Basic Analysis

Kenneth Kuttler

April 16, 2001

Contents

Basic Set theory

We think of a set as a collection of things called elements of the set. For example, we may consider the set of integers, the collection of signed whole numbers such as 1,2,-4, etc. This set which we will believe in is denoted by $\mathbb{Z}$. Other sets could be the set of people in a family or the set of donuts in a display case at the store. Sometimes we use parentheses, $\{\ \}$ to specify a set. When we do this, we list the things which are in the set between the parentheses. For example the set of integers between -1 and 2, including these numbers could be denoted as $\{-1, 0, 1, 2\}$. We say x is an element of a set S, and write $x \in S$ if x is one of the things in S. Thus, $1 \in \{-1, 0, 1, 2, 3\}$. Here are some axioms about sets. Axioms are statements we will agree to believe.

1. Two sets are equal if and only if they have the same elements.
2. To every set, A, and to every condition $S(x)$ there corresponds a set, B, whose elements are exactly those elements x of A for which $S(x)$ holds.
3. For every collection of sets there exists a set that contains all the elements that belong to at least one set of the given collection.
4. The Cartesian product of a nonempty family of nonempty sets is nonempty.
5. If A is a set there exists a set, $\mathcal{P}(A)$ such that $\mathcal{P}(A)$ is the set of all subsets of A.

These axioms are referred to as the axiom of extension, axiom of specification, axiom of unions, axiom of choice, and axiom of powers respectively.

It seems fairly clear we should want to believe in the axiom of extension. It is merely saying, for example, that $\{1, 2, 3\} = \{2, 3, 1\}$ since these two sets have the same elements in them. Similarly, it would seem we would want to specify a new set from a given set using some "condition" which can be used as a test to determine whether the element in question is in the set. For example, we could consider the set of all integers which are multiples of 2. This set could be specified as follows.

$$\{x \in \mathbb{Z} : x = 2y \text{ for some } y \in \mathbb{Z}\}.$$

In this notation, the colon is read as "such that" and in this case the condition is being a multiple of 2. Of course, there could be questions about what constitutes a "condition". Just because something is grammatically correct does not mean it makes any sense. For example consider the following nonsense.

$$S = \{x \in \text{set of dogs} \ : \ \text{it is colder in the mountains than in the winter}\}.$$

We will leave these sorts of considerations however and assume our conditions make sense. The axiom of unions states that if we have any collection of sets, there is a set consisting of all the elements in each of the sets in the collection. Of course this is also open to further consideration. What is a collection? Maybe it would be better to say "set of sets" or, given a set whose elements are sets there exists a set whose elements consist of exactly those things which are elements of at least one of these sets. If $\mathcal{S}$ is such a set whose elements are sets, we write the union of all these sets in the following way.

$$\cup \{A : A \in \mathcal{S}\}$$

or sometimes as

$$\cup \mathcal{S}.$$

Something is in the Cartesian product of a set or "family" of sets if it consists of a single thing taken from each set in the family. Thus $(1, 2, 3) \in \{1, 4, .2\} \times \{1, 2, 7\} \times \{4, 3, 7, 9\}$ because it consists of exactly one element from each of the sets which are separated by $\times$. Also, this is the notation for the Cartesian product of finitely many sets. If $\mathcal{S}$ is a set whose elements are sets, we could write

$$\prod_{A \in \mathcal{S}} A$$

for the Cartesian product. We can think of the Cartesian product as the set of choice functions, a choice function being a function which selects exactly one element of each set of $\mathcal{S}$. You may think the axiom of choice, stating that the Cartesian product of a nonempty family of nonempty sets is nonempty, is innocuous but there was a time when many mathematicians were ready to throw it out because it implies things which are very hard to believe.

We say A is a subset of B and write $A \subseteq B$ if every element of A is also an element of B. This can also be written as $B \supseteq A$. We say A is a proper subset of B and write $A \subset B$ or $B \supset A$ if A is a subset of B but A is not equal to $B, A \neq B$. The intersection of two sets is a set denoted as $A \cap B$ and it means the set of elements of A which are also elements of B. The axiom of specification shows this is a set. The empty set is the set which has no elements in it, denoted as $\emptyset$. The union of two sets is denoted as $A \cup B$ and it means the set of all elements which are in either of the sets. We know this is a set by the axiom of unions.

The complement of a set, (the set of things which are not in the given set) must be taken with respect to a given set called the universal set which is a set which contains the one whose complement is being taken. Thus, if we want to take the complement of a set A, we can say its complement, denoted as A^C (or more precisely as $X \setminus A$) is a set by using the axiom of specification to write

$$A^C \equiv \{x \in X : x \notin A\}$$

The symbol $\notin$ is read as "is not an element of". Note the axiom of specification takes place relative to a given set which we believe exists. Without this universal set we cannot use the axiom of specification to speak of the complement.

Words such as "all" or "there exists" are called quantifiers and they must be understood relative to some given set. Thus we can speak of the set of all integers larger than 3. Or we can say there exists an integer larger than 7. Such statements have to do with a given set, in this case the integers. Failure to have a reference set when quantifiers are used turns out to be illogical even though such usage may be grammatically correct. Quantifiers are used often enough that there are symbols for them. The symbol $\forall$ is read as "for all" or "for every" and the symbol $\exists$ is read as "there exists". Thus $\forall\forall\exists\exists$ could mean for every upside down A there exists a backwards E.

1.1 Exercises

1. There is no set of all sets. This was not always known and was pointed out by Bertrand Russell. Here is what he observed. Suppose there were. Then we could use the axiom of specification to consider the set of all sets which are not elements of themselves. Denoting this set by S, determine whether S is an element of itself. Either it is or it isn't. Show there is a contradiction either way. This is known as Russell's paradox.

2. The above problem shows there is no universal set. Comment on the statement "Nothing contains everything." What does this show about the precision of standard English?

3. Do you believe each person who has ever lived on this earth has the right to do whatever he or she wants? (Note the use of the universal quantifier with no set in sight.) If you believe this, do you really believe what you say you believe? What of those people who want to deprive others their right to do what they want? Do people often use quantifiers this way?

4. DeMorgan's laws are very useful in mathematics. Let $\mathcal{S}$ be a set of sets each of which is contained in some universal set, U. Show

$$\cup\left\{A^C : A \in \mathcal{S}\right\} = \left(\cap\{A : A \in \mathcal{S}\}\right)^C$$

and

$$\cap\left\{A^C : A \in \mathcal{S}\right\} = \left(\cup\{A : A \in \mathcal{S}\}\right)^C.$$

5. Let $\mathcal{S}$ be a set of sets show

$$B \cup \cup\{A : A \in \mathcal{S}\} = \cup\{B \cup A : A \in \mathcal{S}\}.$$

6. Let $\mathcal{S}$ be a set of sets show

$$B \cap \cup\{A : A \in \mathcal{S}\} = \cup\{B \cap A : A \in \mathcal{S}\}.$$

1.2 The Schroder Bernstein theorem

It is very important to be able to compare the size of sets in a rational way. The most useful theorem in this context is the Schroder Bernstein theorem which is the main result to be presented in this section. To aid in this endeavor and because it is important for its own sake, we give the following definition.

Definition 1.1 *Let X and Y be sets.*

$$X \times Y \equiv \{(x, y) : x \in X \text{ and } y \in Y\}$$

A relation is defined to be a subset of $X \times Y$. A function, f is a relation which has the property that if (x, y) and (x, y_1) are both elements of the f, then $y = y_1$. The domain of f is defined as

$$D(f) \equiv \{x : (x, y) \in f\}.$$

and we write $f : D(f) \to Y$.

It is probably safe to say that most people do not think of functions as a type of relation which is a subset of the Cartesian product of two sets. A function is a mapping, sort of a machine which takes inputs, x and makes them into a unique output, $f(x)$. Of course, that is what the above definition says with more precision. An ordered pair, (x, y) which is an element of the function has an input, x and a unique output, y which we denote as $f(x)$ while the name of the function is f.

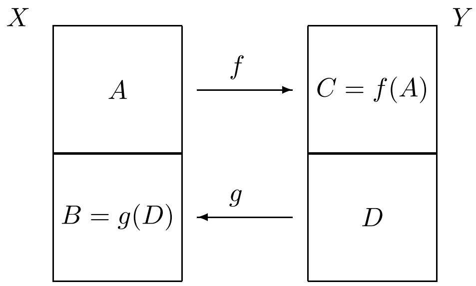
The following theorem which is interesting for its own sake will be used to prove the Schroder Bernstein theorem.

