mathcal{F} f, u\delta_M \rangle = \int_M \hat{f}(\xi)u(\xi) \, dS(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \left( \int_M f(x)e^{-ix\cdot \xi} \, dx \right) u(\xi) \, dS(\xi).\]

Now covering \(M\) with coordinate charts and chopping \(u\) into pieces supported on coordinate patches, we can repeatedly apply the Fubini theorem to write

\[(14.119)\quad \langle f, \mathcal{F}(u\delta_M) \rangle = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \left( \int_M u(\xi)e^{-ix\cdot \xi} \, dS(\xi) \right) f(x) \, dx,\]

so \(\mathcal{F}(u\delta_M)\) is (the image under \(\iota\) in (14.108) of) an element of \(BC(\mathbb{R}^n)\), given by

\[(14.120)\quad \mathcal{F}(u\delta_M)(x) = (2\pi)^{-n/2} \int_M u(\xi)e^{-ix\cdot \xi} \, dS(\xi).\]

In particular,

\[(14.121)\quad \mathcal{F}\delta_M(x) = (2\pi)^{-n/2} \int_M e^{-ix\cdot \xi} \, dS(\xi).\]

**Tempered distributions**

Objects such as \(\delta_p\) and \(\delta_M\) are examples of “distributions.” A beautiful theory of Fourier analysis on the space of “tempered distributions,” which includes some more singular objects, was constructed by L. Schwartz. We present some of his results here. We start with the space \(S(\mathbb{R}^n)\), defined as

\[(14.122)\quad S(\mathbb{R}^n) = \{ f \in C^\infty(\mathbb{R}^n) : x^\beta f^{(\alpha)} \text{ bounded}, \forall \alpha, \beta \}.\]
By this definition, we clearly have

\[ f \in \mathcal{S}(\mathbb{R}^n) \implies x^\beta f^{(\alpha)} \in \mathcal{S}(\mathbb{R}^n), \quad \forall \alpha, \beta. \]

Using the integrability of \((1 + |x|)^{-(n+1)}\) on \(\mathbb{R}^n\), one can show that

\[ f \in \mathcal{S}(\mathbb{R}^n) \implies x^\beta f^{(\alpha)} \in \mathcal{R}(\mathbb{R}^n), \quad \forall \alpha, \beta. \tag{14.123} \]

Also, the integration by parts argument in (14.39) extends:

\[ f \in \mathcal{S}(\mathbb{R}^n) \implies (2\pi)^{-n/2} \int_{\mathbb{R}^n} f^{(\alpha)}(x)e^{-ix\cdot\xi} \, dx = (i\xi) \hat{f}(\xi), \quad \forall \alpha. \tag{14.124} \]

In particular, \(f \in \mathcal{S}(\mathbb{R}^n) \implies \hat{f} \in \mathcal{R}(\mathbb{R}^n)\), so

\[ \mathcal{S}(\mathbb{R}^n) \subset \mathcal{A}(\mathbb{R}^n). \tag{14.125} \]

Thus, by Proposition 14.4,

\[ f \in \mathcal{S}(\mathbb{R}^n) \implies f = \mathcal{F}^* \mathcal{F}f = \mathcal{F}^\ast \mathcal{F}f. \tag{14.126} \]

We can also complement (14.124) by

\[ f \in \mathcal{S}(\mathbb{R}^n) \implies (2\pi)^{-n/2} \int_{\mathbb{R}^n} x^\beta f(x)e^{-ix\cdot\xi} \, dx = i|\beta| \partial_\xi^\beta \hat{f}(\xi), \tag{14.127} \]

and deduce that

\[ \mathcal{F}, \mathcal{F}^* : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n). \tag{14.128} \]

By (14.126), the operators \(\mathcal{F}\) and \(\mathcal{F}^\ast\) are inverses of each other on \(\mathcal{S}(\mathbb{R}^n)\).

The space \(\mathcal{S}(\mathbb{R}^n)\) carries the following sequence of norms:

\[ p_k(f) = \max_{|\alpha| + |\beta| \leq k} \sup_{x \in \mathbb{R}^n} |x^\beta f^{(\alpha)}(x)|. \tag{14.129} \]

One says a linear map \(T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)\) is continuous provided that, for each \(k \in \mathbb{Z}^+\), there exists \(\ell \in \mathbb{Z}^+\) and \(C_k < \infty\) such that

\[ p_k(Tf) \leq C_k p_\ell(f), \quad \forall f \in \mathcal{S}(\mathbb{R}^n). \tag{14.130} \]

The observations leading to (14.123) and (14.127) show that

\[ p_k(\mathcal{F}f) \leq C_k p_{k+n+1}(f), \quad \forall f \in \mathcal{S}(\mathbb{R}^n), \tag{14.131} \]
so $F$ in (14.128) is continuous. The same goes for $F^*$.

Now a tempered distribution $w \in S'(\mathbb{R}^n)$ is a continuous linear functional
\begin{equation}
(14.132) \quad w : S(\mathbb{R}^n) \rightarrow \mathbb{C},
\end{equation}
that is to say, $w$ is a linear map from $S(\mathbb{R}^n)$ to $\mathbb{C}$ with the property that there exists $k \in \mathbb{Z}^+$ and $C < \infty$ such that
\begin{equation}
(14.133) \quad |w(f)| \leq C p_k(f), \quad \forall f \in S(\mathbb{R}^n).
\end{equation}
As in (14.102), we use the notation
\begin{equation}
(14.134) \quad \langle f, w \rangle = w(f), \quad f \in S(\mathbb{R}^n), \ w \in S'(\mathbb{R}^n).
\end{equation}
Analysis behind (14.125) gives
\begin{equation}
(14.135) \quad \|f\|_A \leq C p_{2n+2}(f),
\end{equation}
so each $w \in A'(\mathbb{R}^n)$ also defines an element of $S'(\mathbb{R}^n)$. Thus $\delta_p$ in (14.112) and $\delta_M$ in (14.117) are examples of tempered distributions. To produce more singular tempered distributions, we can define
\begin{equation}
(14.136) \quad \partial^\alpha : S'(\mathbb{R}^n) \rightarrow S'(\mathbb{R}^n)
\end{equation}
by
\begin{equation}
(14.137) \quad \langle f, \partial^\alpha w \rangle = (-1)^{|\alpha|} \langle f^{(\alpha)}, w \rangle.
\end{equation}
In this way, we get, for example, $\delta' \in S'(\mathbb{R})$. We can also define $x^\beta w$ for $w \in S'(\mathbb{R}^n)$ by
\begin{equation}
(14.138) \quad \langle f, x^\beta w \rangle = \langle x^\beta f, w \rangle.
\end{equation}
We now define
\begin{equation}
(14.139) \quad F, \ F^* : S'(\mathbb{R}^n) \rightarrow S'(\mathbb{R}^n)
\end{equation}
by
\begin{equation}
(14.140) \quad \langle f, Fw \rangle = \langle Ff, w \rangle, \quad \langle f, F^*w \rangle = \langle F^*f, w \rangle,
\end{equation}
for $f \in S(\mathbb{R}^n)$, $w \in S'(\mathbb{R}^n)$. Note that, given $w \in S'(\mathbb{R}^n)$,
\begin{equation}
(14.141) \quad |\langle f, w \rangle| \leq C p_k(f), \quad \forall f \in S(\mathbb{R}^n)
\end{equation}
\Rightarrow |\langle f, Fw \rangle| \leq C p_k(Ff) \leq CC_k p_{k+n+1}(f),
by (14.131), so indeed if $w \in S'(\mathbb{R}^n)$, (14.130) defines $Fw$ as an element of $S'(\mathbb{R}^n)$. The same goes for $F^*w$.

Here is a further extension of the Fourier inversion formula.
Proposition 14.13. Given \( w \in \mathcal{S}'(\mathbb{R}^n) \),

\[
\mathcal{F}^* \mathcal{F} w = \mathcal{F} \mathcal{F}^* w = w. \tag{14.142}
\]

Proof. Parallel to (14.107), we have, for \( f \in \mathcal{S}(\mathbb{R}^n), \ w \in \mathcal{S}'(\mathbb{R}^n) \),

\[
\langle f, \mathcal{F}^* \mathcal{F} w \rangle = \langle \mathcal{F}^* f, \mathcal{F} w \rangle = \langle \mathcal{F} \mathcal{F}^* f, w \rangle = \langle f, w \rangle, \tag{14.143}
\]

the last identity by (14.126). The same goes for \( \langle f, \mathcal{F} \mathcal{F}^* w \rangle \).

We can extend (14.124) and (14.127) from \( \mathcal{S}(\mathbb{R}^n) \) to \( \mathcal{S}'(\mathbb{R}^n) \), using (14.137)–(14.138) and an argument parallel to (14.143), to obtain:

Proposition 14.14. Given \( w \in \mathcal{S}'(\mathbb{R}^n) \),

\[
\mathcal{F}(\partial^\alpha w) = (i\xi)^\alpha \mathcal{F} w, \tag{14.144}
\]

and

\[
\mathcal{F}(x^\beta w) = i^{|eta|} \partial_{\xi}^\beta \mathcal{F} w, \tag{14.145}
\]

with similar formulas involving \( \mathcal{F}^* \).

For example, with \( \delta \) as in (14.112) \((p = 0)\), and \( n = 1 \), we have from (14.114) and (14.144)

\[
\mathcal{F} \delta'(\xi) = (2\pi)^{-1/2} i\xi. \tag{14.146}
\]

Poisson summation formulas

Comparing Fourier transforms of functions on \( \mathbb{R}^n \) with Fourier series of related functions on \( \mathbb{T}^n \) leads to highly nontrivial identities, known as Poisson summation formulas. We derive some of them here.

To start, we take

\[
f \in \mathcal{S}(\mathbb{R}^n), \ \varphi(x) = \sum_{\ell \in \mathbb{Z}^n} f(x + 2\pi \ell). \tag{14.147}
\]

We have

\[
\varphi \in C^\infty(\mathbb{R}^n), \ \varphi(x) = \varphi(x + 2\pi k), \ \forall k \in \mathbb{Z}^n, \tag{14.148}
\]

hence (with slight abuse of notation)

\[
\varphi \in C^\infty(\mathbb{T}^n), \ \mathbb{T}^n = \mathbb{R}^n/(2\pi \mathbb{Z}^n). \tag{14.149}
\]
We next observe that
\[(14.140) \int_{\mathbb{T}^n} \varphi(x)e^{-ik \cdot x} \, dx = \int_{\mathbb{R}^n} f(x)e^{-ik \cdot x} \, dx, \quad \forall k \in \mathbb{Z}^n.\]

Consequently, with \(\hat{\varphi}(k)\) defined as in (13.1) and \(\hat{f}(\xi)\) defined as in (14.1),
\[(14.151) \hat{\varphi}(k) = (2\pi)^{-n/2}\hat{f}(k), \quad \forall k \in \mathbb{Z}^n.\]

Now the Fourier inversion formula, in the form of Proposition 13.1, applies to \(\varphi\):
\[(14.152) \varphi(x) = \sum_{k \in \mathbb{Z}^n} \hat{\varphi}(k)e^{ik \cdot x}, \quad \forall x \in \mathbb{T}^n.\]

Putting this together with (14.147) and (14.151) gives the following general Poisson summation formula.

**Proposition 14.15.** Given \(f \in \mathcal{S}(\mathbb{R}^n)\), we have, for each \(x \in \mathbb{R}^n\),
\[(14.153) \sum_{\ell \in \mathbb{Z}^n} f(x + 2\pi \ell) = (2\pi)^{-n/2} \sum_{k \in \mathbb{Z}^n} \hat{f}(k)e^{ik \cdot x}.\]

In particular,
\[(14.154) \sum_{\ell \in \mathbb{Z}^n} f(2\pi \ell) = (2\pi)^{-n/2} \sum_{k \in \mathbb{Z}^n} \hat{f}(k).\]

We can apply this to
\[(14.155) f(x) = e^{-t|x|^2},\]
and use (14.16) to evaluate \(\hat{f}(k)\). This leads to
\[(14.156) \sum_{\ell \in \mathbb{Z}^n} e^{-4\pi^2 t |\ell|^2} = (4\pi t)^{-n/2} \sum_{k \in \mathbb{Z}^n} e^{-|k|^2/4t}, \quad t > 0.\]

Taking \(\tau = 4\pi t\), we can rewrite this as
\[(14.157) \sum_{\ell \in \mathbb{Z}^n} e^{-\pi \tau |\ell|^2} = \tau^{-n/2} \sum_{k \in \mathbb{Z}^n} e^{-\pi |k|^2/\tau}, \quad \tau > 0.\]

This result is known as the Jacobi inversion formula.

The Riemann functional equation
The Riemann zeta function $\zeta(s)$ is defined by

$$\zeta(s) = \sum_{k=1}^{\infty} k^{-s}, \quad \text{Re } s > 1.$$  

(14.158)

This defines $\zeta(s)$ as a function holomorphic in $\{ s \in \mathbb{C} : \text{Re } s > 1 \}$. See (10.52). Here we establish a formula of Riemann that extends $\zeta(s)$ beyond the half plane $\text{Re } s > 1$.

To start the analysis, we relate $\zeta(s)$ to the function

$$g(t) = \sum_{k=1}^{\infty} e^{-\pi k^2 t}. \quad \text{We have}$$  

(14.159)

$$\int_0^{\infty} g(t) t^{s-1} \, dt = \sum_{k=1}^{\infty} k^{-2s} \int_0^{\infty} e^{-t} t^{s-1} \, dt = \zeta(2s) \pi^{-s} \Gamma(s),$$  

(14.160)

for $\text{Re } s > 1/2$. The Gamma function $\Gamma(s)$ is as in (10.50). This gives rise to further identities, via the $n = 1$ case of the Jacobi inversion formula (14.157), i.e.,

$$\sum_{\ell=-\infty}^{\infty} e^{-\pi \ell^2 t} = \sqrt{\frac{1}{t}} \sum_{k=-\infty}^{\infty} e^{-\pi k^2/t},$$  

(14.161)

which implies

$$g(t) = -\frac{1}{2} + \frac{1}{2} t^{-1/2} + t^{-1/2} g\left(\frac{1}{t}\right).$$  

(14.162)

To use this, we first note from (14.160) that, for $\text{Re } s > 1$,

$$\Gamma\left(\frac{s}{2}\right) \pi^{-s/2} \zeta(s) = \int_0^{\infty} g(t) t^{s/2-1} \, dt$$  

(14.163)

$$= \int_0^{1} g(t) t^{s/2-1} \, dt + \int_1^{\infty} g(t) t^{s/2-1} \, dt.$$  

Into the integral over $[0, 1]$, we substitute the right side of (14.162) for $g(t)$, to obtain

$$\Gamma\left(\frac{s}{2}\right) \pi^{-s/2} \zeta(s) = \int_0^{1} \left(-\frac{1}{2} + \frac{1}{2} t^{-1/2}\right) t^{s/2-1} \, dt$$  

(14.164)

$$+ \int_1^{\infty} g(t) t^{s/2-3/2} \, dt + \int_1^{\infty} g(t) t^{s/2-1} \, dt.$$  

We evaluate the first integral on the right and replace $t$ by $1/t$ in the second integral, to obtain, for $\text{Re } s > 1$,

$$\Gamma\left(\frac{s}{2}\right) \pi^{-s/2} \zeta(s) = \frac{1}{s-1} - \frac{1}{s} + \int_1^{\infty} \left[ t^{s/2} + t^{(1-s)/2} \right] g(t) t^{-1} \, dt.$$  

(14.165)

Note that $g(t) \leq C e^{-\pi t}$ for $t \in [1, \infty)$, so the integral on the right side of (14.165) defines a function holomorphic for all $s \in \mathbb{C}$. As seen in §10, $\Gamma(z)$ is holomorphic on $\mathbb{C} \setminus \{0, -1, -2, -3, \ldots \}$. Further results on the Gamma function include the following.
Lemma 14.16. The function $1/\Gamma(z)$ extends to be holomorphic on all of $\mathbb{C}$, with zeros at \{0, -1, -2, -3, \ldots\}. 