Theorem 1.2 *Let $f : X \to Y$ and $g : Y \to X$ be two mappings. Then there exist sets A, B, C, D, such that*

$$A \cup B = X,\ C \cup D = Y,\ A \cap B = \emptyset,\ C \cap D = \emptyset,$$

$$f(A) = C,\ g(D) = B.$$

The following picture illustrates the conclusion of this theorem.



Proof: We will say $A_0 \subseteq X$ satisfies $\mathcal{P}$ if whenever $y \in Y \setminus f(A_0)$, $g(y) \notin A_0$. Note $\emptyset$ satisfies $\mathcal{P}$.

$$\mathcal{A} \equiv \{A_0 \subseteq X : A_0 \text{ satisfies } \mathcal{P}\}.$$

Let $A = \cup\mathcal{A}$. If $y \in Y \setminus f(A)$, then for each $A_0 \in \mathcal{A}$, $y \in Y \setminus f(A_0)$ and so $g(y) \notin A_0$. Since $g(y) \notin A_0$ for all $A_0 \in \mathcal{A}$, it follows $g(y) \notin A$. Hence A satisfies $\mathcal{P}$ and is the largest subset of X which does so. Define

$$C \equiv f(A),\ D \equiv Y \setminus C,\ B \equiv X \setminus A.$$

Thus all conditions of the theorem are satisfied except for $g(D) = B$ and we verify this condition now.

Suppose $x \in B = X \setminus A$. Then $A \cup \{x\}$ does not satisfy $\mathcal{P}$ because this set is larger than A.Therefore there exists

$$y \in Y \setminus f(A \cup \{x\}) \subseteq Y \setminus f(A) \equiv D$$

such that $g(y) \in A \cup \{x\}$. But $g(y) \notin A$ because $y \in Y \setminus f(A)$ and A satisfies $\mathcal{P}$. Hence $g(y) = x$ and this proves the theorem.

Theorem 1.3 *(Schroder Bernstein) If $f : X \to Y$ and $g : Y \to X$ are one to one, then there exists $h : X \to Y$ which is one to one and onto.*

Proof: Let A, B, C, D be the sets of Theorem1.2 and define

$$h(x) \equiv \begin{cases} f(x) & \text{if } x \in A \\ g^{-1}(x) & \text{if } x \in B \end{cases}$$

It is clear h is one to one and onto.

Recall that the Cartesian product may be considered as the collection of choice functions. We give a more precise description next.

Definition 1.4 *Let I be a set and let X_i be a set for each $i \in I$. We say that f is a choice function and write*

$$f \in \prod_{i \in I} X_i$$

if $f(i) \in X_i$ for each $i \in I$.

The axiom of choice says that if $X_i \neq \emptyset$ for each $i \in I$, for I a set, then

$$\prod_{i \in I} X_i \neq \emptyset.$$

The symbol above denotes the collection of all choice functions. Using the axiom of choice, we can obtain the following interesting corollary to the Schroder Bernstein theorem.

Corollary 1.5 *If $f : X \to Y$ is onto and $g : Y \to X$ is onto, then there exists $h : X \to Y$ which is one to one and onto.*

Proof: For each $y \in Y$, let

$$f_0^{-1}(y) \in f^{-1}(y) \equiv \{x \in X : f(x) = y\}$$

and similarly let $g_0^{-1}(x) \in g^{-1}(x)$. We used the axiom of choice to pick a single element, $f_0^{-1}(y)$ in $f^{-1}(y)$ and similarly for $g^{-1}(x)$. Then f_0^{-1} and g_0^{-1} are one to one so by the Schroder Bernstein theorem, there exists $h : X \to Y$ which is one to one and onto.

Definition 1.6 *We say a set S, is finite if there exists a natural number n and a map θ which maps $\{1, \cdots, n\}$ one to one and onto S. We say S is infinite if it is not finite. A set S, is called countable if there exists a map θ mapping $\mathbb{N}$ one to one and onto S.(When θ maps a set A to a set B, we will write $\theta : A \to B$ in the future.) Here $\mathbb{N} \equiv \{1, 2, \cdots\}$, the natural numbers. If there exists a map $\theta : \mathbb{N} \to S$ which is onto, we say that S is at most countable.*

In the literature, the property of being at most countable is often referred to as being countable. When this is done, there is usually no harm incurred from this sloppiness because the question of interest is normally whether one can list all elements of the set, designating a first, second, third etc. in such a way as to exhaust the entire set. The possibility that a single element of the set may occur more than once in the list is often not important.

Theorem 1.7 *If X and Y are both at most countable, then $X \times Y$ is also at most countable.*

Proof: We know there exists a mapping $\eta : \mathbb{N} \to X$ which is onto. If we define $\eta(i) \equiv x_i$ we may consider X as the sequence $\{x_i\}_{i=1}^{\infty}$, written in the traditional way. Similarly, we may consider Y as the sequence $\{y_j\}_{j=1}^{\infty}$. It follows we can represent all elements of $X \times Y$ by the following infinite rectangular array.

$$\begin{array}{cccc}
(x_1, y_1) & (x_1, y_2) & (x_1, y_3) & \cdots \\
(x_2, y_1) & (x_2, y_2) & (x_2, y_3) & \cdots \\
(x_3, y_1) & (x_3, y_2) & (x_3, y_3) & \cdots \\
\vdots & \vdots & \vdots &
\end{array}.$$

We follow a path through this array as follows.

$$\begin{array}{ccccccc}
(x_1, y_1) & \to & (x_1, y_2) & & (x_1, y_3) & \to \\
& \swarrow & & \nearrow & & \\
(x_2, y_1) & & (x_2, y_2) & & & \\
\downarrow & \nearrow & & & & \\
(x_3, y_1) & & & & &
\end{array}$$

Thus the first element of $X \times Y$ is (x_1, y_1), the second element of $X \times Y$ is (x_1, y_2), the third element of $X \times Y$ is (x_2, y_1) etc. In this way we see that we can assign a number from $\mathbb{N}$ to each element of $X \times Y$. In other words there exists a mapping from $\mathbb{N}$ onto $X \times Y$. This proves the theorem.

Corollary 1.8 *If either X or Y is countable, then $X \times Y$ is also countable.*

Proof: By Theorem 1.7, there exists a mapping $\theta : \mathbb{N} \to X \times Y$ which is onto. Suppose without loss of generality that X is countable. Then there exists $\alpha : \mathbb{N} \to X$ which is one to one and onto. Let $\beta : X \times Y \to \mathbb{N}$ be defined by $\beta((x, y)) \equiv \alpha^{-1}(x)$. Then by Corollary 1.5, there is a one to one and onto mapping from $X \times Y$ to $\mathbb{N}$. This proves the corollary.

Theorem 1.9 *If X and Y are at most countable, then $X \cup Y$ is at most countable.*

Proof: Let $X = \{x_i\}_{i=1}^{\infty}, Y = \{y_j\}_{j=1}^{\infty}$ and consider the following array consisting of $X \cup Y$ and path through it.

$$\begin{array}{cccccc} x_1 & \rightarrow & x_2 & & x_3 & \rightarrow \\ & \swarrow & & \nearrow & & \\ y_1 & \rightarrow & y_2 & & & \end{array}$$

Thus the first element of $X \cup Y$ is x_1, the second is x_2 the third is y_1 the fourth is y_2 etc. This proves the theorem.

Corollary 1.10 *If either X or Y are countable, then $X \cup Y$ is countable.*

Proof: There is a map from $\mathbb{N}$ onto $X \times Y$. Suppose without loss of generality that X is countable and $\alpha : \mathbb{N} \rightarrow X$ is one to one and onto. Then define $\beta(y) \equiv 1$, for all $y \in Y$,and $\beta(x) \equiv \alpha^{-1}(x)$. Thus, β maps $X \times Y$ onto $\mathbb{N}$ and applying Corollary 1.5 yields the conclusion and proves the corollary.

1.3 Exercises

1. Show the rational numbers, $\mathbb{Q}$, are countable.

2. We say a number is an algebraic number if it is the solution of an equation of the form

$$a_n x^n + \cdots + a_1 x + a_0 = 0$$

   where all the a_j are integers and all exponents are also integers. Thus $\sqrt{2}$ is an algebraic number because it is a solution of the equation $x^2 - 2 = 0$. Using the fundamental theorem of algebra which implies that such equations or order n have at most n solutions, show the set of all algebraic numbers is countable.

3. Let A be a set and let $\mathcal{P}(A)$ be its power set, the set of all subsets of A. Show there does not exist any function f, which maps A onto $\mathcal{P}(A)$. Thus the power set is always strictly larger than the set from which it came. **Hint:** Suppose f is onto. Consider $S \equiv \{x \in A : x \notin f(x)\}$. If f is onto, then $f(y) = S$ for some $y \in A$. Is $y \in f(y)$? Note this argument holds for sets of any size.

4. The empty set is said to be a subset of every set. Why? Consider the statement: If pigs had wings, then they could fly. Is this statement true or false?

5. If $S = \{1, \cdots, n\}$, show $\mathcal{P}(S)$ has exactly 2^n elements in it. **Hint:** You might try a few cases first.

6. Show the set of all subsets of $\mathbb{N}$, the natural numbers, which have 3 elements, is countable. Is the set of all subsets of $\mathbb{N}$ which have finitely many elements countable? How about the set of all subsets of $\mathbb{N}$?

Linear Algebra

2.1 Vector Spaces

A vector space is an Abelian group of "vectors" satisfying the axioms of an Abelian group,

$$\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v},$$

the commutative law of addition,

$$(\mathbf{v}+\mathbf{w})+\mathbf{z}=\mathbf{v}+(\mathbf{w}+\mathbf{z}),$$

the associative law for addition,

$$\mathbf{v}+\mathbf{0}=\mathbf{v},$$

the existence of an additive identity,

$$\mathbf{v}+(-\mathbf{v})=\mathbf{0},$$

the existence of an additive inverse, along with a field of "scalars", $\mathbb{F}$ which are allowed to multiply the vectors according to the following rules. (The Greek letters denote scalars.)

$$\alpha(\mathbf{v}+\mathbf{w})=\alpha\mathbf{v}+\alpha\mathbf{w}, \tag{2.1}$$

$$(\alpha+\beta)\mathbf{v}=\alpha\mathbf{v}+\beta\mathbf{v}, \tag{2.2}$$

$$\alpha(\beta\mathbf{v})=\alpha\beta(\mathbf{v}), \tag{2.3}$$

$$1\mathbf{v}=\mathbf{v}. \tag{2.4}$$

The field of scalars will always be assumed to be either $\mathbb{R}$ or $\mathbb{C}$ and the vector space will be called real or complex depending on whether the field is $\mathbb{R}$ or $\mathbb{C}$. A vector space is also called a linear space.

For example, $\mathbb{R}^n$ with the usual conventions is an example of a real vector space and $\mathbb{C}^n$ is an example of a complex vector space.

Definition 2.1 *If* $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}\subseteq V$, *a vector space, then*

$$\mathit{span}(\mathbf{v}_1,\cdots,\mathbf{v}_n)\equiv\{\sum_{i=1}^{n}\alpha_i\mathbf{v}_i:\alpha_i\in\mathbb{F}\}.$$

A subset, $W\subseteq V$ *is said to be a subspace if it is also a vector space with the same field of scalars. Thus* $W\subseteq V$ *is a subspace if* $\alpha\mathbf{u}+\beta\mathbf{v}\in W$ *whenever* $\alpha,\beta\in\mathbb{F}$ *and* $\mathbf{u},\mathbf{v}\in W$. *The span of a set of vectors as just described is an example of a subspace.*

Definition 2.2 *If* $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \subseteq V$, *we say the set of vectors is linearly independent if*

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

implies

$$\alpha_1 = \cdots = \alpha_n = 0$$

and we say $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ *is a basis for* V *if*

$$\mathit{span}\,(\mathbf{v}_1, \cdots, \mathbf{v}_n) = V$$

and $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ *is linearly independent. We say the set of vectors is linearly dependent if it is not linearly independent.*

Theorem 2.3 *If*

$$\mathit{span}\,(\mathbf{u}_1, \cdots, \mathbf{u}_m) = \mathit{span}\,(\mathbf{v}_1, \cdots, \mathbf{v}_n)$$

and $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ *are linearly independent, then* $m \leq n$.

Proof: Let $V \equiv \mathit{span}\,(\mathbf{v}_1, \cdots, \mathbf{v}_n)$. Then

$$\mathbf{u}_1 = \sum_{i=1}^{n} c_i \mathbf{v}_i$$

and one of the scalars c_i is non zero. This must occur because of the assumption that $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ is linearly independent. (We cannot have any of these vectors the zero vector and still have the set be linearly independent.) Without loss of generality, we assume $c_1 \neq 0$. Then solving for $\mathbf{v}_1$ we find

$$\mathbf{v}_1 \in \mathit{span}\,(\mathbf{u}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)$$

and so

$$V = \mathit{span}\,(\mathbf{u}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)\,.$$

Thus, there exist scalars $c_1, \cdots, c_n$ such that

$$\mathbf{u}_2 = c_1 \mathbf{u}_1 + \sum_{k=2}^{n} c_k \mathbf{v}_k.$$

By the assumption that $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ is linearly independent, we know that at least one of the c_k for $k \geq 2$ is non zero. Without loss of generality, we suppose this scalar is c_2. Then as before,

$$\mathbf{v}_2 \in \mathit{span}\,(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \cdots, \mathbf{v}_n)$$

and so $V = \mathit{span}\,(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \cdots, \mathbf{v}_n)$. Now suppose $m > n$. Then we can continue this process of replacement till we obtain

$$V = \mathit{span}\,(\mathbf{u}_1, \cdots, \mathbf{u}_n) = \mathit{span}\,(\mathbf{u}_1, \cdots, \mathbf{u}_m)\,.$$

Thus, for some choice of scalars, $c_1 \cdots c_n$,

$$\mathbf{u}_m = \sum_{i=1}^{n} c_i \mathbf{u}_i$$

which contradicts the assumption of linear independence of the vectors $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$. Therefore, $m \leq n$ and this proves the Theorem.

Corollary 2.4 *If* $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ *and* $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ *are two bases for* V, *then* $m = n$.

Proof: By Theorem 2.3, $m \leq n$ and $n \leq m$.

Definition 2.5 *We say a vector space* V *is of dimension* n *if it has a basis consisting of* n *vectors. This is well defined thanks to Corollary 2.4. We assume here that* $n < \infty$ *and say such a vector space is finite dimensional.*

Theorem 2.6 *If* $V = span(\mathbf{u}_1, \cdots, \mathbf{u}_n)$ *then some subset of* $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ *is a basis for* V. *Also, if* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\} \subseteq V$ *is linearly independent and the vector space is finite dimensional, then the set,* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$, *can be enlarged to obtain a basis of* V.

Proof: Let

$$S = \{E \subseteq \{\mathbf{u}_1, \cdots, \mathbf{u}_n\} \text{ such that } span(E) = V\}.$$

For $E \in S$, let $|E|$ denote the number of elements of E. Let

$$m \equiv \min\{|E| \text{ such that } E \in S\}.$$

Thus there exist vectors

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_m\} \subseteq \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$$

such that

$$span(\mathbf{v}_1, \cdots, \mathbf{v}_m) = V$$

and m is as small as possible for this to happen. If this set is linearly independent, it follows it is a basis for V and the theorem is proved. On the other hand, if the set is not linearly independent, then there exist scalars,

$$c_1, \cdots, c_m$$

such that

$$\mathbf{0} = \sum_{i=1}^{m} c_i \mathbf{v}_i$$

and not all the c_i are equal to zero. Suppose $c_k \neq 0$. Then we can solve for the vector, $\mathbf{v}_k$ in terms of the other vectors. Consequently,

$$V = span(\mathbf{v}_1, \cdots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \cdots, \mathbf{v}_m)$$

contradicting the definition of m. This proves the first part of the theorem.

To obtain the second part, begin with $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$. If

$$span(\mathbf{u}_1, \cdots, \mathbf{u}_k) = V,$$

we are done. If not, there exists a vector,

$$\mathbf{u}_{k+1} \notin span(\mathbf{u}_1, \cdots, \mathbf{u}_k).$$

Then $\{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ is also linearly independent. Continue adding vectors in this way until the resulting list spans the space V. Then this list is a basis and this proves the theorem.

2.2 Linear Transformations

Definition 2.7 *Let V and W be two finite dimensional vector spaces. We say*

$$L \in \mathcal{L}(V, W)$$

if for all scalars α and β, and vectors $\mathbf{v}, \mathbf{w}$,

$$L(\alpha\mathbf{v}+\beta\mathbf{w}) = \alpha L(\mathbf{v}) + \beta L(\mathbf{w}).$$

We will sometimes write $L\mathbf{v}$ when it is clear that $L(\mathbf{v})$ is meant.

An example of a linear transformation is familiar matrix multiplication. Let $A = (a_{ij})$ be an $m \times n$ matrix. Then we may define a linear transformation $L : \mathbb{F}^n \to \mathbb{F}^m$ by

$$(L\mathbf{v})_i \equiv \sum_{j=1}^{n} a_{ij} v_j.$$

Here

$$\mathbf{v} \equiv \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Also, if V is an n dimensional vector space and $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis for V, there exists a linear map

$$q : \mathbb{F}^n \to V$$

defined as

$$q(\mathbf{a}) \equiv \sum_{i=1}^{n} a_i \mathbf{v}_i$$

where

$$\mathbf{a} = \sum_{i=1}^{n} a_i \mathbf{e}_i,$$

for $\mathbf{e}_i$ the standard basis vectors for $\mathbb{F}^n$ consisting of

$$\mathbf{e}_i \equiv \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

where the one is in the *ith* slot. It is clear that q defined in this way, is one to one, onto, and linear. For $\mathbf{v} \in V$, $q^{-1}(\mathbf{v})$ is a list of scalars called the components of $\mathbf{v}$ with respect to the basis $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$.