We refer to [T6], §18, for a proof. Given this, we have from (14.165) that $\zeta(s)$ extends to be holomorphic on $\mathbb{C} \setminus \{1\}$. 

The formula (14.165) does more than establish such a holomorphic extension of the zeta function. Note that the right side of (14.165) is invariant under replacing $s$ by $1 - s$. Thus we have the following identity, known as Riemann’s functional equation:

\[
\Gamma\left(\frac{s}{2}\right)\pi^{-s/2}\zeta(s) = \Gamma\left(\frac{1-s}{2}\right)\pi^{-(1-s)/2}\zeta(1-s).
\]

The Riemann zeta function plays a central role in the study of prime numbers. Some basic material on this can be found in §19 of [T6], and a great deal more in [Ed].

**More general sufficient condition for $f \in \mathcal{A}(\mathbb{R}^n)$**

We return to the task of identifying elements of $\mathcal{A}(\mathbb{R}^n)$, and establish a result substantially sharper than Proposition 14.9. We mention that an analogous result holds for Fourier series. The interested reader can investigate this.

To set things up, given $f \in \mathcal{R}(\mathbb{R}^n)$, let

\[
f_h(x) = f(x + h).
\]

Here is our result in the case $n = 1$.

**Proposition 14.17.** If $f \in \mathcal{R}(\mathbb{R})$ and there exists $C < \infty$ such that

\[
\|f - f_h\|_{L^2} \leq C|h|^\alpha, \quad \text{for } |h| \leq 1,
\]

with

\[
\alpha > \frac{1}{2},
\]

then $f \in \mathcal{A}(\mathbb{R})$.

**Proof.** A calculation gives

\[
\hat{f}_h(\xi) = e^{ih\xi} \hat{f}(\xi),
\]

so, by the Plancherel identity,

\[
\|f - f_h\|^2_{L^2} = \int_{-\infty}^{\infty} |1 - e^{ih\xi}|^2 |\hat{f}(\xi)|^2 d\xi.
\]

Now,

\[
\frac{\pi}{2} \leq |h\xi| \leq \frac{3\pi}{2} \implies |1 - e^{ih\xi}|^2 \geq 2,
\]
If (14.168) holds, we deduce that, for $0 < |h| < 1$,
\begin{equation}
(14.173) \quad \int_{|\xi| \leq \frac{4}{|h|}} |\hat{f}(\xi)|^2 d\xi \leq C|h|^{2\alpha},
\end{equation}
hence (setting $|h| = 2^{-\ell+1}$), for $\ell \geq 1$,
\begin{equation}
(14.174) \quad \int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} |\hat{f}(\xi)|^2 d\xi \leq C2^{-2\alpha \ell}.
\end{equation}
Cauchy’s inequality gives
\begin{equation}
(14.175) \quad \int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} \left| |\hat{f}(\xi)| \right| d\xi \\
\leq \left\{ \int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} |\hat{f}(\xi)|^2 d\xi \right\}^{1/2} \times \left\{ \int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} 1 d\xi \right\}^{1/2} \\
\leq C2^{-\alpha \ell} \cdot 2^\ell/2 \\
= C2^{-(\alpha-1/2)\ell}.
\end{equation}
Summing over $\ell \in \mathbb{N}$ and using (again by Cauchy’s inequality)
\begin{equation}
(14.176) \quad \int_{|\xi| \leq 2} |\hat{f}| d\xi \leq C\|\hat{f}\|_{L^2} = C\|f\|_{L^2},
\end{equation}
then gives the proof.

To see how close to sharp Proposition 14.17 is, consider
\begin{equation}
(14.177) \quad f(x) = \chi_I(x) = 1 \text{ if } 0 \leq x \leq 1, \quad 0 \text{ otherwise.}
\end{equation}
We have, for $|h| \leq 1$,
\begin{equation}
(14.178) \quad \|f - f_h\|_{L^2}^2 = 2|h|,
\end{equation}
so (14.168) holds, with $\alpha = 1/2$. Since $A(\mathbb{R}) \subset C(\mathbb{R})$, this function does not belong to $A(\mathbb{R})$, so the condition (14.168A) is about as sharp as it could be.

To produce the appropriate generalization to $n$ variables, let us focus on (14.175), and note that when $\mathbb{R}$ is replaced by $\mathbb{R}^n$, the integral of $1 \, d\xi$ becomes $\sim 2^n \ell$, so to obtain the result

$$
\sum_{\ell \geq 1} \int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} |\hat{f}(\xi)| \, d\xi < \infty,
$$

we want

$$
\int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} |\hat{f}(\xi)|^2 \, d\xi \leq C 2^{-2\gamma \ell}, \quad \gamma > \frac{n}{2}.
$$

It is convenient to rewrite this as

$$
\int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} |\xi|^{2k} |\hat{f}(\xi)|^2 \, d\xi \leq C 2^{-2\alpha \ell},
$$

where

$$
\alpha > 0, \quad \text{if} \quad n = 2k,
$$

$$
\alpha > \frac{1}{2}, \quad \text{if} \quad n = 2k + 1.
$$

Now we bring in Proposition 14.14. Assume

$$
\partial^\beta f = f_\beta \in \mathcal{R}(\mathbb{R}^n), \quad \text{for} \quad |\beta| \leq k,
$$

where a priori $\partial^\beta f$ is defined as an element of $\mathcal{S}'(\mathbb{R}^n)$, by (14.137). Then calculations parallel to (14.169)–(14.174), applied to $f_\beta$ in place of $f$, show that, if

$$
||f_\beta - (f_\beta)_h||_{L^2} \leq C|h|^{\alpha}, \quad \text{for} \quad |\beta| \leq k,
$$

then

$$
\int_{2^\ell \leq |\xi| \leq 2^{\ell+1}} |\xi^\beta \hat{f}(\xi)|^2 \, d\xi \leq C 2^{-\alpha \ell},
$$

for $\ell \geq 1$. Summing over $|\beta| \leq k$ then yields (14.181). We hence have the following higher dimensional extension of Proposition 14.17.
**Proposition 14.18.** Assume $n = 2k$ or $n = 2k + 1$. Take $f \in \mathcal{R}(\mathbb{R}^n)$ and assume that

\begin{equation}
\partial^\beta f = f_\beta \in \mathcal{R}(\mathbb{R}^n), \quad \text{for } |\beta| \leq k,
\end{equation}

and that, for each such $\beta$,

\begin{equation}
\|f_\beta - (f_\beta)_h\|_{L^2} \leq C|h|^\alpha, \quad \text{for } |h| \leq 1,
\end{equation}

where $\alpha$ satisfies (14.182). Then $f \in \mathcal{A}(\mathbb{R}^n)$.

We mention a result that refines Proposition 14.18 when $n > 1$. To state it, we bring in the difference operators

\begin{equation}
\Delta^\beta_{j, \varepsilon} f(x) = \varepsilon^{-1}(f(x + \varepsilon e_j) - f(x)), \quad \Delta^\beta_{\varepsilon} = \Delta^\beta_{1, \varepsilon} \cdots \Delta^\beta_{n, \varepsilon},
\end{equation}

where $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^n$. Here is the result.

**Proposition 14.19.** Assume $n = 2k$ or $n = 2k + 1$. Take $f \in \mathcal{R}(\mathbb{R}^n)$ and assume that there exists $C < \infty$ such that, for $|\beta| \leq k$,

\begin{equation}
\|\Delta^\beta_{\varepsilon} f - (\Delta^\beta_{\varepsilon} f)_h\|_{L^2} \leq C|h|^\alpha, \quad \text{for } |h| \leq 1, \ 0 < \varepsilon \leq 1,
\end{equation}

where $\alpha$ satisfies (14.182). Then $f \in \mathcal{A}(\mathbb{R}^n)$.

We will not present a proof of Proposition 14.19, though the ambitious reader might take it up.

**Exercises**

1. Let $f : \mathbb{R}^n \to \mathbb{C}$ satisfy

\[ |f^{(\alpha)}(x)| \leq C(1 + |x|)^{-(n+1)} \quad \text{for } |\alpha| \leq n + 1. \]

Show that

\[ |\hat{f}(\xi)| \leq C(1 + |\xi|)^{-(n+1)}. \]

Deduce that $f \in \mathcal{A}(\mathbb{R}^n)$.

2. Sharpen the result of Exercise 1 as follows. Assume $f$ satisfies

\[ |f^{(\alpha)}(x)| \leq C(1 + |x|)^{-(n+1)} \quad \text{for } |\alpha| \leq \left\lceil \frac{n}{2} \right\rceil + 1. \]

Then show that $f \in \mathcal{A}(\mathbb{R}^n)$. 
3. Take $n = 1$. For each of the following functions $f : \mathbb{R} \to \mathbb{C}$, compute $\hat{f}(\xi)$.

(a) \( f(x) = e^{-|x|} \),
(b) \( f(x) = \frac{1}{1 + x^2} \),
(c) \( f(x) = \chi_{[-1,1]}(x) \),
(d) \( f(x) = (1 - |x|)\chi_{[-1,1]}(x) \).

4. In each case (a)–(d) of Exercise 3, record the identity that follows from the Plancherel identity (13.50), established in Proposition 13.5.

5. Show that the Poisson summation formula (14.154) applies when

\[ f \in \mathcal{A}(\mathbb{R}^n) \quad \text{and} \quad \varphi \in \mathcal{A}(\mathbb{T}^n), \]

where, as in (14.147), \( \varphi(x) = \sum_\ell f(x + 2\pi \ell) \).

The next exercises bear on the function

\[ \Phi_n(x) = \int_{S^{n-1}} e^{ix \cdot \xi} \, dS(\xi), \quad x \in \mathbb{R}^n. \]

6. Show that

(a) \( \Phi_n \in C^\infty(\mathbb{R}^n) \).
(b) \( (\Delta + 1)\Phi_n = 0 \) on \( \mathbb{R}^n \).
(c) \( \Phi_n \) is radial, i.e., \( \Phi_n(x) = \varphi_n(|x|) \).

7. Deduce from Exercise 6 and the formula for \( \Delta \) in spherical polar coordinates that

\[ \varphi_n''(s) + \frac{n-1}{s} \varphi_n'(s) + \varphi_n(s) = 0. \]

Note that \( \varphi_n(s) = \Phi_n(se_1) \) is a smooth, even function of \( s \).

8. Convert the calculation

\[ \varphi_n(s) = \int_{S^{n-1}} e^{is \xi_1} \, dS(\xi) \]

\[ = \sum_{k=0}^{\infty} \frac{(is)^k}{k!} \int_{S^{n-1}} \xi_1^k \, dS(\xi) \]

\[ = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell)!} s^{2\ell} \int_{S^{n-1}} \xi_1^{2\ell} \, dS(\xi) \]
into a formula for
\[ \int_{S^{n-1}} \xi_1^{2\ell} \, dS(\xi), \quad \ell \in \mathbb{N}. \]

9. More generally, produce a formula for
\[ \int_{S^{n-1}} \xi^\alpha dS(\xi), \quad \alpha = (\alpha_1, \ldots, \alpha_n), \quad \alpha_j \in \mathbb{Z}^+, \]
in terms of \( \Phi_n^{(\alpha)}(0) \).

10. In the spirit of Exercises 8–9, use
\[ e^{-|x|^2/4} = \pi^{-n/2} \int_{\mathbb{R}^n} e^{-|\xi|^2} e^{i x \cdot \xi} \, d\xi \]
to produce a formula for
\[ \int_{\mathbb{R}^n} \xi^\alpha e^{-|\xi|^2} \, d\xi \]
in terms of derivatives of \( e^{-|x|^2/4} \) at \( x = 0 \).

11. Show that \(|x|^{2-n}\) defines an element of \( S'(\mathbb{R}^n) \), and
\[ \Delta(|x|^{2-n}) = C_n \delta \quad \text{on} \quad \mathbb{R}^n, \quad C_n = -(n-2)A_{n-1}, \]
for \( n \geq 3 \). Similarly, show that
\[ \Delta(\log |x|) = 2\pi \delta \quad \text{on} \quad \mathbb{R}^2. \]

\textit{Hint.} Check Exercises 13–14 of §9.

12. Define \( f_r \in C(\mathbb{R}) \) by
\[ f_r(x) = (1 - x^2)^r \quad \text{for} \quad |x| \leq 1, \]
\[ 0 \quad \text{for} \quad |x| > 1. \]

Show that \( f_r \in A(\mathbb{R}) \) for each \( r > 0 \), as a consequence of Proposition 14.17. What is the best conclusion one could draw from Proposition 14.9?
A. Metric spaces, convergence, and compactness

A metric space is a set \( X; \) together with a distance function \( d : X \times X \to [0, \infty) \), having the properties that

\[
\begin{align*}
  d(x, y) &= 0 \iff x = y, \\
  d(x, y) &= d(y, x), \\
  d(x, y) &\leq d(x, z) + d(y, z).
\end{align*}
\]

The third of these properties is called the triangle inequality. An example of a metric space is the set of rational numbers \( \mathbb{Q} \), with \( d(x, y) = |x - y| \). Another example is \( X = \mathbb{R}^n \), with

\[
d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.
\]

If \( (x_\nu) \) is a sequence in \( X \), indexed by \( \nu = 1, 2, 3, \ldots \), i.e., by \( \nu \in \mathbb{Z}^+ \), one says \( x_\nu \to y \) if \( d(x_\nu, y) \to 0 \), as \( \nu \to \infty \). One says \( (x_\nu) \) is a Cauchy sequence if \( d(x_\nu, x_\mu) \to 0 \) as \( \mu, \nu \to \infty \). One says \( X \) is a complete metric space if every Cauchy sequence converges to a limit in \( X \). Some metric spaces are not complete; for example, \( \mathbb{Q} \) is not complete. You can take a sequence \( (x_\nu) \) of rational numbers such that \( x_\nu \to \sqrt{2} \), which is not rational. Then \( (x_\nu) \) is Cauchy in \( \mathbb{Q} \), but it has no limit in \( \mathbb{Q} \).

If a metric space \( X \) is not complete, one can construct its completion \( \tilde{X} \) as follows. Let an element \( \xi \) of \( \tilde{X} \) consist of an equivalence class of Cauchy sequences in \( X \), where we say \( (x_\nu) \sim (y_\nu) \) provided \( d(x_\nu, y_\nu) \to 0 \). We write the equivalence class containing \( (x_\nu) \) as \( [x_\nu] \). If \( \xi = [x_\nu] \) and \( \eta = [y_\nu] \), we can set \( d(\xi, \eta) = \lim_{\mu, \nu \to \infty} d(x_\nu, y_\mu) \), and verify that this is well defined, and makes \( \tilde{X} \) a complete metric space.

If the completion of \( \mathbb{Q} \) is constructed by this process, you get \( \mathbb{R} \), the set of real numbers. This construction provides a good way to develop the basic theory of the real numbers.

There are a number of useful concepts related to the notion of closeness. We define some of them here. First, if \( p \) is a point in a metric space \( X \) and \( r \in (0, \infty) \), the set

\[
B_r(p) = \{ x \in X : d(x, p) < r \}
\]

is called the open ball about \( p \) of radius \( r \). Generally, a neighborhood of \( p \in X \) is a set containing such a ball, for some \( r > 0 \).