Definition 2.8 *Given a linear transformation L, mapping V to W, where $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis of V and $\{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ is a basis for W, an $m \times n$ matrix $A = (a_{ij})$is called the matrix of the transformation L with respect to the given choice of bases for V and W, if whenever $\mathbf{v} \in V$, then multiplication of the components of $\mathbf{v}$ by (a_{ij}) yields the components of $L\mathbf{v}$.*

The following diagram is descriptive of the definition. Here q_V and q_W are the maps defined above with reference to the bases, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ respectively.

$$\begin{array}{ccccc} & & L & & \\ \{\mathbf{v}_1, \cdots, \mathbf{v}_n\} & V & \to & W & \{\mathbf{w}_1, \cdots, \mathbf{w}_m\} \\ & q_V \uparrow & \circ & \uparrow q_W & \\ & \mathbb{F}^n & \to & \mathbb{F}^m & \\ & & A & & \end{array} \tag{2.5}$$

Letting $\mathbf{b} \in \mathbb{F}^n$, this requires

$$\sum_{i,j} a_{ij} b_j \mathbf{w}_i = L \sum_j b_j \mathbf{v}_j = \sum_j b_j L \mathbf{v}_j.$$

Now

$$L\mathbf{v}_j = \sum_i c_{ij} \mathbf{w}_i \tag{2.6}$$

for some choice of scalars c_{ij} because $\{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ is a basis for W. Hence

$$\sum_{i,j} a_{ij} b_j \mathbf{w}_i = \sum_j b_j \sum_i c_{ij} \mathbf{w}_i = \sum_{i,j} c_{ij} b_j \mathbf{w}_i.$$

It follows from the linear independence of $\{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ that

$$\sum_j a_{ij} b_j = \sum_j c_{ij} b_j$$

for any choice of $\mathbf{b} \in \mathbb{F}^n$ and consequently

$$a_{ij} = c_{ij}$$

where c_{ij} is defined by (2.6). It may help to write (2.6) in the form

$$\begin{pmatrix} L\mathbf{v}_1 & \cdots & L\mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_m \end{pmatrix} C = \begin{pmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_m \end{pmatrix} A \tag{2.7}$$

where $C = (c_{ij})\,, A = (a_{ij})\,.$

Example 2.9 *Let*

$$V \equiv \{ \text{ polynomials of degree 3 or less}\},$$

$$W \equiv \{ \text{ polynomials of degree 2 or less}\},$$

and $L \equiv D$ where D is the differentiation operator. A basis for V is $\{1, x, x^2, x^3\}$ and a basis for W is $\{1, x, x^2\}$.

What is the matrix of this linear transformation with respect to this basis? Using (2.7),

$$\begin{pmatrix} 0 & 1 & 2x & 3x^2 \end{pmatrix} = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} C.$$

It follows from this that

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Now consider the important case where $V=\mathbb{F}^n$, $W=\mathbb{F}^m$, and the basis chosen is the standard basis of vectors $\mathbf{e}_i$ described above. Let L be a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$ and let A be the matrix of the transformation with respect to these bases. In this case the coordinate maps q_V and q_W are simply the identity map and we need

$$\pi_i(L\mathbf{b})=\pi_i(A\mathbf{b})$$

where π_i denotes the map which takes a vector in $\mathbb{F}^m$ and returns the ith entry in the vector, the ith component of the vector with respect to the standard basis vectors. Thus, if the components of the vector in $\mathbb{F}^n$ with respect to the standard basis are $(b_1,\cdots,b_n)$,

$$\mathbf{b}=\begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}^T=\sum_i b_i\mathbf{e}_i,$$

then

$$\pi_i(L\mathbf{b})\equiv(L\mathbf{b})_i=\sum_j a_{ij}b_j.$$

What about the situation where different pairs of bases are chosen for V and W? How are the two matrices with respect to these choices related? Consider the following diagram which illustrates the situation.

$$\begin{array}{ccc}
\mathbb{F}^n & \underrightarrow{A_2} & \mathbb{F}^m \\
q_2\downarrow & \circ & p_2\downarrow \\
V & \underrightarrow{L} & W \\
q_1\uparrow & \circ & p_1\uparrow \\
\mathbb{F}^n & \underrightarrow{A_1} & \mathbb{F}^m
\end{array}$$

In this diagram q_i and p_i are coordinate maps as described above. We see from the diagram that

$$p_1^{-1}p_2A_2q_2^{-1}q_1=A_1,$$

where $q_2^{-1}q_1$ and $p_1^{-1}p_2$ are one to one, onto, and linear maps.

In the special case where $V=W$ and only one basis is used for $V=W$, this becomes

$$q_1^{-1}q_2A_2q_2^{-1}q_1=A_1.$$

Letting S be the matrix of the linear transformation $q_2^{-1}q_1$ with respect to the standard basis vectors in $\mathbb{F}^n$, we get

$$S^{-1}A_2S=A_1.$$

When this occurs, we say that A_1 is similar to A_2 and we call $A\to S^{-1}AS$ a similarity transformation.

Theorem 2.10 *In the vector space of $n\times n$ matrices, we say*

$$A\sim B$$

if there exists an invertible matrix S such that

$$A=S^{-1}BS.$$

Then $\sim$ is an equivalence relation and $A\sim B$ if and only if whenever V is an n dimensional vector space, there exists $L\in\mathcal{L}(V,V)$ and bases $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ and $\{\mathbf{w}_1,\cdots,\mathbf{w}_n\}$ such that A is the matrix of L with respect to $\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$ and B is the matrix of L with respect to $\{\mathbf{w}_1,\cdots,\mathbf{w}_n\}$.

Proof: $A \sim A$ because $S = I$ works in the definition. If $A \sim B$, then $B \sim A$, because

$$A = S^{-1}BS$$

implies

$$B = SAS^{-1}.$$

If $A \sim B$ and $B \sim C$, then

$$A = S^{-1}BS,\ B = T^{-1}CT$$

and so

$$A = S^{-1}T^{-1}CTS = (TS)^{-1}\,CTS$$

which implies $A \sim C$. This verifies the first part of the conclusion.

Now let V be an n dimensional vector space, $A \sim B$ and pick a basis for V,

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}.$$

Define $L \in \mathcal{L}(V, V)$ by

$$L\mathbf{v}_i \equiv \sum_j a_{ji}\mathbf{v}_j$$

where $A = (a_{ij})$. Then if $B = (b_{ij})$, and $S = (s_{ij})$ is the matrix which provides the similarity transformation,

$$A = S^{-1}BS,$$

between A and B, it follows that

$$L\mathbf{v}_i = \sum_{r,s,j} s_{ir}b_{rs}\left(s^{-1}\right)_{sj}\mathbf{v}_j. \tag{2.8}$$

Now define

$$\mathbf{w}_i \equiv \sum_j \left(s^{-1}\right)_{ij}\mathbf{v}_j.$$

Then from (2.8),

$$\sum_i \left(s^{-1}\right)_{ki} L\mathbf{v}_i = \sum_{i,j,r,s} \left(s^{-1}\right)_{ki} s_{ir}b_{rs}\left(s^{-1}\right)_{sj}\mathbf{v}_j$$

and so

$$L\mathbf{w}_k = \sum_s b_{ks}\mathbf{w}_s.$$

This proves the theorem because the if part of the conclusion was established earlier.

2.3 Inner product spaces

Definition 2.11 *A vector space X is said to be a normed linear space if there exists a function, denoted by $|\cdot| : X \to [0,\infty)$ which satisfies the following axioms.*

1. $|x| \geq 0$ *for all* $x \in X$, *and* $|x| = 0$ *if and only if* $x = 0$.
2. $|ax| = |a|\,|x|$ *for all* $a \in \mathbb{F}$.
3. $|x+y| \leq |x| + |y|$.

Note that we are using the same notation for the norm as for the absolute value. This is because the norm is just a generalization to vector spaces of the concept of absolute value. However, the notation $||x||$ is also often used. Not all norms are created equal. There are many geometric properties which they may or may not possess. There is also a concept called an inner product which is discussed next. It turns out that the best norms come from an inner product.