A set \( U \subset X \) is called open if it contains a neighborhood of each of its points. The complement of an open set is said to be closed. The following result characterizes closed sets.

**Proposition A.1.** A subset \( K \subset X \) of a metric space \( X \) is closed if and only if

\[
x_j \in K, \ x_j \to p \in X \implies p \in K.
\]
Proof. Assume $K$ is closed, $x_j \in K$, $x_j \to p$. If $p \notin K$, then $p \in X \setminus K$, which is open, so some $B_\varepsilon(p) \subset X \setminus K$, and $d(x_j, p) \geq \varepsilon$ for all $j$. This contradiction implies $p \in K$.

Conversely, assume (A.3) holds, and let $q \in U = X \setminus K$. If $B_{1/n}(q)$ is not contained in $U$ for any $n$, then there exists $x_n \in K \cap B_{1/n}(q)$, hence $x_n \to q$, contradicting (A.3). This completes the proof.

The following is straightforward.

**Proposition A.2.** If $U_\alpha$ is a family of open sets in $X$, then $\bigcup_{\alpha} U_\alpha$ is open. If $K_\alpha$ is a family of closed subsets of $X$, then $\bigcap_{\alpha} K_\alpha$ is closed.

Given $S \subset X$, we denote by $\overline{S}$ (the closure of $S$) the smallest closed subset of $X$ containing $S$, i.e., the intersection of all the closed sets $K_\alpha \subset X$ containing $S$. The following result is straightforward.

**Proposition A.3.** Given $S \subset X$, $p \in \overline{S}$ if and only if there exist $x_j \in S$ such that $x_j \to p$.

Given $S \subset X$, $p \in X$, we say $p$ is an accumulation point of $S$ if and only if, for each $\varepsilon > 0$, there exists $q \in S \cap B_\varepsilon(p)$, $q \neq p$. It follows that $p$ is an accumulation point of $S$ if and only if each $B_\varepsilon(p)$, $\varepsilon > 0$, contains infinitely many points of $S$. One straightforward observation is that all points of $\overline{S} \setminus S$ are accumulation points of $S$.

The interior of a set $S \subset X$ is the largest open set contained in $S$, i.e., the union of all the open sets contained in $S$. Note that the complement of the interior of $S$ is equal to the closure of $X \setminus S$.

We now turn to the notion of compactness. We say a metric space $X$ is compact provided the following property holds:

(A) Each sequence $(x_k)$ in $X$ has a convergent subsequence.

We will establish various properties of compact metric spaces, and provide various equivalent characterizations. For example, it is easily seen that (A) is equivalent to:

(B) Each infinite subset $S \subset X$ has an accumulation point.

The following property is known as total boundedness:

**Proposition A.4.** If $X$ is a compact metric space, then

(C) Given $\varepsilon > 0$, $\exists$ finite set $\{x_1, \ldots, x_N\}$ such that $B_\varepsilon(x_1), \ldots, B_\varepsilon(x_N)$ covers $X$.

Proof. Take $\varepsilon > 0$ and pick $x_1 \in X$. If $B_\varepsilon(x_1) = X$, we are done. If not, pick $x_2 \in X \setminus B_\varepsilon(x_1)$. If $B_\varepsilon(x_1) \cup B_\varepsilon(x_2) = X$, we are done. If not, pick $x_3 \in X \setminus [B_\varepsilon(x_1) \cup B_\varepsilon(x_2)]$. Continue, taking $x_{k+1} \in X \setminus [B_\varepsilon(x_1) \cup \cdots \cup B_\varepsilon(x_k)]$, if $B_\varepsilon(x_1) \cup \cdots \cup B_\varepsilon(x_k) \neq X$. Note that, for $1 \leq i, j \leq k$,

$$i \neq j \implies d(x_i, x_j) \geq \varepsilon.$$

If one never covers $X$ this way, consider $S = \{x_j : j \in \mathbb{N}\}$. This is an infinite set with no accumulation point, so property (B) is contradicted.
Corollary A.5. If $X$ is a compact metric space, it has a countable dense subset.

Proof. Given $\varepsilon = 2^{-n}$, let $S_n$ be a finite set of points $x_j$ such that $\{B_\varepsilon(x_j)\}$ covers $X$. Then $C = \cup_n S_n$ is a countable dense subset of $X$.

Here is another useful property of compact metric spaces, which will eventually be generalized even further, in (E) below.

Proposition A.6. Let $X$ be a compact metric space. Assume $K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots$ form a decreasing sequence of closed subsets of $X$. If each $K_n \neq \emptyset$, then $\cap_n K_n \neq \emptyset$.

Proof. Pick $x_n \in K_n$. If (A) holds, $(x_n)$ has a convergent subsequence, $x_{n_k} \to y$. Since $\{x_{n_k} : k \geq \ell\} \subset K_{n_\ell}$, which is closed, we have $y \in \cap_n K_n$.

Corollary A.7. Let $X$ be a compact metric space. Assume $U_1 \subset U_2 \subset U_3 \subset \cdots$ form an increasing sequence of open subsets of $X$. If $\cup_n U_n = X$, then $U_N = X$ for some $N$.

Proof. Consider $K_n = X \setminus U_n$.

The following is an important extension of Corollary A.7.

Proposition A.8. If $X$ is a compact metric space, then it has the property:

\[(D) \quad \text{Every open cover } \{U_\alpha : \alpha \in \mathcal{A}\} \text{ of } X \text{ has a finite subcover.}\]

Proof. Each $U_\alpha$ is a union of open balls, so it suffices to show that (A) implies the following:

\[(D') \quad \text{Every cover } \{B_\alpha : \alpha \in \mathcal{A}\} \text{ of } X \text{ by open balls has a finite subcover.}\]

Let $C = \{z_j : j \in \mathbb{N}\} \subset X$ be a countable dense subset of $X$, as in Corollary A.2. Each $B_\alpha$ is a union of balls $B_{r_j}(z_j)$, with $z_j \in C \cap B_\alpha$, $r_j$ rational. Thus it suffices to show that

\[(D'') \quad \text{Every countable cover } \{B_j : j \in \mathbb{N}\} \text{ of } X \text{ by open balls has a finite subcover.}\]

For this, we set $U_n = B_1 \cup \cdots \cup B_n$ and apply Corollary A.7.

The following is a convenient alternative to property (D):

\[(E) \quad \text{If } K_\alpha \subset X \text{ are closed and } \bigcap_\alpha K_\alpha = \emptyset, \text{ then some finite intersection is empty.}\]

Considering $U_\alpha = X \setminus K_\alpha$, we see that

\[(D) \iff (E).\]

The following result completes Proposition A.8.
Theorem A.9. For a metric space $X$,

$$(A) \iff (D).$$

Proof. By Proposition A.8, $(A) \Rightarrow (D)$. To prove the converse, it will suffice to show that $(E) \Rightarrow (B)$. So let $S \subset X$ and assume $S$ has no accumulation point. We claim:

Such $S$ must be closed.

Indeed, if $z \in \overline{S}$ and $z \notin S$, then $z$ would have to be an accumulation point. Say $S = \{x_\alpha : \alpha \in A\}$. Then each $K_\alpha$ has no accumulation point, hence $K_\alpha \subset X$ is closed. Also $\cap_\alpha K_\alpha = \emptyset$. Hence there exists a finite set $F \subset A$ such that $\cap_{\alpha \in F} K_\alpha = \emptyset$, if $(E)$ holds. Hence $S = \bigcup_{\alpha \in F} \{x_\alpha\}$ is finite, so indeed $(E) \Rightarrow (B)$.

Remark. So far we have that for every metric space $X$,

$$(A) \iff (B) \iff (D) \iff (E) \iff (C).$$

We claim that $(C)$ implies the other conditions if $X$ is complete. Of course, compactness implies completeness, but $(C)$ may hold for incomplete $X$, e.g., $X = (0, 1) \subset \mathbb{R}$.

Proposition A.10. If $X$ is a complete metric space with property (C), then $X$ is compact.

Proof. It suffices to show that $(C) \Rightarrow (B)$ if $X$ is a complete metric space. So let $S \subset X$ be an infinite set. Cover $X$ by balls $B_{1/2}(x_1), \ldots, B_{1/2}(x_N)$. One of these balls contains infinitely many points of $S$, and so does its closure, say $X_1 = B_{1/2}(y_1)$. Now cover $X$ by finitely many balls of radius $1/4$; their intersection with $X_1$ provides a cover of $X_1$. One such set contains infinitely many points of $S$, and so does its closure $X_2 = \overline{B_{1/4}(y_2)} \cap X_1$. Continue in this fashion, obtaining

$$X_1 \supset X_2 \supset X_3 \supset \cdots \supset X_k \supset X_{k+1} \supset \cdots, \quad X_j \subset \overline{B_{2^{-j}}(y_j)},$$

each containing infinitely many points of $S$. One sees that $(y_j)$ forms a Cauchy sequence. If $X$ is complete, it has a limit, $y_j \to z$, and $z$ is seen to be an accumulation point of $S$.

If $X_j$, $1 \leq j \leq m$, is a finite collection of metric spaces, with metrics $d_j$, we can define a Cartesian product metric space

$$X = \prod_{j=1}^m X_j, \quad d(x, y) = d_1(x_1, y_1) + \cdots + d_m(x_m, y_m).$$

Another choice of metric is $\delta(x, y) = \sqrt{d_1(x_1, y_1)^2 + \cdots + d_m(x_m, y_m)^2}$. The metrics $d$ and $\delta$ are equivalent, i.e., there exist constants $C_0, C_1 \in (0, \infty)$ such that

$$(A.5) \quad C_0 \delta(x, y) \leq d(x, y) \leq C_1 \delta(x, y), \quad \forall x, y \in X.$$

A key example is $\mathbb{R}^m$, the Cartesian product of $m$ copies of the real line $\mathbb{R}$.

We describe some important classes of compact spaces.
Proposition A.11. If $X_j$ are compact metric spaces, $1 \leq j \leq m$, so is $X = \prod_{j=1}^{m} X_j$.

Proof. If $(x_\nu)$ is an infinite sequence of points in $X$, say $x_\nu = (x_{1\nu}, \ldots, x_{m\nu})$, pick a convergent subsequence of $(x_{1\nu})$ in $X_1$, and consider the corresponding subsequence of $(x_\nu)$, which we relabel $(x_\nu)$. Using this, pick a convergent subsequence of $(x_{2\nu})$ in $X_2$. Continue. Having a subsequence such that $x_{j\nu} \to y_j$ in $X_j$ for each $j = 1, \ldots, m$, we then have a convergent subsequence in $X$.

The following result is useful for calculus on $\mathbb{R}^n$.

Proposition A.12. If $K$ is a closed bounded subset of $\mathbb{R}^n$, then $K$ is compact.

Proof. The discussion above reduces the problem to showing that any closed interval $I = [a, b]$ in $\mathbb{R}$ is compact. This compactness is a corollary of Proposition A.10. For pedagogical purposes, we redo the argument here, since in this concrete case it can be streamlined.

Suppose $S$ is a subset of $I$ with infinitely many elements. Divide $I$ into 2 equal subintervals, $I_1 = [a, b_1]$, $I_2 = [b_1, b]$, $b_1 = (a+b)/2$. Then either $I_1$ or $I_2$ must contain infinitely many elements of $S$. Say $I_1$ does. Let $x_1$ be any element of $S$ lying in $I_1$. Now divide $I_1$ into two equal pieces, $I_j = I_{j1} \cup I_{j2}$. One of these intervals (say $I_{j1}$) contains infinitely many points of $S$. Pick $x_2 \in I_{j1}$ to be one such point (different from $x_1$). Then subdivide $I_{j1}$ into two equal subintervals, and continue. We get an infinite sequence of distinct points $x_\nu \in S$, and $|x_\nu - x_{\nu+k}| \leq 2^{-\nu}(b-a)$, for $k \geq 1$. Since $\mathbb{R}$ is complete, $(x_\nu)$ converges, say to $y \in I$. Any neighborhood of $y$ contains infinitely many points in $S$, so we are done.

If $X$ and $Y$ are metric spaces, a function $f : X \to Y$ is said to be continuous provided $x_\nu \to x$ in $X$ implies $f(x_\nu) \to f(x)$ in $Y$. An equivalent condition, which the reader is invited to verify, is

\begin{equation}
U \text{ open in } Y \implies f^{-1}(U) \text{ open in } X.
\end{equation}

Proposition A.13. If $X$ and $Y$ are metric spaces, $f : X \to Y$ continuous, and $K \subset X$ compact, then $f(K)$ is a compact subset of $Y$.

Proof. If $(y_\nu)$ is an infinite sequence of points in $f(K)$, pick $x_\nu \in K$ such that $f(x_\nu) = y_\nu$. If $K$ is compact, we have a subsequence $x_{\nu j} \to p$ in $X$, and then $y_{\nu j} \to f(p)$ in $Y$.

If $F : X \to \mathbb{R}$ is continuous, we say $f \in C(X)$. A useful corollary of Proposition A.13 is:

Proposition A.14. If $X$ is a compact metric space and $f \in C(X)$, then $f$ assumes a maximum and a minimum value on $X$.

Proof. We know from Proposition A.13 that $f(X)$ is a compact subset of $\mathbb{R}$. Hence $f(X)$ is bounded, say $f(X) \subset I = [a, b]$. Repeatedly subdividing $I$ into equal halves, as in the proof of Proposition A.12, at each stage throwing out intervals that do not intersect $f(X)$, and keeping only the leftmost and rightmost interval amongst those remaining, we obtain points $\alpha \in f(X)$ and $\beta \in f(X)$ such that $f(X) \subset [\alpha, \beta]$. Then $\alpha = f(x_0)$ for some $x_0 \in X$ is the minimum and $\beta = f(x_1)$ for some $x_1 \in X$ is the maximum.

At this point, the reader might take a look at the proof of the Mean Value Theorem, given in §0, which applies this result.
If \( S \subset \mathbb{R} \) is a nonempty, bounded set, Proposition A.12 implies \( \overline{S} \) is compact. The function \( \eta : \overline{S} \to \mathbb{R}, \eta(x) = x \) is continuous, so by Proposition A.14 it assumes a maximum and a minimum on \( \overline{S} \). We set

\[
\sup S = \max_{s \in S} x, \quad \inf S = \min_{x \in S} x,
\]

when \( S \) is bounded. More generally, if \( S \subset \mathbb{R} \) is nonempty and bounded from above, say \( S \subset (-\infty, B] \), we can pick \( A < B \) such that \( S \cap [A, B] \) is nonempty, and set

\[
\sup S = \sup S \cap [A, B].
\]

Similarly, if \( S \subset \mathbb{R} \) is nonempty and bounded from below, say \( S \subset [A, \infty) \), we can pick \( B > A \) such that \( S \cap [A, B] \) is nonempty, and set

\[
\inf S = \inf S \cap [A, B].
\]

If \( X \) is a nonempty set and \( f : X \to \mathbb{R} \) is bounded from above, we set

\[
\sup_{x \in X} f(x) = \sup f(X),
\]

and if \( f : X \to \mathbb{R} \) is bounded from below, we set

\[
\inf_{x \in X} f(x) = \inf f(X).
\]

If \( f \) is not bounded from above, we set \( \sup f = +\infty \), and if \( f \) is not bounded from below, we set \( \inf f = -\infty \).

Given a set \( X, f : X \to \mathbb{R} \), and \( x_n \to x \), we set

\[
\limsup_{n \to \infty} f(x_n) = \lim_{n \to \infty} \left( \sup_{k \geq n} f(x_k) \right),
\]

and

\[
\liminf_{n \to \infty} f(x_n) = \lim_{n \to \infty} \left( \inf_{k \geq n} f(x_k) \right).
\]

We return to the notion of continuity. A function \( f \in C(X) \) is said to be \emph{uniformly continuous} provided that, for any \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that

\[
x, y \in X, \quad d(x, y) \leq \delta \implies |f(x) - f(y)| \leq \varepsilon.
\]

An equivalent condition is that \( f \) have a \emph{modulus of continuity}, i.e., a monotonic function \( \omega : [0, 1) \to [0, \infty) \) such that \( \delta \searrow 0 \implies \omega(\delta) \searrow 0 \), and such that

\[
x, y \in X, \quad d(x, y) \leq \delta \leq 1 \implies |f(x) - f(y)| \leq \omega(\delta).
\]

Not all continuous functions are uniformly continuous. For example, if \( X = (0, 1) \subset \mathbb{R} \), then \( f(x) = \sin 1/x \) is continuous, but not uniformly continuous, on \( X \). The following result is useful, for example, in the development of the Riemann integral in §1.
Proposition A.15. If $X$ is a compact metric space and $f \in C(X)$, then $f$ is uniformly continuous.

Proof. If not, there exist $x_\nu, y_\nu \in X$ and $\varepsilon > 0$ such that $d(x_\nu, y_\nu) \leq 2^{-\nu}$ but

$$\tag{A.14} |f(x_\nu) - f(y_\nu)| \geq \varepsilon. $$

Taking a convergent subsequence $x_{\nu_j} \to p$, we also have $y_{\nu_j} \to p$. Now continuity of $f$ at $p$ implies $f(x_{\nu_j}) \to f(p)$ and $f(y_{\nu_j}) \to f(p)$, contradicting (A.14).

If $X$ and $Y$ are metric spaces, the space $C(X, Y)$ of continuous maps $f : X \to Y$ has a natural metric structure, under some additional hypotheses. We use

$$\tag{A.15} D(f, g) = \sup_{x \in X} d(f(x), g(x)).$$

This sup exists provided $f(X)$ and $g(X)$ are bounded subsets of $Y$, where to say $B \subset Y$ is bounded is to say $d : B \times B \to [0, \infty)$ has bounded image. In particular, this supremum exists if $X$ is compact. The following result is useful in the proof of the fundamental local existence result for ODE, in §3.

Proposition A.16. If $X$ is a compact metric space and $Y$ is a complete metric space, then $C(X, Y)$, with the metric (A.9), is complete.

Proof. That $D(f, g)$ satisfies the conditions to define a metric on $C(X, Y)$ is straightforward. We check completeness. Suppose $(f_\nu)$ is a Cauchy sequence in $C(X, Y)$, so, as $\nu \to \infty$,

$$\tag{A.16} \sup_{k \geq 0} \sup_{x \in X} d(f_{\nu_k}(x), f_\nu(x)) \leq \varepsilon_\nu \to 0. $$

Then in particular $(f_\nu(x))$ is a Cauchy sequence in $Y$ for each $x \in X$, so it converges, say to $g(x) \in Y$. It remains to show that $g \in C(X, Y)$ and that $f_\nu \to g$ in the metric (A.9).

In fact, taking $k \to \infty$ in the estimate above, we have

$$\tag{A.17} \sup_{x \in X} d(g(x), f_\nu(x)) \leq \varepsilon_\nu \to 0, $$

i.e., $f_\nu \to g$ uniformly. It remains only to show that $g$ is continuous. For this, let $x_j \to x$ in $X$ and fix $\varepsilon > 0$. Pick $N$ so that $\varepsilon_N < \varepsilon$. Since $f_N$ is continuous, there exists $J$ such that $J \geq J \Rightarrow d(f_N(x_j), f_N(x)) < \varepsilon$. Hence

$$ j \geq J \Rightarrow d(g(x_j), g(x)) \leq d(g(x_j), f_N(x_j)) + d(f_N(x_j), f_N(x)) + d(f_N(x), g(x)) < 3\varepsilon. $$

This completes the proof.

In case $Y = \mathbb{R}$, $C(X, \mathbb{R}) = C(X)$, introduced earlier in this appendix. The distance function (A.15) can be written

$$D(f, g) = \|f - g\|_{\text{sup}}, \quad \|f\|_{\text{sup}} = \sup_{x \in X} |f(x)|.$$
\[ \|f\|_{\text{sup}} \text{ is a norm on } C(X). \]

Generally, a norm on a vector space \( V \) is an assignment \( f \mapsto \|f\| \in [0, \infty) \), satisfying

\[ \|f\| = 0 \iff f = 0, \quad \|af\| = |a| \|f\|, \quad \|f + g\| \leq \|f\| + \|g\|; \]

given \( f, g \in V \) and \( a \) a scalar (in \( \mathbb{R} \) or \( \mathbb{C} \)). A vector space equipped with a norm is called a normed vector space. It is then a metric space, with distance function \( D(f, g) = \|f - g\| \).

If the space is complete, one calls \( V \) a Banach space.

In particular, by Proposition A.16, \( C(X) \) is a Banach space, when \( X \) is a compact metric space.

We next give a couple of slightly more sophisticated results on compactness. The following extension of Proposition A.11 is a special case of Tychonov’s Theorem.

**Proposition A.17.** If \( \{X_j : j \in \mathbb{Z}^+\} \) are compact metric spaces, so is \( X = \prod_{j=1}^{\infty} X_j \).

Here, we can make \( X \) a metric space by setting

\[
(A.18) \quad d(x, y) = \sum_{j=1}^{\infty} 2^{-j} \frac{d_j(p_j(x), p_j(y))}{1 + d_j(p_j(x), p_j(y))},
\]

where \( p_j : X \to X_j \) is the projection onto the \( j \)th factor. It is easy to verify that, if \( x_\nu \in X \), then \( x_\nu \to y \) in \( X \), as \( \nu \to \infty \), if and only if, for each \( j \), \( p_j(x_\nu) \to p_j(y) \) in \( X_j \).

**Proof.** Following the argument in Proposition A.11, if \( (x_\nu) \) is an infinite sequence of points in \( X \), we obtain a nested family of subsequences

\[
(A.19) \quad (x_\nu) \supset (x_1^1_\nu) \supset (x_2^2_\nu) \supset \cdots \supset (x_j^j_\nu) \supset \cdots
\]

such that \( p_\ell(x_\nu^j) \) converges in \( X_\ell \), for \( 1 \leq \ell \leq j \). The next step is a diagonal construction. We set

\[
(A.20) \quad \xi_\nu = x_\nu^j_\nu \in X.
\]

Then, for each \( j \), after throwing away a finite number \( N(j) \) of elements, one obtains from \( (\xi_\nu) \) a subsequence of the sequence \( (x_\nu^j) \) in (A.19), so \( p_\ell(\xi_\nu) \) converges in \( X_\ell \) for all \( \ell \). Hence \( (\xi_\nu) \) is a convergent subsequence of \( (x_\nu) \).

The next result is a special case of Ascoli’s Theorem.

**Proposition A.18.** Let \( X \) and \( Y \) be compact metric spaces, and fix a modulus of continuity \( \omega(\delta) \). Then

\[
(A.21) \quad C_\omega = \{ f \in C(X, Y) : d(f(x), f(x')) \leq \omega(d(x, x')) \ \forall x, x' \in X \}
\]

is a compact subset of \( C(X, Y) \).

**Proof.** Let \( (f_\nu) \) be a sequence in \( C_\omega \). Let \( \Sigma \) be a countable dense subset of \( X \), as in Corollary A.5. For each \( x \in \Sigma \), \( (f_\nu(x)) \) is a sequence in \( Y \), which hence has a convergent subsequence.
Using a diagonal construction similar to that in the proof of Proposition A.17, we obtain a subsequence \((\varphi_\nu)\) of \((f_\nu)\) with the property that \(\varphi_\nu(x)\) converges in \(Y\), for each \(x \in \Sigma\), say

\[(A.22) \quad \varphi_\nu(x) \to \psi(x),\]

for all \(x \in \Sigma\), where \(\psi : \Sigma \to Y\).

So far, we have not used \((A.21)\). This hypothesis will now be used to show that \(\varphi_\nu\) converges uniformly on \(X\). Pick \(\varepsilon > 0\). Then pick \(\delta > 0\) such that \(\omega(\delta) < \varepsilon/3\). Since \(X\) is compact, we can cover \(X\) by finitely many balls \(B_\delta(x_j), \ 1 \leq j \leq N, \ x_j \in \Sigma\). Pick \(M\) so large that \(\varphi_\nu(x_j)\) is within \(\varepsilon/3\) of its limit for all \(\nu \geq M\) (when \(1 \leq j \leq N\)). Now, for any \(x \in X\), picking \(\ell \in \{1, \ldots, N\}\) such that \(d(x, x_\ell) \leq \delta\), we have, for \(k \geq 0, \ \nu \geq M\),

\[
(A.23) \quad d(\varphi_{\nu+k}(x), \varphi_\nu(x)) \leq d(\varphi_{\nu+k}(x_\ell), \varphi_\nu(x_\ell)) + d(\varphi_{\nu+k}(x_\ell), \varphi_\nu(x_\ell)) + d(\varphi_\nu(x_\ell), \varphi_\nu(x)) \leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3.
\]

Thus \((\varphi_\nu(x))\) is Cauchy in \(Y\) for all \(x \in X\), hence convergent. Call the limit \(\psi(x)\), so we now have \((A.22)\) for all \(x \in X\). Letting \(k \to \infty\) in \((A.23)\) we have uniform convergence of \(\varphi_\nu\) to \(\psi\). Finally, passing to the limit \(\nu \to \infty\) in \((A.24)\)

\[
(A.24) \quad d(\varphi_\nu(x), \varphi_\nu(x')) \leq \omega(d(x, x'))
\]

gives \(\psi \in C_\omega\).

We want to re-state Proposition A.18, bringing in the notion of equicontinuity. Given metric spaces \(X\) and \(Y\), and a set of maps \(\mathcal{F} \subset C(X,Y)\), we say \(\mathcal{F}\) is equicontinuous at a point \(x_0 \in X\) provided

\[
(A.25) \quad \forall \varepsilon > 0, \ \exists \delta > 0 \text{ such that } \forall x \in X, \ f \in \mathcal{F}, \ d_X(x, x_0) < \delta \implies d_Y(f(x), f(x_0)) < \varepsilon.
\]

We say \(\mathcal{F}\) is equicontinuous on \(X\) if it is equicontinuous at each point of \(X\). We say \(\mathcal{F}\) is uniformly equicontinuous on \(X\) provided

\[
(A.26) \quad \forall \varepsilon > 0, \ \exists \delta > 0 \text{ such that } \forall x, x' \in X, \ f \in \mathcal{F}, \ d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \varepsilon.
\]

Note that \((A.26)\) is equivalent to the existence of a modulus of continuity \(\omega\) such that \(\mathcal{F} \subset C_\omega\), given by \((A.21)\). It is useful to record the following result.

**Proposition A.19.** Let \(X\) and \(Y\) be metric spaces, \(\mathcal{F} \subset C(X,Y)\). Assume \(X\) is compact. then

\[
(A.27) \quad \mathcal{F} \text{ equicontinuous } \implies \mathcal{F} \text{ is uniformly equicontinuous.}
\]
Proof. The argument is a variant of the proof of Proposition A.15. In more detail, suppose there exist \( x, x' \in X, \) \( \varepsilon > 0, \) and \( f \in F \) such that \( d(x, x') \leq 2^{-\nu} \) but

\[
\text{(A.28)} \quad d(f(x), f(x')) \geq \varepsilon.
\]

Taking a convergent subsequence \( x_{\nu_j} \to p \in X, \) we also have \( x'_{\nu_j} \to p. \) Now equicontinuity of \( F \) at \( p \) implies that there exists \( N < \infty \) such that

\[
\text{(A.29)} \quad d(g(x_{\nu_j}), g(p)) < \frac{\varepsilon}{2}, \quad \forall j \geq N, \ g \in F,
\]

contradicting (A.28).

Putting together Propositions A.18 and A.19 then gives the following.

**Proposition A.20.** Let \( X \) and \( Y \) be compact metric spaces. If \( F \subset C(X,Y) \) is equicontinuous on \( X, \) then it has compact closure in \( C(X,Y). \)

We next define the notion of a connected space. A metric space \( X \) is said to be connected provided that it cannot be written as the union of two disjoint nonempty open subsets. The following is a basic class of examples.

**Proposition A.21.** Each interval \( I \) in \( \mathbb{R} \) is connected.

**Proof.** Suppose \( A \subset I \) is nonempty, with nonempty complement \( B \subset I, \) and both sets are open. Take \( a \in A, \ b \in B; \) we can assume \( a < b. \) Let \( \xi = \sup\{x \in [a, b] : x \in A\} \) This exists, as a consequence of the basic fact that \( \mathbb{R} \) is complete.

Now we obtain a contradiction, as follows. Since \( A \) is closed \( \xi \in A. \) But then, since \( A \) is open, there must be a neighborhood \((\xi - \varepsilon, \xi + \varepsilon)\) contained in \( A; \) this is not possible.

We say \( X \) is path-connected if, given any \( p, q \in X, \) there is a continuous map \( \gamma : [0, 1] \to X \) such that \( \gamma(0) = p \) and \( \gamma(1) = q. \) It is an easy consequence of Proposition A.21 that \( X \) is connected whenever it is path-connected.

The next result, known as the Intermediate Value Theorem, is frequently useful.

**Proposition A.22.** Let \( X \) be a connected metric space and \( f : X \to \mathbb{R} \) continuous. Assume \( p, q \in X, \) and \( f(p) = a < f(q) = b. \) Then, given any \( c \in (a, b), \) there exists \( z \in X \) such that \( f(z) = c. \)

**Proof.** Under the hypotheses, \( A = \{x \in X : f(x) < c\} \) is open and contains \( p, \) while \( B = \{x \in X : f(x) > c\} \) is open and contains \( q. \) If \( X \) is connected, then \( A \cup B \) cannot be all of \( X; \) so any point in its complement has the desired property.

**Exercises**

1. If \( X \) is a metric space, with distance function \( d, \) show that

\[
|d(x, y) - d(x', y')| \leq d(x, x') + d(y, y'),
\]
and hence 
\[ d : X \times X \rightarrow [0, \infty) \] is continuous.

2. Let \( \varphi : [0, \infty) \rightarrow [0, \infty) \) be a \( C^2 \) function. Assume 
\[ \varphi(0) = 0, \quad \varphi' > 0, \quad \varphi'' < 0. \]
Prove that if \( d(x, y) \) is symmetric and satisfies the triangle inequality, so does 
\[ \delta(x, y) = \varphi(d(x, y)). \]

*Hint.* Show that such \( \varphi \) satisfies \( \varphi(s + t) \leq \varphi(s) + \varphi(t) \), for \( s, t \in \mathbb{R}^+ \).

3. Show that the function \( d(x, y) \) defined by (A.18) satisfies (A.1).

*Hint.* Consider \( \varphi(r) = r/(1 + r) \).

4. Let \( X \) be a compact metric space. Assume \( f_j, f \in C(X) \) and 
\[ f_j(x) \nearrow f(x), \quad \forall x \in X. \]
Prove that \( f_j \to f \) uniformly on \( X \). (This result is called Dini’s theorem.)

*Hint.* For \( \varepsilon > 0 \), let \( K_j(\varepsilon) = \{ x \in X : f(x) - f_j(x) \geq \varepsilon \} \). Note that \( K_j(\varepsilon) \supset K_{j+1}(\varepsilon) \supset \cdots \).