Definition 2.12 *A mapping $(\cdot,\cdot) : V \times V \to \mathbb{F}$ is called an inner product if it satisfies the following axioms.*

1. $(x,y) = \overline{(y,x)}$.
2. $(x,x) \geq 0$ *for all* $x \in V$ *and equals zero if and only if* $x = 0$.
3. $(ax+by,z) = a\,(x,z) + b\,(y,z)$ *whenever* $a, b \in \mathbb{F}$.

Note that 2 and 3 imply $(x, ay+bz) = \overline{a}(x,y) + \overline{b}(x,z)$.
We will show that if $(\cdot,\cdot)$ is an inner product, then

$$(x,x)^{1/2} \equiv |x|$$

defines a norm.

Definition 2.13 *A normed linear space in which the norm comes from an inner product as just described is called an inner product space. A Hilbert space is a complete inner product space. Recall this means that every Cauchy sequence,$\{x_n\}$, one which satisfies*

$$\lim_{n,m\to\infty} |x_n - x_m| = 0,$$

converges.

Example 2.14 *Let $V = \mathbb{C}^n$ with the inner product given by*

$$(\mathbf{x},\mathbf{y}) \equiv \sum_{k=1}^{n} x_k \overline{y}_k.$$

. This is an example of a complex Hilbert space.

Example 2.15 *Let $V = \mathbb{R}^n$,*

$$(\mathbf{x},\mathbf{y}) = \mathbf{x}\cdot\mathbf{y} \equiv \sum_{j=1}^{n} x_j y_j.$$

. This is an example of a real Hilbert space.

Theorem 2.16 *(Cauchy Schwartz) In any inner product space*

$$|(x,y)| \leq |x||y|.$$

where $|x| \equiv (x,x)^{1/2}$.

Proof: Let $\omega \in \mathbb{C}$, $|\omega| = 1$, and $\overline{\omega}(x,y) = |(x,y)| = \operatorname{Re}(x, y\omega)$. Let

$$F(t) = (x + ty\omega, x + t\omega y).$$

If $y = 0$ there is nothing to prove because

$$(x,0) = (x, 0+0) = (x,0) + (x,0)$$

and so $(x,0) = 0$. Thus, we may assume $y \neq 0$. Then from the axioms of the inner product,

$$F(t) = |x|^2 + 2t \operatorname{Re}(x, \omega y) + t^2 |y|^2 \geq 0.$$

This yields

$$|x|^2 + 2t|(x,y)| + t^2|y|^2 \geq 0.$$

Since this inequality holds for all $t \in \mathbb{R}$, it follows from the quadratic formula that

$$4|(x,y)|^2 - 4|x|^2|y|^2 \leq 0.$$

This yields the conclusion and proves the theorem.

Earlier it was claimed that the inner product defines a norm. In this next proposition this claim is proved.

Proposition 2.17 *For an inner product space,* $|x| \equiv (x,x)^{1/2}$ *does specify a norm.*

Proof: All the axioms are obvious except the triangle inequality. To verify this,

$$\begin{aligned} |x+y|^2 &\equiv (x+y, x+y) \equiv |x|^2 + |y|^2 + 2\operatorname{Re}(x,y) \\ &\leq |x|^2 + |y|^2 + 2|(x,y)| \\ &\leq |x|^2 + |y|^2 + 2|x||y| = (|x| + |y|)^2. \end{aligned}$$

The best norms of all are those which come from an inner product because of the following identity which is known as the parallelogram identity.

Proposition 2.18 *If* $(V, (\cdot,\cdot))$ *is an inner product space then for* $|x| \equiv (x,x)^{1/2}$, *the following identity holds.*

$$|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2.$$

It turns out that the validity of this identity is equivalent to the existence of an inner product which determines the norm as described above. These sorts of considerations are topics for more advanced courses on functional analysis.

Definition 2.19 *We say a basis for an inner product space,* $\{u_1, \cdots, u_n\}$ *is an orthonormal basis if*

$$(u_k, u_j) = \delta_{kj} \equiv \begin{cases} 1 \text{ if } k = j \\ 0 \text{ if } k \neq j \end{cases}.$$

Note that if a list of vectors satisfies the above condition for being an orthonormal set, then the list of vectors is automatically linearly independent. To see this, suppose

$$\sum_{j=1}^{n} c_j u_j = 0$$

Then taking the inner product of both sides with u_k, we obtain

$$0 = \sum_{j=1}^{n} c_j (u_j, u_k) = \sum_{j=1}^{n} c_j \delta_{jk} = c_k.$$

Lemma 2.20 *Let X be a finite dimensional inner product space of dimension n. Then there exists an orthonormal basis for X, $\{u_1, \cdots, u_n\}$.*

Proof: Let $\{x_1, \cdots, x_n\}$ be a basis for X. Let $u_1 \equiv x_1 / |x_1|$. Now suppose for some $k < n$, $u_1, \cdots, u_k$ have been chosen such that $(u_j, u_k) = \delta_{jk}$. Then we define

$$u_{k+1} \equiv \frac{x_{k+1} - \sum_{j=1}^{k} (x_{k+1}, u_j) u_j}{\left| x_{k+1} - \sum_{j=1}^{k} (x_{k+1}, u_j) u_j \right|},$$

where the numerator is not equal to zero because the x_j form a basis. Then if $l \leq k$,

$$\begin{aligned}
(u_{k+1}, u_l) &= C \left((x_{k+1}, u_l) - \sum_{j=1}^{k} (x_{k+1}, u_j)(u_j, u_l) \right) \\
&= C \left((x_{k+1}, u_l) - \sum_{j=1}^{k} (x_{k+1}, u_j) \delta_{lj} \right) \\
&= C \left((x_{k+1}, u_l) - (x_{k+1}, u_l) \right) = 0.
\end{aligned}$$

The vectors, $\{u_j\}_{j=1}^{n}$, generated in this way are therefore an orthonormal basis.

The process by which these vectors were generated is called the Gramm Schmidt process.

Lemma 2.21 *Suppose $\{u_j\}_{j=1}^{n}$ is an orthonormal basis for an inner product space X. Then for all $x \in X$,*

$$x = \sum_{j=1}^{n} (x, u_j) u_j.$$

Proof: By assumption that this is an orthonormal basis,

$$\sum_{j=1}^{n} (x, u_j)(u_j, u_l) = (x, u_l).$$

Letting $y = \sum_{j=1}^{n} (x, u_j) u_j$, it follows $(x - y, u_j) = 0$ for all j. Hence, for any choice of scalars, $c_1, \cdots, c_n$,

$$\left(x - y, \sum_{j=1}^{n} c_j u_j \right) = 0$$

and so $(x - y, z) = 0$ for all $z \in X$. Thus this holds in particular for $z = x - y$. Therefore, $x = y$ and this proves the theorem.

The next theorem is one of the most important results in the theory of inner product spaces. It is called the Riesz representation theorem.

Theorem 2.22 *Let $f \in \mathcal{L}(X, \mathbb{F})$ where X is a finite dimensional inner product space. Then there exists a unique $z \in X$ such that for all $x \in X$,*

$$f(x) = (x, z).$$

Proof: First we verify uniqueness. Suppose z_j works for $j = 1, 2$. Then for all $x \in X$,

$$0 = f(x) - f(x) = (x, z_1 - z_2)$$

and so $z_1 = z_2$.

It remains to verify existence. By Lemma 2.20, there exists an orthonormal basis, $\{u_j\}_{j=1}^n$. Define

$$z \equiv \sum_{j=1}^n \overline{f(u_j)} u_j.$$

Then using Lemma 2.21,

$$\begin{aligned}(x, z) &= \left(x, \sum_{j=1}^n \overline{f(u_j)} u_j\right) = \sum_{j=1}^n f(u_j)(x, u_j) \\ &= f\left(\sum_{j=1}^n (x, u_j) u_j\right) = f(x).\end{aligned}$$

This proves the theorem.

Corollary 2.23 *Let $A \in \mathcal{L}(X, Y)$ where X and Y are two finite dimensional inner product spaces. Then there exists a unique $A^* \in \mathcal{L}(Y, X)$ such that*

$$(Ax, y)_Y = (x, A^*y)_X$$

for all $x \in X$ and $y \in Y$.

Proof: Let $f_y \in \mathcal{L}(X, \mathbb{F})$ be defined as

$$f_y(x) \equiv (Ax, y)_Y.$$

Then by the Riesz representation theorem, there exists a unique element of X, $A^*(y)$ such that

$$(Ax, y)_Y = (x, A^*(y))_X.$$

It only remains to verify that A^* is linear. Let a and b be scalars. Then for all $x \in X$,

$$(x, A^*(ay_1 + by_2))_X \equiv \overline{a}(Ax, y_1) + \overline{b}(Ax, y_2)$$

$$\overline{a}(x, A^*(y_1)) + \overline{b}(x, A^*(y_2)) = (x, aA^*(y_1) + bA^*(y_2)).$$

By uniqueness, $A^*(ay_1 + by_2) = aA^*(y_1) + bA^*(y_2)$ which shows A^* is linear as claimed.

The linear map, A^* is called the adjoint of A. In the case when $A : X \to X$ and $A = A^*$, we call A a self adjoint map. The next theorem will prove useful.

Theorem 2.24 *Suppose V is a subspace of $\mathbb{F}^n$ having dimension $p \leq n$. Then there exists a $Q \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ such that $QV \subseteq \mathbb{F}^p$ and $|Q\mathbf{x}| = |\mathbf{x}|$ for all $\mathbf{x}$. Also*

$$Q^*Q = QQ^* = I.$$

Proof: By Lemma 2.20 there exists an orthonormal basis for $V, \{\mathbf{v}_i\}_{i=1}^p$. By using the Gramm Schmidt process we may extend this orthonormal basis to an orthonormal basis of the whole space, $\mathbb{F}^n$,

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_p, \mathbf{v}_{p+1}, \cdots, \mathbf{v}_n\}.$$

Now define $Q \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ by $Q(\mathbf{v}_i) \equiv \mathbf{e}_i$ and extend linearly. If $\sum_{i=1}^n x_i \mathbf{v}_i$ is an arbitrary element of $\mathbb{F}^n$,

$$\left|Q\left(\sum_{i=1}^n x_i \mathbf{v}_i\right)\right|^2 = \left|\sum_{i=1}^n x_i \mathbf{e}_i\right|^2 = \sum_{i=1}^n |x_i|^2 = \left|\sum_{i=1}^n x_i \mathbf{v}_i\right|^2.$$

It remains to verify that $Q^*Q = QQ^* = I$. To do so, let $\mathbf{x}, \mathbf{y} \in \mathbb{F}^p$. Then

$$(Q(\mathbf{x}+\mathbf{y}), Q(\mathbf{x}+\mathbf{y})) = (\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}).$$

Thus

$$|Q\mathbf{x}|^2 + |Q\mathbf{y}|^2 + \operatorname{Re}(Q\mathbf{x}, Q\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + \operatorname{Re}(\mathbf{x}, \mathbf{y})$$

and since Q preserves norms, it follows that for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$,

$$\operatorname{Re}(Q\mathbf{x}, Q\mathbf{y}) = \operatorname{Re}(\mathbf{x}, Q^*Q\mathbf{y}) = \operatorname{Re}(\mathbf{x}, \mathbf{y}).$$

Therefore, since this holds for all $\mathbf{x}$, it follows that $Q^*Q\mathbf{y} = \mathbf{y}$ showing that $Q^*Q = I$. Now

$$Q = Q(Q^*Q) = (QQ^*)Q.$$

Since Q is one to one, this implies

$$I = QQ^*$$

and proves the theorem.

This case of a self adjoint map turns out to be very important in applications. It is also easy to discuss the eigenvalues and eigenvectors of such a linear map. For $A \in \mathcal{L}(X, X)$, we give the following definition of eigenvalues and eigenvectors.

Definition 2.25 *A non zero vector, y is said to be an eigenvector for $A \in \mathcal{L}(X, X)$ if there exists a scalar, λ, called an eigenvalue, such that*

$$Ay = \lambda y.$$

The important thing to remember about eigenvectors is that they are **never** equal to zero. The following theorem is about the eigenvectors and eigenvalues of a self adjoint operator. The proof given generalizes to the situation of a compact self adjoint operator on a Hilbert space and leads to many very useful results. It is also a very elementary proof because it does not use the fundamental theorem of algebra and it contains a way, very important in applications, of finding the eigenvalues. We will use the following notation.

Definition 2.26 *Let X be an inner product space and let $S \subseteq X$. Then*

$$S^\perp \equiv \{x \in X : (x, s) = 0 \text{ for all } s \in S\}.$$

Note that even if S is not a subspace, $S^\perp$ is.

Definition 2.27 *Let X be a finite dimensional inner product space and let $A \in \mathcal{L}(X, X)$. We say A is self adjoint if $A^* = A$.*

Theorem 2.28 *Let $A \in \mathcal{L}(X, X)$ be self adjoint. Then there exists an orthonormal basis of eigenvectors, $\{u_j\}_{j=1}^n$.*

Proof: Consider (Ax, x). This quantity is always a real number because

$$\overline{(Ax, x)} = (x, Ax) = (x, A^* x) = (Ax, x)$$

thanks to the assumption that A is self adjoint. Now define

$$\lambda_1 \equiv \inf \{(Ax, x) : |x| = 1, x \in X_1 \equiv X\}.$$

Claim: λ_1 is finite and there exists $v_1 \in X$ with $|v_1| = 1$ such that $(Av_1, v_1) = \lambda_1$.

Proof: Let $\{u_j\}_{j=1}^n$ be an orthonormal basis for X and for $x \in X$, let $(x_1, \cdots, x_n)$ be defined as the components of the vector x. Thus,

$$x = \sum_{j=1}^n x_j u_j.$$

Since this is an orthonormal basis, it follows from the axioms of the inner product that

$$|x|^2 = \sum_{j=1}^n |x_j|^2.$$

Thus

$$(Ax, x) = \left(\sum_{k=1}^n x_k A u_k, \sum_{j=1} x_j u_j \right) = \sum_{k,j} x_k \overline{x_j} (Au_k, u_j),$$

a continuous function of $(x_1, \cdots, x_n)$. Thus this function achieves its minimum on the closed and bounded subset of $\mathbb{F}^n$ given by

$$\{(x_1, \cdots, x_n) \in \mathbb{F}^n : \sum_{j=1}^n |x_j|^2 = 1\}.$$

Then $v_1 \equiv \sum_{j=1}^n x_j u_j$ where $(x_1, \cdots, x_n)$ is the point of $\mathbb{F}^n$ at which the above function achieves its minimum. This proves the claim.

Continuing with the proof of the theorem, let $X_2 \equiv \{v_1\}^\perp$ and let

$$\lambda_2 \equiv \inf \{(Ax, x) : |x| = 1, x \in X_2\}$$

As before, there exists $v_2 \in X_2$ such that $(Av_2, v_2) = \lambda_2$. Now let $X_2 \equiv \{v_1, v_2\}^\perp$ and continue in this way. This leads to an increasing sequence of real numbers, $\{\lambda_k\}_{k=1}^n$ and an orthonormal set of vectors, $\{v_1, \cdots, v_n\}$. It only remains to show these are eigenvectors and that the λ_j are eigenvalues.

Consider the first of these vectors. Letting $w \in X_1 \equiv X$, the function of the real variable, t, given by

$$f(t) \equiv \frac{(A(v_1 + tw), v_1 + tw)}{|v_1 + tw|^2}$$

$$= \frac{(Av_1, v_1) + 2t \operatorname{Re}(Av_1, w) + t^2 (Aw, w)}{|v_1|^2 + 2t \operatorname{Re}(v_1, w) + t^2 |w|^2}$$

achieves its minimum when $t=0$. Therefore, the derivative of this function evaluated at $t=0$ must equal zero. Using the quotient rule, this implies

$$2\operatorname{Re}(Av_1,w)-2\operatorname{Re}(v_1,w)(Av_1,v_1)$$

$$=2\left(\operatorname{Re}(Av_1,w)-\operatorname{Re}(v_1,w)\lambda_1\right)=0.$$

Thus $\operatorname{Re}(Av_1-\lambda_1 v_1,w)=0$ for all $w\in X$. This implies $Av_1=\lambda_1 v_1$. To see this, let $w\in X$ be arbitrary and let θ be a complex number with $|\theta|=1$ and

$$|(Av_1-\lambda_1 v_1,w)|=\theta(Av_1-\lambda_1 v_1,w).$$

Then

$$|(Av_1-\lambda_1 v_1,w)|=\operatorname{Re}\left(Av_1-\lambda_1 v_1,\overline{\theta}w\right)=0.$$

Since this holds for all w, $Av_1=\lambda_1 v_1$. Now suppose $Av_k=\lambda_k v_k$ for all $k<m$. We observe that $A:X_m\to X_m$ because if $y\in X_m$ and $k<m$,

$$(Ay,v_k)=(y,Av_k)=(y,\lambda_k v_k)=0,$$

showing that $Ay\in\{v_1,\cdot\cdot\cdot,v_{m-1}\}^{\perp}\equiv X_m$. Thus the same argument just given shows that for all $w\in X_m$,

$$(Av_m-\lambda_m v_m,w)=0. \tag{2.9}$$

For arbitrary $w\in X$.

$$w=\left(w-\sum_{k=1}^{m-1}(w,v_k)v_k\right)+\sum_{k=1}^{m-1}(w,v_k)v_k\equiv w_{\perp}+w_m$$

and the term in parenthesis is in $\{v_1,\cdot\cdot\cdot,v_{m-1}\}^{\perp}\equiv X_m$ while the other term is contained in the span of the vectors, $\{v_1,\cdot\cdot\cdot,v_{m-1}\}$. Thus by (2.9),

$$(Av_m-\lambda_m v_m,w)=(Av_m-\lambda_m v_m,w_{\perp}+w_m)$$

$$=(Av_m-\lambda_m v_m,w_m)=0$$

because

$$A:X_m\to X_m\equiv\{v_1,\cdot\cdot\cdot,v_{m-1}\}^{\perp}$$

and $w_m\in \mathit{span}(v_1,\cdot\cdot\cdot,v_{m-1})$. Therefore, $Av_m=\lambda_m v_m$ for all m. This proves the theorem.