Given a metric space \( X \) and \( f : X \to [-\infty, \infty] \), we say \( f \) is lower semicontinuous at \( x \in X \) provided 
\[ f^{-1}((c, \infty)) \subset X \] is open, \( \forall c \in \mathbb{R} \).

We say \( f \) is upper semicontinuous provided 
\[ f^{-1}([-\infty, c)) \] is open, \( \forall c \in \mathbb{R} \).

5. Show that 
\[ f \text{ is lower semicontinuous } \iff f^{-1}([-\infty, c]) \text{ is closed, } \forall c \in \mathbb{R}, \]
and 
\[ f \text{ is upper semicontinuous } \iff f^{-1}([c, \infty]) \text{ is closed, } \forall c \in \mathbb{R}. \]

6. Show that 
\[ f \text{ is lower semicontinuous } \iff x_n \to x \text{ implies } \liminf f(x_n) \geq f(x). \]
Show that 
\[ f \text{ is upper semicontinuous } \iff x_n \to x \text{ implies } \limsup f(x_n) \leq f(x). \]
7. Given $S \subset X$, show that

- $\chi_S$ is lower semicontinuous $\iff S$ is open.
- $\chi_S$ is upper semicontinuous $\iff S$ is closed.

8. If $X$ is a compact metric space, show that

$$f : X \to \mathbb{R}$$

is lower semicontinuous $\implies \min f$ is achieved.

9. In the setting of (A.4), let

$$\delta(x, y) = \left\{d_1(x_1, y_1)^2 + \cdots + d_m(x_m, y_m)^2\right\}^{1/2}.$$ 

Show that

$$\delta(x, y) \leq d(x, y) \leq \sqrt{m} \delta(x, y).$$

10. Let $X$ and $Y$ be compact metric spaces. Show that if $\mathcal{F} \subset C(X, Y)$ is compact, then $\mathcal{F}$ is equicontinuous. (This is a converse to Proposition A.20.)

11. Recall that a Banach space is a complete normed linear space. Consider $C^1(I)$, where $I = [0, 1]$, with norm

$$\|f\|_{C^1} = \sup_I |f| + \sup_I |f'|.$$ 

Show that $C^1(I)$ is a Banach space.

12. Let $\mathcal{F} = \{f \in C^1(I) : \|f\|_{C^1} \leq 1\}$. Show that $\mathcal{F}$ has compact closure in $C(I)$. Find a function in the closure of $\mathcal{F}$ that is not in $C^1(I)$. 
B. Partitions of unity

In the text we have made occasional use of partitions of unity, and we include some material on this topic here. We begin by defining and constructing a continuous partition of unity on a compact metric space, subordinate to a open cover \(\{U_j : 1 \leq j \leq N\}\) of \(X\). By definition, this is a family of continuous functions \(\varphi_j : X \to \mathbb{R}\) such that

\[
\varphi_j \geq 0, \quad \text{supp} \varphi_j \subset U_j, \quad \sum_j \varphi_j = 1.
\]

To construct such a partition of unity, we do the following. First, it can be shown that there is an open cover \(\{V_j : 1 \leq j \leq N\}\) of \(X\) and open sets \(W_j\) such that

\[
V_j \subset W_j \subset \overline{W}_j \subset U_j.
\]

Given this, let \(\psi_j(x) = \text{dist}(x, X \setminus W_j)\). Then \(\psi_j\) is continuous, supp \(\psi_j \subset \overline{W}_j \subset U_j\), and \(\psi_j\) is strictly positive on \(V_j\). Hence \(\Psi = \sum_j \psi_j\) is continuous and strictly positive on \(X\), and we see that

\[
\varphi_j(x) = \Psi(x)^{-1}\psi_j(x)
\]

yields such a partition of unity.

We indicate how to construct the sets \(V_j\) and \(W_j\) used above, starting with \(V_1\) and \(W_1\). Note that the set \(K_1 = X \setminus (U_2 \cup \cdots \cup U_N)\) is a compact subset of \(U_1\). Assume it is nonempty; otherwise just throw \(U_1\) out and relabel the sets \(U_j\). Now set

\[
V_1 = \{x \in U_1 : \text{dist}(x, K_1) < \frac{1}{3}\text{dist}(x, X \setminus U_1)\},
\]

and

\[
W_1 = \{x \in U_1 : \text{dist}(x, K_1) < \frac{2}{3}\text{dist}(x, X \setminus U_1)\}.
\]

To construct \(V_2\) and \(W_2\), proceed as above, but use the cover \(\{U_2, \ldots, U_N, V_1\}\). Continue until done.

Given a smooth compact surface \(M\) (perhaps with boundary), covered by coordinate patches \(U_j\) \((1 \leq j \leq N)\), one can construct a smooth partition of unity on \(M\), subordinate to this cover. The main additional tool for this is the construction of a function \(\psi \in C_0^\infty(\mathbb{R}^n)\) such that

\[
\psi(x) = 1 \quad \text{for} \quad |x| \leq \frac{1}{2}, \quad \psi(x) = 0 \quad \text{for} \quad |x| \geq 1.
\]

One way to get this is to start with the function on \(\mathbb{R}\) given by

\[
f_0(x) = e^{-1/x} \quad \text{for} \quad x > 0,
\]

\[
0 \quad \text{for} \quad x \leq 0.
\]
It is an exercise to show that

$$f_0 \in C^\infty(\mathbb{R}).$$

Now the function

$$f_1(x) = f_0(x)f_0(\frac{1}{2} - x)$$

belongs to $C^\infty(\mathbb{R})$ and is zero outside the interval $[0, 1/2]$. Hence the function

$$f_2(x) = \int_{-\infty}^{x} f_1(s) \, ds$$

belongs to $C^\infty(\mathbb{R})$, is zero for $x \leq 0$, and equals some positive constant (say $C_2$) for $x \geq 1/2$. Then

$$\psi(x) = \frac{1}{C_2} f_2(1 - |x|)$$

is a function on $\mathbb{R}^n$ with the desired properties.

With this function in hand, to construct the smooth partition of unity mentioned above is an exercise we recommend to the reader.
C. Differential forms and the change of variable formula

The change of variable formula for one-variable integrals,

\[ \int_a^t f(\varphi(x)) \varphi'(x) \, dx = \int_{\varphi(a)}^{\varphi(t)} f(x) \, dx; \]

given \( f \) continuous and \( \varphi \) of class \( C^1 \), is easily established, via the fundamental theorem of calculus and the chain rule. We recall how this was done in §0. If we denote the left side of (C.1) by \( A(t) \) and the right by \( B(t) \), we apply these results to get

\[ A'(t) = f(\varphi(t))\varphi'(t) = B'(t), \]

and since \( A(a) = B(a) = 0 \), another application of the fundamental theorem of calculus (or simply the mean value theorem) gives \( A(t) = B(t) \).

For multiple integrals, the change of variable formula takes the following form, given in Proposition 4.13:

**Theorem C.1.** Let \( O, \Omega \) be connected, open sets in \( \mathbb{R}^n \) and let \( \varphi : O \to \Omega \) be a \( C^1 \) diffeomorphism. Given \( f \) continuous on \( \Omega \), with compact support, we have

\[ \int_O f(\varphi(x)) \det D\varphi(x) \, dx = \int_\Omega f(x) \, dx; \]

There are many variants of Theorem C.1. In particular one wants to extend the class of functions \( f \) for which (C.3) holds, but once one has Theorem C.1 as stated, such extensions are relatively painless. See the derivation of Theorem 4.15.

Let’s face it; the proof of Theorem C.1 given in §4 was a grim affair, involving careful estimates of volumes of images of small cubes under the map \( \varphi \) and numerous pesky details. Recently, P. Lax [L] found a fresh approach to the proof of the multidimensional change of variable formula. More precisely, [L] established the following result, from which Theorem C.1 is a consequence.

**Theorem C.2.** Let \( \varphi : \mathbb{R}^n \to \mathbb{R}^n \) be a \( C^1 \) map. Assume \( \varphi(x) = x \) for \( |x| \geq R \). Let \( f \) be a continuous function on \( \mathbb{R}^n \) with compact support. Then

\[ \int f(\varphi(x)) \det D\varphi(x) \, dx = \int f(x) \, dx. \]

We will give a variant of the proof of [L]. One difference between this proof and that of [L] is that we use the language of differential forms.
Proof of Theorem C.2. Via standard approximation arguments, it suffices to prove this when \( \varphi \) is \( C^2 \) and \( f \in C^1_0(\mathbb{R}^n) \), which we will assume from here on.

To begin, pick \( A > 0 \) such that \( f(x - Ae_1) \) is supported in \( \{ x : |x| > R \} \), where \( e_1 = (1,0,\ldots,0) \). Also take \( A \) large enough that the image of \( \{ x : |x| \leq R \} \) under \( \varphi \) does not intersect the support of \( f(x - Ae_1) \). We can set

\[ F(x) = f(x) - f(x - Ae_1) = \frac{\partial \psi}{\partial x_1}(x), \]

where

\[ \psi(x) = \int_0^A f(x - se_1) \, ds, \quad \psi \in C^1_0(\mathbb{R}^n). \]

Then we have the following identities involving \( n \)-forms:

\[ \alpha = F(x) \, dx_1 \wedge \cdots \wedge dx_n = \frac{\partial \psi}{\partial x_1} \, dx_1 \wedge \cdots \wedge dx_n \]

\[ = d\psi \wedge dx_2 \wedge \cdots \wedge dx_n \]

\[ = d(\psi \wedge dx_2 \wedge \cdots \wedge dx_n), \]

i.e., \( \alpha = d\beta \), with \( \beta = \psi \wedge dx_2 \wedge \cdots \wedge dx_n \) a compactly supported \((n - 1)\)-form of class \( C^1 \).

Now the pull-back of \( \alpha \) under \( \varphi \) is given by

\[ \varphi^* \alpha = F(\varphi(x)) \det D\varphi(x) \, dx_1 \wedge \cdots \wedge dx_n. \]

Furthermore, the right side of (C.8) is equal to

\[ f(\varphi(x)) \det D\varphi(x) \, dx_1 \wedge \cdots \wedge dx_n - f(x - Ae_1) \, dx_1 \wedge \cdots \wedge dx_n. \]

Hence we have

\[ \int f(\varphi(x)) \det D\varphi(x) \, dx_1 \cdots dx_n - \int f(x) \, dx_1 \cdots dx_n \]

\[ = \int \varphi^* \alpha = \int \varphi^* d\beta = \int d(\varphi^* \beta), \]

where we use the general identity

\[ \varphi^* d\beta = d(\varphi^* \beta), \]

a consequence of the chain rule. On the other hand, a very special case of Stokes’ theorem applies to

\[ \varphi^* \beta = \gamma = \sum_j \gamma_j(x) \, d\overline{x_1} \wedge \cdots \wedge \overline{dx_j} \wedge \cdots \wedge dx_n, \]
with $\gamma_j \in C^1_0(\mathbb{R}^n)$. Namely

$$d\gamma = \sum_j (-1)^{j-1} \frac{\partial \gamma_j}{\partial x_j} dx_1 \wedge \cdots \wedge dx_n,$$

and hence, by the fundamental theorem of calculus,

$$\int d\gamma = 0.$$

This gives the desired identity (C.4), from (C.10).

We make some remarks on Theorem C.2. Note that $\varphi$ is not assumed to be one-to-one or onto. In fact, as noted in [L], the identity (C.4) implies that such $\varphi$ must be onto, and this has important topological implications. We mention that, if one puts absolute values around $\det D\varphi(x)$ in (C.4), the appropriate formula is

$$\int f(\varphi(x)) |\det D\varphi(x)| dx = \int f(x) n(x) dx,$$

where $n(x) = \# \{ y : \varphi(y) = x \}$. A proof of (C.15) can be found in texts on geometrical measure theory.

As noted in [L], Theorem C.2 was proven in [B-D]. The proof there makes use of differential forms and Stokes’ theorem, but it is quite different from the proof given here. A crucial difference is that the proof in [B-D] requires that one knows the change of variable formula as formulated in Theorem C.1.

We now show how Theorem C.1 can be deduced from Theorem C.2. We will use the following lemma.

Lemma C.3. In the setting of Theorem C.1, and with $\det D\varphi > 0$, given $p \in \Omega$, there exists a neighborhood $U$ of $p$ and a $C^1$ map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\Phi = \varphi \text{ on } \varphi^{-1}(U), \quad \Phi(x) = x \text{ for } |x| \text{ large},$$

and

$$\Phi(x) \in U \implies x \in \varphi^{-1}(U).$$

Granted the lemma, we proceed as follows. Assume $\det D\varphi > 0$ on $\Omega$. Given $f \in C(\Omega)$, $\text{supp } f \subset K$, compact in $\Omega$, cover $K$ with a finite number of subsets $U_j$ as in Lemma C.3, and, using a continuous partition of unity (cf. Appendix B), write $f = \sum_j f_j$, $\text{supp } f_j \subset U_j$. Also, let $\Phi_j$ have the obvious significance. By Theorem C.2, we have

$$\int f_j(\Phi_j(x)) |\det D\Phi_j(x)| dx = \int f_j \, dx.$$
But we also have
\[
(C.19) \quad \int f_j(\Phi_j(x)) \det D\Phi_j(x) \, dx = \int f_j(\varphi(x)) \det D\varphi(x) \, dx.
\]

Now summing over \(j\) gives (C.3).

If we do not have \(\det D\varphi > 0\) on \(\mathcal{O}\), then \(\det D\varphi < 0\) on \(\mathcal{O}\). In this case, one can compose with the map
\[
(C.20) \quad \kappa : \mathbb{R}^n \to \mathbb{R}^n, \quad \kappa(x, x') = (-x_1, x'),
\]
(for which Theorem C.1 is elementary) and readily recover the desired result.

We turn to the proof of Lemma C.3. Say \(q = \varphi^{-1}(p)\), \(D\varphi(q) = A \in G\ell_+(n, \mathbb{R})\), i.e., \(A \in G\ell(n, \mathbb{R})\) and \(\det A > 0\). Translating coordinates, we can assume \(p = q = 0\). We set
\[
(C.21) \quad \Psi(x) = \beta(x)\varphi(x) + (1 - \beta(x))Ax,
\]
where \(\beta \in C_0^\infty(\mathbb{R}^n)\) has support in a small neighborhood of \(q\) and \(\beta \equiv 1\) on a smaller neighborhood \(V = \varphi^{-1}(U)\), chosen so that we can apply Corollary 2.4, to deduce that
\[
(C.22) \quad \Psi \text{ maps } \mathbb{R}^n \text{ diffeomorphically onto its image, an open set in } \mathbb{R}^n.
\]

In fact, estimates behind the proof of Proposition 2.2 imply that, for appropriately chosen \(\beta\), there exists \(b > 0\) such that \(|\Psi(x) - \Psi(y)| \geq b|x - y|\) for all \(x, y \in \mathbb{R}^n\). Hence the image \(\Psi(\mathbb{R}^n)\) is closed in \(\mathbb{R}^n\), as well as open, so actually \(\Psi\) maps \(\mathbb{R}^n\) diffeomorphically onto \(\mathbb{R}^n\).

Note that \(\Psi = \varphi\) on \(V = \varphi^{-1}(U)\). We want to alter \(\Psi(x)\) for large \(|x|\) to obtain \(\Phi\), satisfying (C.16)–(C.17). To do this, we use the fact that \(G\ell_+(n, \mathbb{R})\) is connected (see Proposition 5.14). Pick a smooth path \(\Gamma : [0, 1] \to G\ell_+(n, \mathbb{R})\) such that \(\Gamma(t) = A\) for \(t \in [0, 1/4]\) and \(\Gamma(t) = I\) for \(t \in [3/4, 1]\). Let
\[
(C.23) \quad M = \sup_{0 \leq t \leq 1} \|\Gamma(t)^{-1}\|, \text{ so } |\Gamma(t)x| \geq M^{-1}|x|, \forall x \in \mathbb{R}^n.
\]

Now assume \(U \subset B_{R_2} = \{x \in \mathbb{R}^n : |x| < R_1\}\), so \(\Psi(V) \subset B_{R_1}\). Next, take \(R_2\) so large that \(V = \varphi^{-1}(U) \subset B_{R_2}\) and
\[
(C.24) \quad |x| \geq R_2 \implies |\Psi(x)| > MR_1 \text{ and } \Psi(x) = Ax.
\]

Now set
\[
(C.25) \quad \Phi(x) = \Psi(x) \text{ for } |x| \leq R_2,
\]
\[
\Gamma(t)x \text{ for } |x| = R_2 + t, \ 0 \leq t \leq 1,
\]
\[
x \text{ for } |x| \geq R_2 + 1.
\]

This map has the properties (C.16)–(C.17).
D. Complements on power series

If a function \( f \) is sufficiently differentiable on an interval in \( \mathbb{R} \) containing \( x \) and \( y \), the Taylor expansion about \( y \) reads

\[
\text{(D.1)} \quad f(x) = f(y) + f'(y)(x - y) + \cdots + \frac{1}{n!} f^{(n)}(y)(x - y)^n + R_n(x, y).
\]

Here, \( T_n(x; y) = f(y) + \cdots + \frac{f^{(n)}(y)}{n!}(x - y)^n \) is that polynomial of degree \( n \) in \( x \) all of whose \( x \)-derivatives of order \( \leq n \), evaluated at \( y \), coincide with those of \( f \). This prescription makes the formula for \( T_n(x; y) \) easy to derive. The analysis of the remainder term \( R_n(x, y) \) is more subtle. One useful result about this remainder is the following. Say \( x > y \), and for simplicity assume \( f^{(n+1)} \) is continuous on \([y; x] \); we say \( f \in C^{n+1}([y; x]) \). Then

\[
\text{(D.2)} \quad m \leq f^{(n+1)}(\xi) \leq M, \quad \forall \, \xi \in [y; x] \implies m \frac{(x - y)^{n+1}}{(n + 1)!} \leq R_n(x, y) \leq M \frac{(x - y)^{n+1}}{(n + 1)!}.
\]

Under our hypotheses, this result is equivalent to the Lagrange form of the remainder:

\[
\text{(D.3)} \quad R_n(x, y) = \frac{1}{(n + 1)!} (x - y)^{n+1} f^{(n+1)}(\zeta_n),
\]

for some \( \zeta_n \) between \( x \) and \( y \). There are various proofs of (D.3). One will be given below.

One of our purposes here is to comment on how effective estimates on \( R_n(x, y) \) are in determining the convergence of the infinite series

\[
\text{(D.4)} \quad \sum_{k=0}^{\infty} \frac{f^{(k)}(y)}{k!} (x - y)^k
\]

to \( f(x) \). That is to say, we want to perceive that \( R_n(x, y) \to 0 \) as \( n \to \infty \), in appropriate circumstances. Before we look at how effective the estimate (D.2) is at this job, we want to introduce another player, and along the way discuss the derivation of various formulas for the remainder in (D.1).

A simple formula for \( R_n(x, y) \) follows upon taking the \( y \)-derivative of both sides of (D.1); we are assuming that \( f \) is at least \( (n+1) \)-fold differentiable. When we do this (applying the Leibniz formula to those terms that are products) an enormous amount of cancellation arises, and the formula collapses to

\[
\text{(D.5)} \quad \frac{\partial R_n}{\partial y} = -\frac{1}{n!} f^{(n+1)}(y)(x - y)^n, \quad R_n(x, x) = 0.
\]

If we concentrate on \( R_n(x, y) \) as a function of \( y \) and look at the difference quotient \([R_n(x, y) - R_n(x, x)]/(y - x)\), an immediate consequence of the mean value theorem is that

\[
\text{(D.6)} \quad R_n(x, y) = \frac{1}{n!} (x - y) (x - \xi_n)^n f^{(n+1)}(\xi_n),
\]
for some $\xi_n$ between $x$ and $y$. This result, known as Cauchy’s formula for the remainder, has a slightly more complicated appearance than (D.3), but as we will see it has advantages over Lagrange’s formula. The application of the mean value theorem to obtain (D.6) does not require the continuity of $f^{(n+1)}$, but we do not want to dwell on that point.

If $f^{(n+1)}$ is continuous, we can apply the Fundamental Theorem of Calculus to (D.5), in the $y$-variable, and obtain the basic integral formula

$$R_n(x, y) = \frac{1}{n!} \int_y^x (x-s)^n f^{(n+1)}(s) \, ds.$$  

Another proof of (D.7) is indicated in Exercise 9 of §0. If we think of the integral in (D.7) as $(x-y)$ times the mean value of the integrand, we see (D.6) as a consequence. On the other hand, if we want to bring a factor of $(x-y)^{n+1}$ outside the integral in (D.7), the change of variable $x-s = t(x-y)$ gives the integral formula

$$R_n(x, y) = \frac{1}{n!} (x-y)^{n+1} \int_0^1 t^n f^{(n+1)}(ty + (1-t)x) \, dt.$$  

If we think of this integral as $1/(n+1)$ times a weighted mean value of $f^{(n+1)}$, we recover the Lagrange formula (D.3).

From the Lagrange form (D.3) of the remainder in the Taylor series (D.1) we have the estimate

$$|R_n(x, y)| \leq \frac{|x-y|^{n+1}}{(n+1)!} \sup_{\xi \in I(x,y)} |f^{(n+1)}(\xi)|,$$

where $I(x,y)$ is the open interval from $x$ to $y$ (either $(x,y)$ or $(y,x)$, disregarding the trivial case $x = y$). Meanwhile, from the Cauchy form (D.6) of the remainder we have the estimate

$$|R_n(x, y)| \leq \frac{|x-y|}{n!} \sup_{\xi \in I(x,y)} |(x-\xi)^n f^{(n+1)}(\xi)|.$$

We now study how effective these estimates are in determining that various power series converge.

We begin with a look at these remainder estimates for the power series expansion about the origin of the simple function

$$f(x) = \frac{1}{1-x}.$$  

We have, for $x \neq 1$,

$$f^{(k)}(x) = k! (1-x)^{-k-1},$$
and formula (D.1) becomes

\[
\frac{1}{1-x} = 1 + x + \cdots + x^n + R_n(x,0).
\]

Of course, everyone knows that the infinite series

\[
1 + x + \cdots + x^n + \cdots
\]

converges to \( f(x) \) in (D.11), precisely for \( x \in (-1,1) \). What we are interested in is what can be deduced from the estimate (D.9), which, for the function (D.11), takes the form

\[
|R_n(x,0)| \leq |x|^{n+1} \sup_{0 \leq \zeta \leq x} |1 - \zeta|^{-n-2}.
\]

We consider two cases. First, if \( x \leq 0 \), then \( |1 - \zeta| \geq 1 \) for \( \zeta \in I(x,0) \), so

\[
x \leq 0 \implies |R_n(x,0)| \leq |x|^{n+1}.
\]

Thus the estimate (D.9) implies that \( R_n(x,0) \to 0 \) in (D.13), for all \( x \in (-1,0] \). Suppose however that \( x \geq 0 \). What we have from (D.15) is

\[
x \geq 0 \implies |R_n(x,0)| \leq |x|^{n+1} \sup_{0 \leq \zeta \leq x} |1 - \zeta|^{-n-2}
\]

\[
= \frac{1}{1-x} \left( \frac{x}{1-x} \right)^{n+1}.
\]

This tends to 0 as \( n \to \infty \) if and only if \( x < 1 - x \), i.e., if and only if \( x < 1/2 \). What we have is the following:

**Conclusion.** The estimate (D.9) implies the convergence of the Taylor series (about the origin) for the function \( f(x) = 1/(1-x) \), only for \( -1 < x < 1/2 \).

This example points to a weakness in the estimate (D.9). Now let us see how well we can do with the estimate (D.10). For the function (D.11), this takes the form

\[
|R_n(x,0)| \leq (n+1) |x| \sup_{\xi \in I(x,0)} \frac{|x-\xi|^n}{|1-\xi|^{n+2}}.
\]

For \( -1 < x \leq 0 \) one has an estimate like (D.16), with a harmless factor of \((n+1)\) thrown in. On the other hand, one readily verifies that

\[
0 \leq \xi \leq x < 1 \implies \frac{x-\xi}{1-\xi} \leq x,
\]

so we deduce from (D.18) that

\[
0 \leq x < 1 \implies |R_n(x,0)| \leq (n+1) \frac{x^{n+1}}{1-x},
\]
which does tend to 0 for all \( x \in [0, 1) \).

One might be wondering if one could come up with some more complicated example, for which Cauchy’s form is effective only on an interval shorter than the interval of convergence. In fact, you can’t. Cauchy’s form of the remainder is always effective in the interior of the interval of convergence. A proof of this, making use of some complex analysis, is given in [T3].

We look at some more power series, and see when convergence can be established at an endpoint of an interval of convergence, using the estimate (D.10), i.e.,

\[(D.20) \quad |R_n(x, y)| \leq C_n(x, y), \quad C_n(x, y) = \frac{|x - y|}{n!} \sup_{\xi \in I(x, y)} |(x - \xi)^n f^{(n+1)}(\xi)|.\]

We consider the following family of examples:

\[(D.21) \quad f(x) = (1 - x)^a, \quad a > 0.\]

The power series expansion has radius of convergence 1 (if \( a \) is not an integer) and, as we will see, one has convergence at both endpoints, +1 and −1, whenever \( a > 0 \). Let us see when \( C_n(\pm1, 0) \to 0 \). We have

\[(D.22) \quad f^{(n+1)}(x) = (-1)^{n+1} a(a - 1) \cdots (a - n) (1 - x)^{a-n-1}.\]

Hence

\[(D.23) \quad C_n(-1, 0) = \left| \frac{a(a - 1) \cdots (a - n)}{n!} \right| \sup_{-1 < \xi < 0} \left| \frac{-1 - \xi}{1 - \xi^{n+1-a}} \right| ^n \]

\[= \left| a(1 - a)\left(1 - \frac{a}{2}\right) \cdots \left(1 - \frac{a}{n}\right) \right| \]

\[= \mathcal{O}(n^{-a}),\]

as one can see by applying the log, and using \( \log(1 - a/k) \leq -a/k \) for \( k > a \). (Compare the proof of Proposition D.1.) Hence \( C_n(-1, 0) \to 0 \) as \( n \to \infty \), whenever \( a > 0 \) in (D.21). On the other hand,

\[(D.24) \quad C_n(1, 0) = \left| \frac{a(a - 1) \cdots (a - n)}{n!} \right| \sup_{0 < \xi < 1} (1 - \xi)^{a-1}.\]

If \( a \in (0, 1) \), this is identically +∞, while if \( a \geq 1 \) it is \( \mathcal{O}(n^{-a}) \), as above.

**Conclusion.** The estimate (D.20) is successful at establishing the convergence of the Taylor series (about the origin) for the function \( f(x) = (1 - x)^a \), at \( x = -1 \), whenever \( a > 0 \). It fails to establish the convergence at \( x = +1 \), when \( 0 < a < 1 \), but it is successful when \( a \geq 1 \).

We mention that convergence for \( x \in (-1, 1) \) is easily checked for the power series of the functions (D.21). The failure of (D.20) to establish convergence at \( x = +1 \) does not imply failure of such convergence. In fact we have the following result, which will be useful in Appendix E.
Proposition D.1. Given \( a > 0 \), the Taylor series about the origin for the function \( f(x) = (1 - x)^a \) converges absolutely and uniformly to \( f(x) \) on the closed interval \([-1, 1]\).

Proof. As noted, the series is

\[
\sum_{n=0}^{\infty} c_n x^n, \quad c_n = (-1)^n \frac{a(a-1) \cdots (a-n+1)}{n!},
\]
and, by an analysis parallel to (D.23),

\[
|c_n| \leq C n^{-a}.
\]

In more detail, if \( n - 1 > a \),

\[
c_n = -\frac{a}{n} \prod_{1 \leq k \leq a} \left(1 - \frac{a}{k}\right) \prod_{a < k \leq n-1} \left(1 - \frac{a}{k}\right),
\]

which we can write as \( c_n = (A/n)b_n \), where \( b_n \) denotes the last product in (D.26A). Then

\[
\log b_n \leq - \sum_{a < k \leq n-1} \frac{a}{k} \leq -a \log n + \beta,
\]

so

\[
b_n \leq e^{-a \log n + \beta} = \gamma n^{-a},
\]
giving (D.26).

Since the right side of (D.26) is summable (by the integral test) whenever \( a > 0 \), we see that the series (D.25) does converge absolutely and uniformly on \([-1, 1]\); so its limit is a continuous function \( f_a \) on \([-1, 1]\). The remark above has shown that \( f_a(x) = (1 - x)^a \) for \( x \in [-1, 1] \); by continuity this identity also holds at \( x = 1 \).

The material above has emphasized the study of the expansion (D.1) for infinitely smooth \( f \), concentrating on the issue of convergence as \( n \to \infty \). The behavior for fixed \( n \) (e.g., \( n = 2 \)) as \( x \to y \) is also of great interest, and in this connection it is important to note that (D.1) holds, with a useful formula for \( R_n(x, y) \), when \( f \) is merely \( C^n \), not necessarily \( C^{n+1} \). So suppose \( f \in C^n \), i.e., \( f, f', \ldots, f^{(n)} \) are continuous on an interval \( I \) about \( y \). Then the result (D.7) holds, with \( n \) replaced by \( n - 1 \); i.e., for \( x \in I \) we have

\[
f(x) = f(y) + f'(y)(x - y) + \cdots + \frac{1}{(n-1)!} f^{(n-1)}(y)(x - y)^{n-1} + R_{n-1}(x, y),
\]

with

\[
R_{n-1}(x, y) = \frac{1}{(n-1)!} \int_y^x (x - s)^{n-1} f^{(n)}(s) \, ds.
\]

Now we can add and subtract \( f^{(n)}(y) \) to the factor \( f^{(n)}(s) \) in the integrand above, and obtain

\[
R_{n-1}(x, y) = \frac{1}{n!} f^{(n)}(y)(x - y)^n + \frac{1}{(n-1)!} \int_y^x (x - s)^{n-1} \left[ f^{(n)}(s) - f^{(n)}(y) \right] \, ds.
\]

This establishes the following.
Proposition D.2. Assume \( f \) has \( n \) continuous derivatives on an interval \( I \) containing \( y \). Then, for \( x \in I \), the formula (D.1) holds, with

\[
R_n(x, y) = \frac{1}{(n-1)!} \int_{y}^{x} (x - s)^{n-1} \left[ f^{(n)}(s) - f^{(n)}(y) \right] ds.
\]

Note that since the integral in (D.30) equals \( x - y \) times the value of the integrand at some point \( s = \xi \) between \( x \) and \( y \), we can write a “Cauchy form” of the remainder (D.30) as

\[
R_n(x, y) = \frac{1}{(n-1)!} \left[ f^{(n)}(\xi) - f^{(n)}(y) \right] (x - \xi)^{n-1} (x - y).
\]

Alternatively, parallel to (D.8), we can write

\[
R_n(x, y) = \frac{(x - y)^n}{(n-1)!} \int_{0}^{1} \left[ f^{(n)}(sx + (1 - s)y) - f^{(n)}(y) \right] (1 - s)^{n-1} ds,
\]

and obtain a “Lagrange form”:

\[
R_n(x, y) = \frac{(x - y)^n}{n!} \left[ f^{(n)}(\zeta) - f^{(n)}(y) \right],
\]

for some \( \zeta \) between \( x \) and \( y \). Note that (D.33) also follows by replacing \( n \) by \( n - 1 \) in (D.3).
E. The Weierstrass theorem and the Stone-Weierstrass theorem

The following result of Weierstrass is a very useful tool in analysis.