When a matrix has such a basis of eigenvectors, we say it is non defective.

There are more general theorems of this sort involving normal linear transformations. We say a linear transformation is normal if $AA^*=A^*A$. It happens that normal matrices are non defective. The proof involves the fundamental theorem of algebra and is outlined in the exercises.

As an application of this theorem, we give the following fundamental result, important in geometric measure theory and continuum mechanics. It is sometimes called the right polar decomposition.

Theorem 2.29 *Let $F\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ where $m\geq n$. Then there exists $R\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ and $U\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ such that*

$$F=RU,\ U=U^*,$$

all eigen values of U are non negative,

$$U^2=F^*F,R^*R=I,$$

and $|R\mathbf{x}|=|\mathbf{x}|$.

Proof: $(F^*F)^* = F^*F$ and so by linear algebra there is an orthonormal basis of eigenvectors, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ such that

$$F^*F\mathbf{v}_i = \lambda_i \mathbf{v}_i.$$

It is also clear that $\lambda_i \geq 0$ because

$$\lambda_i (\mathbf{v}_i, \mathbf{v}_i) = (F^*F\mathbf{v}_i, \mathbf{v}_i) = (F\mathbf{v}_i, F\mathbf{v}_i) \geq 0.$$

Now if $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we define the tensor product $\mathbf{u} \otimes \mathbf{v} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ by

$$\mathbf{u} \otimes \mathbf{v}(\mathbf{w}) \equiv (\mathbf{w}, \mathbf{v})\,\mathbf{u}.$$

Then $F^*F = \sum_{i=1}^n \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i$ because both linear transformations agree on the basis $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$. Let

$$U \equiv \sum_{i=1}^n \lambda_i^{1/2} \mathbf{v}_i \otimes \mathbf{v}_i.$$

Then $U^2 = F^*F$, $U = U^*$, and the eigenvalues of U, $\left\{\lambda_i^{1/2}\right\}_{i=1}^n$ are all nonnegative.

Now R is defined on $U(\mathbb{R}^n)$ by

$$RU\mathbf{x} \equiv F\mathbf{x}.$$

This is well defined because if $U\mathbf{x}_1 = U\mathbf{x}_2$, then $U^2(\mathbf{x}_1 - \mathbf{x}_2) = 0$ and so

$$0 = (F^*F(\mathbf{x}_1 - \mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2) = |F(\mathbf{x}_1 - \mathbf{x}_2)|^2.$$

Now $|RU\mathbf{x}|^2 = |U\mathbf{x}|^2$ because

$$|RU\mathbf{x}|^2 = |F\mathbf{x}|^2 = (F\mathbf{x}, F\mathbf{x}) = (F^*F\mathbf{x}, \mathbf{x})$$

$$= \left(U^2\mathbf{x}, \mathbf{x}\right) = (U\mathbf{x}, U\mathbf{x}) = |U\mathbf{x}|^2.$$

Let $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ be an orthonormal basis for

$$U(\mathbb{R}^n)^\perp \equiv \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x}, \mathbf{z}) = 0 \text{ for all } \mathbf{z} \in U(\mathbb{R}^n)\}$$

and let $\{\mathbf{y}_1, \cdots, \mathbf{y}_p\}$ be an orthonormal basis for $F(\mathbb{R}^n)^\perp$. Then $p \geq r$ because if $\{F(\mathbf{z}_i)\}_{i=1}^s$ is an orthonormal basis for $F(\mathbb{R}^n)$, it follows that $\{U(\mathbf{z}_i)\}_{i=1}^s$ is orthonormal in $U(\mathbb{R}^n)$ because

$$(U\mathbf{z}_i, U\mathbf{z}_j) = \left(U^2\mathbf{z}_i, \mathbf{z}_j\right) = (F^*F\mathbf{z}_i, \mathbf{z}_j) = (F\mathbf{z}_i, F\mathbf{z}_j).$$

Therefore,

$$p + s = m \geq n = r + \dim U(\mathbb{R}^n) \geq r + s.$$

Now define $R \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ by $R\mathbf{x}_i \equiv \mathbf{y}_i$, $i = 1, \cdots, r$. Thus

$$\left| R\left(\sum_{i=1}^r c_i \mathbf{x}_i + U\mathbf{v}\right)\right|^2 = \left|\sum_{i=1}^r c_i \mathbf{y}_i + F\mathbf{v}\right|^2 = \sum_{i=1}^r |c_i|^2 + |F\mathbf{v}|^2$$

$$= \sum_{i=1}^r |c_i|^2 + |U\mathbf{v}|^2 = \left|\sum_{i=1}^r c_i \mathbf{x}_i + U\mathbf{v}\right|^2,$$

and so $|R\mathbf{z}| = |\mathbf{z}|$ which implies that for all $\mathbf{x}, \mathbf{y}$,

$$|\mathbf{x}|^2 + |\mathbf{y}|^2 + 2(\mathbf{x}, \mathbf{y}) = |\mathbf{x} + \mathbf{y}|^2 = |R(\mathbf{x} + \mathbf{y})|^2$$

$$= |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2(R\mathbf{x}, R\mathbf{y}).$$

Therefore,

$$(\mathbf{x}, \mathbf{y}) = (R^* R\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y}$ and so $R^* R = I$ as claimed. This proves the theorem.

2.4 Exercises

1. Show $(A^*)^* = A$ and $(AB)^* = B^* A^*$.

2. Suppose $A : X \to X$, an inner product space, and $A \geq 0$. By this we mean $(Ax, x) \geq 0$ for all $x \in X$ and $A = A^*$. Show that A has a square root, U, such that $U^2 = A$. **Hint:** Let $\{u_k\}_{k=1}^n$ be an orthonormal basis of eigenvectors with $Au_k = \lambda_k u_k$. Show each $\lambda_k \geq 0$ and consider

$$U \equiv \sum_{k=1}^n \lambda_k^{1/2} u_k \otimes u_k$$

3. In the context of Theorem 2.29, suppose $m \leq n$. Show

$$F = UR$$

where

$$U \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m), R \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m), U = U^*$$

, U has all non negative eigenvalues, $U^2 = FF^*$, and $RR^* = I$. **Hint:** This is an easy corollary of Theorem 2.29.

4. Show that if X is the inner product space $\mathbb{F}^n$, and A is an $n \times n$ matrix, then

$$A^* = \overline{A^T}.$$

5. Show that if A is an $n \times n$ matrix and $A = A^*$ then all the eigenvalues and eigenvectors are real and that eigenvectors associated with distinct eigenvalues are orthogonal, (their inner product is zero). Such matrices are called Hermitian.

6. Let the orthonormal basis of eigenvectors of F^*F be denoted by $\{\mathbf{v}_i\}_{i=1}^n$ where $\{\mathbf{v}_i\}_{i=1}^r$ are those whose eigenvalues are positive and $\{\mathbf{v}_i\}_{i=r}^n$ are those whose eigenvalues equal zero. In the context of the RU decomposition for F, show $\{R\mathbf{v}_i\}_{i=1}^n$ is also an orthonormal basis. Next verify that there exists a solution, $\mathbf{x}$, to the equation, $F\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{b} \in span\{R\mathbf{v}_i\}_{i=1}^r$ if and only is $\mathbf{b} \in (\ker F^*)^\perp$. Here $\ker F^* \equiv \{\mathbf{x} : F^*\mathbf{x} = \mathbf{0}\}$ and $(\ker F^*)^\perp \equiv \{\mathbf{y} : (\mathbf{y}, \mathbf{x}) = 0 \text{ for all } \mathbf{x} \in \ker F^*\}$. **Hint:** Show that $F^*\mathbf{x} = \mathbf{0}$ if and only if $UR^*\mathbf{x} = \mathbf{0}$ if and only if $R^*\mathbf{x} \in span\{\mathbf{v}_i\}_{i=r+1}^n$ if and only if $\mathbf{x} \in span\{R\mathbf{v}_i\}_{i=r+1}^n$.