**Theorem E.1.** Given a compact interval $I$, any continuous function $f$ on $I$ is a uniform limit of polynomials.

Otherwise stated, our goal is to prove that the space $C(I)$ of continuous (real valued) functions on $I$ is equal to $\overline{P}(I)$, the uniform closure in $C(I)$ of the space of polynomials. Our starting point will be the result that the power series for $(1 - x)$ converges uniformly on $[-1, 1]$, for any $a > 0$. This was established in §D, and we will use it, with $a = 1/2$.

From the identity $x^{1/2} = (1 - (1 - x))^{1/2}$, we have $x^{1/2} \in \overline{P}([0, 2])$. More to the point, from the identity

(E.1) $|x| = (1 - (1 - x^2))^{1/2}$,

we have $|x| \in \overline{P}([-\sqrt{2}, \sqrt{2}])$. Using $|x| = b^{-1}|bx|$, for any $b > 0$, we see that $|x| \in \overline{P}(I)$ for any interval $I = [-c, c]$, and also for any closed subinterval, hence for any compact interval $I$. By translation, we have

(E.2) $|x - a| \in \overline{P}(I)$

for any compact interval $I$. Using the identities

(E.3) $\max(x, y) = \frac{1}{2}(x + y) + \frac{1}{2}|x - y|, \quad \min(x, y) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y|,$

we see that for any $a \in \mathbb{R}$ and any compact $I$,

(E.4) $\max(x, a), \ \min(x, a) \in \overline{P}(I)$.

We next note that $\overline{P}(I)$ is an algebra of functions, i.e.,

(E.5) $f, g \in \overline{P}(I), \ c \in \mathbb{R} \implies f + g, fg, cf \in \overline{P}(I)$.

Using this, one sees that, given $f \in \overline{P}(I)$, with range in a compact interval $J$, one has $h \circ f \in \overline{P}(I)$ for all $h \in \overline{P}(J)$. Hence $f \in \overline{P}(I) \implies |f| \in \overline{P}(I)$, and, via (E.3), we deduce that

(E.6) $f, g \in \overline{P}(I) \implies \max(f, g), \ \min(f, g) \in \overline{P}(I)$.

Suppose now that $I' = [a', b']$ is a subinterval of $I = [a, b]$. With the notation $x_+ = \max(x, 0)$, we have

(E.7) $f_{II'}(x) = \min\left((x - a')_+, (b' - x)_+\right) \in \overline{P}(I)$.
This is a piecewise linear function, equal to zero off \( I \setminus I' \), with slope 1 from \( a' \) to the midpoint \( m' \) of \( I' \), and slope −1 from \( m' \) to \( b' \).

Now if \( I \) is divided into \( N \) equal subintervals, any continuous function on \( I \) that is linear on each such subinterval can be written as a linear combination of such “tent functions,” so it belongs to \( \mathcal{P}(I) \). Finally, any \( f \in C(I) \) can be uniformly approximated by such piecewise linear functions, so we have \( f \in \mathcal{P}(I) \), proving the theorem.

A far reaching extension of Theorem E.1, due to M. Stone, is the following result, known as the Stone-Weierstrass theorem.

**Theorem E.2.** Let \( X \) be a compact metric space, \( \mathcal{A} \) a subalgebra of \( C_\mathbb{R}(X) \), the algebra of real valued continuous functions on \( X \). Suppose \( 1 \in \mathcal{A} \) and that \( \mathcal{A} \) separates points of \( X \), i.e., for distinct \( p, q \in X \), there exists \( h_{pq} \in \mathcal{A} \) with \( h_{pq}(p) \neq h_{pq}(q) \). Then the closure \( \overline{\mathcal{A}} \) is equal to \( C_\mathbb{R}(X) \).

We present the proof in eight steps.

**Step 1.** By Theorem E.1, if \( f \in \overline{\mathcal{A}} \) and \( \varphi : \mathbb{R} \to \mathbb{R} \) is continuous, then \( \varphi \circ f \in \mathcal{A} \).

**Step 2.** Consequently, if \( f_j \in \mathcal{A} \), then

\[
\max(f_1, f_2) = \frac{1}{2}|f_1 - f_2| + \frac{1}{2}(f_1 + f_2) \in \overline{\mathcal{A}},
\]

and similarly \( \min(f_1, f_2) \in \overline{\mathcal{A}} \).

**Step 3.** It follows from the hypotheses that if \( p, q \in X \) and \( p \neq q \), then there exists \( f_{pq} \in \mathcal{A} \), equal to 1 at \( p \) and to 0 at \( q \).

**Step 4.** Apply an appropriate continuous \( \varphi : \mathbb{R} \to \mathbb{R} \) to get \( g_{pq} = \varphi \circ f_{pq} \in \overline{\mathcal{A}} \), equal to 1 on a neighborhood of \( p \) and to 0 on a neighborhood of \( q \), and satisfying \( 0 \leq g_{pq} \leq 1 \) on \( X \).

**Step 5.** Fix \( p \in X \) and let \( U \) be an open neighborhood of \( p \). By Step 4, given \( q \in X \setminus U \), there exists \( g_{pq} \in \overline{\mathcal{A}} \) such that \( g_{pq} = 1 \) on a neighborhood \( O_q \) of \( p \), equal to 0 on a neighborhood \( \Omega_q \) of \( q \), satisfying \( 0 \leq g_{pq} \leq 1 \) on \( X \).

Now \( \{\Omega_q\} \) is an open cover of \( X \setminus U \), so there exists a finite subcover \( \Omega_{q_1}, \ldots, \Omega_{q_N} \). Let

\[
 g_{pU} = \min_{1 \leq j \leq N} g_{pq_j} \in \overline{\mathcal{A}}.
\]

Then \( g_{pU} = 1 \) on \( O = \cap_{q=1}^N \Omega_q \), an open neighborhood of \( p \), \( g_{pU} = 0 \) on \( X \setminus U \), and \( 0 \leq g_{pU} \leq 1 \) on \( X \).

**Step 6.** Take \( K \subset U \subset X \), \( K \) closed, \( U \) open. By Step 5, for each \( p \in K \), there exists \( g_{pU} \in \overline{\mathcal{A}} \), equal to 1 on a neighborhood \( O_p \) of \( p \), and equal to 0 on \( X \setminus U \).
Now \( \{O_p\} \) covers \( K \), so there exists a finite subcover \( O_{p_1}, \ldots, O_{p_m} \). Let

\[
g_{KU} = \max_{1 \leq j \leq M} g_{p_j}U \in \mathcal{A}.
\]

We have

\[
g_{KU} = 1 \quad \text{on} \quad K, \quad 0 \quad \text{on} \quad X \setminus U, \quad \text{and} \quad 0 \leq g_{KU} \leq 1 \quad \text{on} \quad X.
\]

Step 7. Take \( f \in C_\mathbb{R}(X) \) such that \( 0 \leq f \leq 1 \) on \( X \). Fix \( k \in \mathbb{N} \) and set

\[
K_\ell = \left\{ x \in X : f(x) \geq \frac{\ell}{k} \right\},
\]

so \( X = K_0 \supset \cdots \supset K_\ell \supset K_{\ell+1} \supset \cdots K_k \supset K_{k+1} = \emptyset \). Define open \( U_\ell \supset K_\ell \) by

\[
U_\ell = \left\{ x \in X : f(x) > \frac{\ell - 1}{k} \right\}, \quad \text{so} \quad X \setminus U_\ell = \left\{ x \in X : f(x) \leq \frac{\ell - 1}{k} \right\}.
\]

By Step 6, there exist \( \psi_\ell \in \mathcal{A} \) such that

\[
\psi_\ell = 1 \quad \text{on} \quad K_\ell, \quad \psi_\ell = 0 \quad \text{on} \quad X \setminus U_\ell, \quad \text{and} \quad 0 \leq \psi_\ell \leq 1 \quad \text{on} \quad X.
\]

Let

\[
f_\ell = \max_{0 \leq \ell \leq k} \frac{\ell}{k} \psi_\ell \in \mathcal{A}.
\]

It follows that \( f_\ell \geq \ell/k \) on \( K_\ell \) and \( f_\ell \leq (\ell - 1)/k \) on \( X \setminus U_\ell \), for all \( \ell \). Hence \( f_\ell \geq (\ell - 1)/k \) on \( K_{\ell-1} \) and \( f_\ell \leq \ell/k \) on \( U_{\ell+1} \). In other words,

\[
\frac{\ell - 1}{k} \leq f(x) \leq \frac{\ell}{k} \implies \frac{\ell - 1}{k} \leq f_\ell(x) \leq \frac{\ell}{k},
\]

so

\[
|f(x) - f_\ell(x)| \leq \frac{1}{k}, \quad \forall x \in X.
\]

Step 8. It follows from Step 7 that if \( f \in C_\mathbb{R}(X) \) and \( 0 \leq f \leq 1 \) on \( X \), then \( f \in \mathcal{A} \). It is an easy final step to see that \( f \in C_\mathbb{R}(X) \Rightarrow f \in \mathcal{A} \).

Theorem E.2 has a complex analogue. In that case, we add the assumption that \( f \in \mathcal{A} \Rightarrow f \in \mathcal{A} \), and conclude that \( \mathcal{A} = \mathcal{C}(X) \). This is easily reduced to the real case.

Here are a couple of applications of Theorem E.2, in its real and complex forms:

**Corollary E.3.** If \( X \) is a compact subset of \( \mathbb{R}^n \), then every \( f \in \mathcal{C}(X) \) is a uniform limit of polynomials on \( \mathbb{R}^n \).

**Corollary E.4.** The space of trigonometric polynomials on the \( n \)-torus \( \mathbb{T}^n \), given by

\[
\sum_{|k| \leq N} a_k e^{ik \cdot \theta},
\]

is dense in \( \mathcal{C}(\mathbb{T}^n) \).
F. Sard’s theorem

Let $F : \mathcal{O} \to \mathbb{R}^n$ be a $C^1$ map, with $\mathcal{O}$ open in $\mathbb{R}^n$. If $p \in \mathcal{O}$ and $DF(p) : \mathbb{R}^n \to \mathbb{R}^n$ is not surjective, then $p$ is said to be a critical point, and $F(p)$ a critical value. The set $C$ of critical points can be a large subset of $\mathcal{O}$, even all of it, but the set of critical values $F(C)$ must be small in $\mathbb{R}^n$, as the following result implies.

**Proposition F.1.** If $F : \mathcal{O} \to \mathbb{R}^n$ is a $C^1$ map, $C \subset \mathcal{O}$ its set of critical points, and $K \subset \mathcal{O}$ compact, then $F(C \cap K)$ is a nil subset of $\mathbb{R}^n$.

**Proof.** Without loss of generality, we can assume $K$ is a cubical cell. Let $\mathcal{P}$ be a partition of $K$ into cubical cells $R_\alpha$, all of diameter $\delta$. Write $\mathcal{P} = \mathcal{P}' \cup \mathcal{P}''$, where cells in $\mathcal{P}'$ are disjoint from $C$, and cells in $\mathcal{P}''$ intersect $C$. Pick $x_\alpha \in R_\alpha \cap C$, for $R_\alpha \in \mathcal{P}''$.

Fix $\varepsilon > 0$. Now we have

$$F(x_\alpha + y) = F(x_\alpha) + DF(x_\alpha) + r_\alpha(y),$$

and, if $\delta > 0$ is small enough, then $|r_\alpha(y)| \leq \varepsilon |y| \leq \varepsilon \delta$, for $x_\alpha + y \in R_\alpha$. Thus $F(R_\alpha)$ is contained in an $\varepsilon \delta$-neighborhood of the set $H_\alpha = F(x_\alpha) + DF(x_\alpha)(R_\alpha - x_\alpha)$, which is a parallelepiped of dimension $\leq n - 1$, and diameter $\leq M \delta$, if $|DF| \leq M$. Hence

$$\text{cont}^+ F(R_\alpha) \leq C \varepsilon \delta^n \leq C'' \varepsilon V(R_\alpha), \quad \text{for } R_\alpha \in \mathcal{P}''.$$

Thus

$$\text{cont}^+ F(C \cap K) \leq \sum_{R_\alpha \in \mathcal{P}''} \text{cont}^+ F(R_\alpha) \leq C'' \varepsilon.$$

Taking $\varepsilon \to 0$, we have the proof.

This is the easy case of a result known as Sard’s Theorem, which also treats the case $F : \mathcal{O} \to \mathbb{R}^n$ when $\mathcal{O}$ is an open set in $\mathbb{R}^m$, $m > n$. Then a more elaborate argument is needed, and one requires more differentiability, namely that $F$ is class $C^k$, with $k = m - n + 1$. A proof can be found in [Mil] or [Stb].
G. Morse functions

If $\Omega \subset \mathbb{R}^n$ is open, a $C^2$ function $f : \Omega \to \mathbb{R}$ is said to be a Morse function if each critical point of $f$ is nondegenerate, i.e.,

$$\forall p \in \Omega, \quad \nabla f(p) = 0 \Rightarrow D^2 f(p) \text{ is invertible,}$$

where $D^2 f(p)$ is the symmetric $n \times n$ matrix of second order partial derivatives defined in §1. More generally, if $M$ is an $n$-dimensional surface, a $C^2$ function $f : M \to \mathbb{R}$ is said to be a Morse function if $f \circ \varphi$ is a Morse function on $\Omega$ for each coordinate patch $\varphi : \Omega \to U \subset M$.

Our goal here is to establish the existence of lots of Morse functions on an $n$-dimensional surface $M$. For simplicity, we restrict attention to the case where $M$ is compact. Here is our main result.

**Theorem G.1.** Let $M \subset \mathbb{R}^N$ be a compact, smooth, $n$-dimensional surface. For $a \in \mathbb{R}^N$, set

$$(G.2) \quad \varphi_a : M \to \mathbb{R}, \quad \varphi_a(x) = a \cdot x, \quad x \in M.$$  

Take $f \in C^2(M)$. Then the set $\mathcal{O}_f$ of $a \in \mathbb{R}^N$ such that

$$(G.3) \quad f + \varphi_a : M \to \mathbb{R} \text{ is a Morse function}$$

is a dense open subset of $\mathbb{R}^N$.

It is easy to verify that $\mathcal{O}_f$ is open, since when (J.1) holds, a small $C^2$ perturbation $g$ of $f$ has the property that $D^2 g(x)$ is invertible for $x$ near $p$. What is not so easy is to show that $\mathcal{O}_f$ is dense (or even nonempty!). Our proof of such denseness will make use of Sard’s theorem, from Appendix I. We begin with an easy special case.

**Proposition G.2.** In the setting of Theorem J.1, assume $N = n + 1$ and $M = \partial \Omega$, with $\Omega \subset \mathbb{R}^{n+1}$ open. Then

$$(G.4) \quad \{a \in S^n : a \notin \mathcal{O}_0\} \text{ is a nil set,}$$

hence has empty interior in the unit sphere $S^n$.

**Proof.** Here we are examining when $\varphi_a$ is a Morse function on $M$. Let $N : M \to S^n$ be the exterior unit normal. Then $x_0 \in M$ is a critical point of $\varphi_a$ if and only if $N(x_0) = \pm a$. Such a point $x_0$ is a nondegenerate critical point of $\varphi_a$ if and only if it is not a critical point of $N$. Hence, if $\pm a$ are regular values of $N$, then $\varphi_a$ is a Morse function, i.e., $a \in \mathcal{O}_0$. By Sard’s theorem, the set of points in $S^n$ that are critical values of $N$ is a nil set, so the proof of Proposition J.2 is done.

We begin to tackle Theorem G.1 with the following result.
**Lemma G.3.** Let \( \Omega \subset \mathbb{R}^n \) be open, and take \( g \in C^2(\Omega) \). Let \( \overline{U} \subset \Omega \) be the closure of a smoothly bounded open \( U \). Set \( g_a(x) = g(x) + a \cdot x \). Let \( \mathcal{O}_g \) denote the set of \( a \in \mathbb{R}^n \) such that \( g_a|_\overline{U} \) has only nondegenerate critical points. Then \( \mathbb{R}^n \setminus \mathcal{O}_g \) is a nil set.

**Proof.** Consider

\[
F(x) = -\nabla g(x), \quad F : \Omega \to \mathbb{R}^n.
\]

A point \( x \in \Omega \) is a critical point of \( g_a \) if and only if \( F(x) = a \), and this critical point is degenerate only if, in addition, \( a \) is a critical value of \( F \). Hence the desired conclusion holds for all \( a \in \mathbb{R}^n \) that are not critical values of \( F|_\overline{U} \). Again Sard’s theorem applies.

**Proof of Theorem G.1.** Each \( p \in M \) has a neighborhood \( U_p \) in \( M \) such that \( \overline{U}_p \subset \Omega_p \subset M \) and some \( n \) of the coordinates \( x_j \) on \( \mathbb{R}^N \) produce coordinates on \( \Omega_p \). Say \( x_1, \ldots, x_n \) do it. Let \( (a_{n+1}, \ldots, a_N) \) be fixed, but arbitrary. Then Lemma J.3 can be applied to \( g = f + \sum_{n+1}^N a_j x_j \), treated as a function of \( (x_1, \ldots, x_n) \). It follows that, for all \( (a_1, \ldots, a_n) \) but a nil set, \( f + \varphi_a \) has only nondegenerate critical points in \( \overline{U}_p \). Thus

\[
\{a \in \mathbb{R}^N : f + \varphi_a \text{ has only nondegenerate critical points in } \overline{U}_p \}
\]

is dense in \( \mathbb{R}^N \). We also know this set is open. Now \( M \) can be covered by a finite collection of such sets \( U_p \), so \( \mathcal{O}_f \), defined in Theorem G.1, is a finite intersection of open dense subsets of \( \mathbb{R}^N \), hence it is open and dense, as asserted.
H. Inner product spaces

On occasion, particularly in Appendix G, we have looked at norms and inner products on spaces of functions, such as $C(S^1)$ and $R(\mathbb{R})$, which are vector spaces. Generally, a complex vector space $V$ is a set on which there are operations of vector addition:

\begin{equation}
(H.1) \quad f, g \in V \implies f + g \in V,
\end{equation}

and multiplication by an element of $\mathbb{C}$ (called scalar multiplication):

\begin{equation}
(H.2) \quad a \in \mathbb{C}, \ f \in V \implies af \in V,
\end{equation}

satisfying the following properties. For vector addition, we have

\begin{equation}
(H.3) \quad f + g = g + f, \ (f + g) + h = f + (g + h), \ f + 0 = f, \ f + (-f) = 0.
\end{equation}

For multiplication by scalars, we have

\begin{equation}
(H.4) \quad a(bf) = (ab)f, \quad 1 \cdot f = f.
\end{equation}

Furthermore, we have two distributive laws:

\begin{equation}
(H.5) \quad a(f + g) = af + ag, \quad (a + b)f = af + bf.
\end{equation}

These properties are readily verified for the function spaces mentioned above.

An inner product on a complex vector space $V$ assigns to elements $f, g \in V$ the quantity $(f, g) \in \mathbb{C}$, in a fashion that obeys the following three rules:

\begin{equation}
(H.6) \quad (a_1 f_1 + a_2 f_2, g) = a_1 (f_1, g) + a_2 (f_2, g),
\end{equation}

\begin{equation}
(f, g) = \overline{(g, f)},
\end{equation}

\begin{equation}
(f, f) > 0 \quad \text{unless} \quad f = 0.
\end{equation}

A vector space equipped with an inner product is called an inner product space. For example,

\begin{equation}
(H.7) \quad (f, g) = \frac{1}{2\pi} \int_{S^1} f(\theta)\overline{g(\theta)} \, d\theta
\end{equation}

defines an inner product on $C(S^1)$, and also on $R(S^1)$, where we identify two functions that differ only on a set of upper content zero. Similarly,

\begin{equation}
(H.8) \quad (f, g) = \int_{-\infty}^{\infty} f(x)\overline{g(x)} \, dx
\end{equation}
defines an inner product on $\mathcal{R}(\mathbb{R})$ (where, again, we identify two functions that differ only on a set of upper content zero).

As another example, in we define $\ell^2$ to consist of sequences $(a_k)_{k \in \mathbb{Z}}$ such that

\[(H.9)\quad \sum_{k=-\infty}^{\infty} |a_k|^2 < \infty.\]

An inner product on $\ell^2$ is given by

\[(H.10)\quad \langle (a_k), (b_k) \rangle = \sum_{k=-\infty}^{\infty} a_k \overline{b_k}.\]

Given an inner product on $V$, one says the object $\|f\|$ defined by

\[(H.11)\quad \|f\| = \sqrt{\langle f, f \rangle}\]

is the norm on $V$ associated with the inner product. Generally, a norm on $V$ is a function $f \mapsto \|f\|$ satisfying

\[(H.12)\quad \|af\| = |a| \cdot \|f\|, \quad a \in \mathbb{C}, \; f \in V;\]

\[(H.13)\quad \|f\| > 0 \quad \text{unless} \; f = 0;\]

\[(H.14)\quad \|f + g\| \leq \|f\| + \|g\|.\]

The property (H.14) is called the triangle inequality. A vector space equipped with a norm is called a normed vector space. We can define a distance function on such a space by

\[(H.15)\quad d(f, g) = \|f - g\|,\]

Properties (H.12)–(H.14) imply that $d : V \times V \to [0, \infty)$ satisfies the properties in (A.1), making $V$ a metric space.

If $\|f\|$ is given by (H.11), from an inner product satisfying (H.6), it is clear that (H.12)–(H.13) hold, but (H.14) requires a demonstration. Note that

\[(H.16)\quad \|f + g\|^2 = \langle f + g, f + g \rangle = \|f\|^2 + \langle f, g \rangle + \langle g, f \rangle + \|g\|^2 = \|f\|^2 + 2 \operatorname{Re}(f, g) + \|g\|^2,\]

while

\[(H.17)\quad (\|f\| + \|g\|)^2 = \|f\|^2 + 2 \|f\| \cdot \|g\| + \|g\|^2.\]

Thus to establish (H.17) it suffices to prove the following, known as Cauchy’s inequality.
**Proposition H.1.** For any inner product on a vector space \( V \), with \( \| \cdot \| \) defined by (H.11),

\[
(f, g) \leq \| f \| \cdot \| g \|, \quad \forall f, g \in V.
\]

**Proof.** We start with

\[
0 \leq \| f - g \|^2 = \| f \|^2 - 2 \text{Re}(f, g) + \| g \|^2,
\]

which implies

\[
2 \text{Re}(f, g) \leq \| f \|^2 + \| g \|^2, \quad \forall f, g \in V.
\]

Replacing \( f \) by \( af \) for arbitrary \( a \in \mathbb{C} \) of absolute value 1 yields

\[
2 \text{Re}(f, g) \leq \| f \|^2 + \| g \|^2, \quad \forall f, g \in V.
\]

Replacing \( f \) by \( tf \) and \( g \) by \( t^{-1}g \) for arbitrary \( t \in (0, \infty) \), we have

\[
2|f, g| \leq t^2 \| f \|^2 + t^{-2} \| g \|^2, \quad \forall f, g \in V, \ t \in (0, \infty).
\]

If we take \( t^2 = \| g \|/\| f \| \), we obtain the desired inequality (H.18). This assumes \( f \) and \( g \) are both nonzero, but (H.18) is trivial if \( f \) or \( g \) is 0.

An inner product space \( V \) is called a Hilbert space if it is a complete metric space, i.e., if every Cauchy sequence \( (f_n) \) in \( V \) has a limit in \( V \). The space \( l^2 \) has this completeness property, but \( C(S^1) \), with inner product (H.7), does not, nor does \( R(S^1) \). Appendix A describes a process of constructing the completion of a metric space. When applied to an incomplete inner product space, it produces a Hilbert space. When this process is applied to \( C(S^1) \), the completion is the space \( L^2(S^1) \). An alternative construction of \( L^2(S^1) \) uses the Lebesgue integral. For this approach, one can consult Chapter 4 of [T2].
References


Index

A

algebra
  of sets 79
analytic function
  complex 35
  real 55
anticommutation relation 147
Arc length 110
area
  of spheres 112
arcsin 72, 173
arctan 73
Ascoli’s theorem 250
autonomous system 61
averaging over rotations 114

B

basis 39, 109
Brouwer fixed-point theorem 186
boundary 83

C

Cauchy sequence 47, 179
Cauchy-Riemann equations 35, 55, 165, 168
Cauchy’s integral formula 165, 170
Cauchy’s integral theorem 165
chain rule 18, 26, 109, 140, 147, 171
change of variable formula 82, 85, 88, 95, 109, 137, 259
characteristic function 10, 78
closed set 245
closure 77, 246
cofactor 41
compact space 246
complete metric space 47, 245
completion of a metric space 245
conformal map 170
connected set 254
content 10, 78
continuous 6, 57
    uniformly 7, 250
Contraction Mapping Principle 47, 58
convolution 220
coordinate chart 108
corner 151
cos 71, 131, 171
Cramer’s formula 37
cross product 41, 108
curl 148, 158, 159
cusp 152

D

Darboux’s Theorem 8, 135
degree of a map 188
derivation 67, 157
derivative 11, 23
determinant 32, 36, 38, 119
diagonal construction 250
diffeomorphism 45, 56, 86, 108
differential equation 58
differential form 137
    closed 147
    exact 147
distribution 211, 212
div 146, 154, 155
Divergence Theorem 11, 121, 156, 159

E

Euclidean norm 23
Euler characteristic 194
Euler’s formula 71, 132
exponential 70, 171
    of a matrix 49, 74, 113
exterior derivative 146

F

fixed point 47, 179
flow 66
Fourier inversion formula 199, 209, 218, 220, 224
Fourier series 199
Fourier transform 218
Fubini’s Theorem 81, 105, 168
Fundamental Theorem of Algebra 168, 193
Fundamental Theorem of Calculus 11, 12, 18, 23, 150, 161

G
Gamma function 112
Gauss curvature 195
Gauss-Bonnet formula 195
Gauss linking number formula 198
Gaussian integral 79
$Gl(n, \mathbb{R})$ 24, 82, 104
$Gl_+(n, \mathbb{R})$ 41
grad 137
Green’s formula 157
Green’s Theorem 154, 160, 162

H
Haar measure 113
harmonic conjugate 169
harmonic function 168, 171
Hessian 29
holomorphic function 35, 55, 165
homotopy 179

I
Implicit Function Theorem 51, 116
index of a vector field 191
initial value problem 58
inner product space 275
integral curve 66
integral equation 58
integrating factor 148
integration by parts 18, 113
interior 77, 246
interior product 145
Intermediate Value Theorem 254
Inverse Function Theorem 45, 108, 112
iterated integral 78
Jordan-Brouwer separation theorem 191
Jordan curve theorem 191

Laplace operator 156
Lebesgue integral 227
Lie bracket 68
Lie derivative 67, 181
Liouville’s Theorem 166
Lipschitz 11, 58, 81, 105
log 70, 172

manifold 124
matrix 23, 30, 104
maximum 13, 31, 249
mean value property 168
Mean Value Theorem 12, 249
metric space 242
metric tensor 109, 113
minimum 13, 31, 249
modulus of continuity 250
monotone convergence theorem 97
Morse function 273
multi-index 27, 28

neighborhood 245
Newton’s method 49, 55
nil set 79, 80
normal derivative 152
normal vector
   outward 154, 155, 160
   positive 158, 159, 160

open set 245
orbit 66
orientation 139, 150, 152, 155
outer measure 101

<table>
<thead>
<tr>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>partial derivative 14, 23</td>
</tr>
<tr>
<td>partition 6, 76, 89</td>
</tr>
<tr>
<td>of unity 146, 257</td>
</tr>
<tr>
<td>permutation 39</td>
</tr>
<tr>
<td>sign of 39</td>
</tr>
<tr>
<td>pi ($\pi$) 72, 73, 112</td>
</tr>
<tr>
<td>Picard iteration 58</td>
</tr>
<tr>
<td>Plancherel identity 224</td>
</tr>
<tr>
<td>$\mathcal{P}(n)$ 127</td>
</tr>
<tr>
<td>Poincaré lemma 147, 152, 180</td>
</tr>
<tr>
<td>Poisson summation formula 236</td>
</tr>
<tr>
<td>polar coordinates 84</td>
</tr>
<tr>
<td>polar decomposition 127</td>
</tr>
<tr>
<td>positive-definite matrix 31, 45, 109</td>
</tr>
<tr>
<td>power series 36, 55, 166, 170, 263</td>
</tr>
<tr>
<td>pull-back 137, 139, 146</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>rational numbers 9, 245</td>
</tr>
<tr>
<td>real numbers 245</td>
</tr>
<tr>
<td>Riemann functional equation 237</td>
</tr>
<tr>
<td>Riemann integral 6, 76</td>
</tr>
<tr>
<td>Riemann zeta function 236</td>
</tr>
<tr>
<td>row reduction 107</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saddle point 31</td>
</tr>
<tr>
<td>Sard’s theorem 272</td>
</tr>
<tr>
<td>sin 71, 131, 171</td>
</tr>
<tr>
<td>Skew($n$) 109</td>
</tr>
<tr>
<td>$SO(n)$ 41, 113</td>
</tr>
<tr>
<td>sphere</td>
</tr>
<tr>
<td>area of 112</td>
</tr>
<tr>
<td>spherical polar coordinates 89, 111</td>
</tr>
<tr>
<td>Stokes formula 149, 159, 160, 179</td>
</tr>
<tr>
<td>Stone-Weierstrass theorem 166, 270</td>
</tr>
</tbody>
</table>
submersion 116
surface 108
surface integral 110, 113, 115
Sym(n) 127

\begin{itemize}
    \item \textbf{T} \hspace{1cm} \tan 73
    \item tangent space 107
    \item tangent vector
        \begin{itemize}
            \item forward 158, 160
        \end{itemize}
    \item Taylor’s formula with remainder 19, 28, 264
    \item tempered distribution 235
    \item triangle inequality 235
    \item trigonometric functions 71
    \item Tychonov’s theorem 252
\end{itemize}

\begin{itemize}
    \item \textbf{V} \hspace{1cm} \textbf{vector field} 66, 120, 138, 150, 155, 158, 159
    \item volume 76, 111, 114
    \item volume form 154, 161
\end{itemize}

\begin{itemize}
    \item \textbf{W} \hspace{1cm} \textbf{wedge product} 145
    \item Weierstrass theorem 269
\end{itemize}