7. Let A and B be $n \times n$ matrices and let the columns of B be

$$\mathbf{b}_1, \cdots, \mathbf{b}_n$$

and the rows of A are

$$\mathbf{a}_1^T, \cdots, \mathbf{a}_n^T.$$

Show the columns of AB are

$$A\mathbf{b}_1 \cdots A\mathbf{b}_n$$

and the rows of AB are

$$\mathbf{a}_1^T B \cdots \mathbf{a}_n^T B.$$

8. Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be an orthonormal basis for $\mathbb{F}^n$. Let Q be a matrix whose ith column is $\mathbf{v}_i$. Show

$$Q^* Q = QQ^* = I.$$

such a matrix is called an orthogonal matrix.

9. Show that a matrix, Q is orthogonal if and only if it preserves distances. By this we mean $|Q\mathbf{v}| = |\mathbf{v}|$. Here $|\mathbf{v}| \equiv (\mathbf{v} \cdot \mathbf{v})^{1/2}$ for the dot product defined above.

10. Suppose $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$ are two orthonormal bases for $\mathbb{F}^n$ and suppose Q is an $n \times n$ matrix satisfying $Q\mathbf{v}_i = \mathbf{w}_i$. Then show Q is orthogonal. If $|\mathbf{v}| = 1$, show there is an orthogonal transformation which maps $\mathbf{v}$ to $\mathbf{e}_1$.

11. Let A be a Hermitian matrix so $A = A^*$ and suppose all eigenvalues of A are larger than δ^2. Show

$$(A\mathbf{v}, \mathbf{v}) \geq \delta^2 |\mathbf{v}|^2$$

Where here, the inner product is

$$(\mathbf{v}, \mathbf{u}) \equiv \sum_{j=1}^{n} v_j \overline{u_j}.$$

12. Let $L \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Let $\{v_1, \cdots, v_n\}$ be a basis for V and let $\{w_1, \cdots, w_m\}$ be a basis for W. Now define $w \otimes v \in \mathcal{L}(V, W)$ by the rule,

$$w \otimes v(u) \equiv (u, v) w.$$

Show $w \otimes v \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ as claimed and that

$$\{w_j \otimes v_i : i = 1, \cdots, n, j = 1, \cdots, m\}$$

is a basis for $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Conclude the dimension of $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ is nm.

13. Let X be an inner product space. Show $|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2$. This is called the parallelogram identity.

2.5 Determinants

Here we give a discussion of the most important properties of determinants. There are more elegant ways to proceed and the reader is encouraged to consult a more advanced algebra book to read these. Another very good source is Apostol [2]. The goal here is to present all of the major theorems on determinants with a minimum of abstract algebra as quickly as possible. In this section and elsewhere $\mathbb{F}$ will denote the field of scalars, usually $\mathbb{R}$ or $\mathbb{C}$. To begin with we make a simple definition.

Definition 2.30 *Let $(k_1, \cdots, k_n)$ be an ordered list of n integers. We define*

$$\pi(k_1, \cdots, k_n) \equiv \prod \{(k_s - k_r) : r < s\}.$$

In words, we consider all terms of the form $(k_s - k_r)$ where k_s comes after k_r in the ordered list and then multiply all these together. We also make the following definition.

$$sgn(k_1, \cdots, k_n) \equiv \begin{cases} 1 \text{ if } \pi(k_1, \cdots, k_n) > 0 \\ -1 \text{ if } \pi(k_1, \cdots, k_n) < 0 \\ 0 \text{ if } \pi(k_1, \cdots, k_n) = 0 \end{cases}$$

This is called the sign of the permutation $\binom{1\cdots n}{k_1\cdots k_n}$ in the case when there are no repeats in the ordered list, $(k_1, \cdots, k_n)$ and $\{k_1, \cdots, k_n\} = \{1, \cdots, n\}$.

Lemma 2.31 *Let $(k_1, \cdots k_i \cdots, k_j, \cdots, k_n)$ be a list of n integers. Then*

$$\pi(k_1, \cdots, k_i \cdots, k_j, \cdots, k_n) = -\pi(k_1, \cdots, k_j \cdots, k_i, \cdots, k_n)$$

and

$$sgn(k_1, \cdots, k_i \cdots, k_j, \cdots, k_n) = -sgn(k_1, \cdots, k_j \cdots, k_i, \cdots, k_n)$$

In words, if we switch two entries the sign changes.

Proof: The two lists are

$$(k_1, \cdots, k_i, \cdots, k_j, \cdots, k_n) \tag{2.10}$$

and

$$(k_1, \cdots, k_j, \cdots, k_i, \cdots, k_n). \tag{2.11}$$

Suppose there are $r-1$ numbers between k_i and k_j in the first list and consequently $r-1$ numbers between k_j and k_i in the second list. In computing $\pi(k_1, \cdots, k_i, \cdots, k_j, \cdots, k_n)$ we have $r-1$ terms of the form $(k_j - k_p)$ where the k_p are those numbers occurring in the list between k_i and k_j. Corresponding to these terms we have $r-1$ terms of the form $(k_p - k_j)$ in the computation of $\pi(k_1, \cdots, k_j, \cdots, k_i, \cdots, k_n)$. These differences produce a $(-1)^{r-1}$ in going from $\pi(k_1, \cdots, k_i, \cdots, k_j, \cdots, k_n)$ to $\pi(k_1, \cdots, k_j, \cdots, k_i, \cdots, k_n)$. We also have the $r-1$ terms $(k_p - k_i)$ in computing $\pi(k_1, \cdots, k_i \cdots, k_j, \cdots, k_n)$ and the $r-1$ terms, $(k_i - k_p)$ in computing $\pi(k_1, \cdots, k_j, \cdots, k_i, \cdots, k_n)$, producing another $(-1)^{r-1}$. Thus, in considering the differences in π, we see these terms just considered do not change the sign. However, we have $(k_j - k_i)$ in the first product and $(k_i - k_j)$ in the second and all other factors in the computation of π match up in the two computations so it follows $\pi(k_1, \cdots, k_i, \cdots, k_j, \cdots, k_n) = -\pi(k_1, \cdots, k_j, \cdots, k_i, \cdots, k_n)$ as claimed.

Corollary 2.32 *Suppose $(k_1, \cdots, k_n)$ is obtained by making p switches in the ordered list, $(1, \cdots, n)$. Then*

$$(-1)^p = sgn(k_1, \cdots, k_n). \tag{2.12}$$

Proof: We observe that $sgn(1, \cdots, n) = 1$ and according to Lemma 2.31, each time we switch two entries we multiply by (-1). Therefore, making the p switches, we obtain $(-1)^p = (-1)^p sgn(1, \cdots, n) = sgn(k_1, \cdots, k_n)$ as claimed.

We now are ready to define the determinant of an $n \times n$ matrix.

Definition 2.33 *Let $(a_{ij}) = A$ denote an $n \times n$ matrix. We define*

$$\det(A) \equiv \sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_n) a_{1k_1} \cdots a_{nk_n}$$

where the sum is taken over all ordered lists of numbers from $\{1, \cdots, n\}$. Note it suffices to take the sum over only those ordererd lists in which there are no repeats because if there are, we know $sgn(k_1, \cdots, k_n) = 0$.

Let A be an $n \times n$ matrix, $A = (a_{ij})$ and let $(r_1, \cdots, r_n)$ denote an ordered list of n numbers from $\{1, \cdots, n\}$. Let $A(r_1, \cdots, r_n)$ denote the matrix whose k^{th} row is the r_k row of the matrix, A. Thus

$$\det (A(r_1, \cdots, r_n)) = \sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_n) a_{r_1 k_1} \cdots a_{r_n k_n} \tag{2.13}$$

and

$$A(1, \cdots, n) = A.$$

Proposition 2.34 *Let $(r_1, \cdots, r_n)$ be an ordered list of numbers from $\{1, \cdots, n\}$. Then*

$$\begin{aligned} sgn(r_1, \cdots, r_n) \det (A) &= \sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_n) a_{r_1 k_1} \cdots a_{r_n k_n} && (2.14) \\ &= \det (A(r_1, \cdots, r_n)). && (2.15) \end{aligned}$$

In words, if we take the determinant of the matrix obtained by letting the p^{th} row be the r_p row of A, then the determinant of this modified matrix equals the expression on the left in (2.14).

Proof: Let $(1, \cdots, n) = (1, \cdots, r, \cdots s, \cdots, n)$ so $r < s$.

$$\det (A(1, \cdots, r, \cdots s, \cdots, n)) = \tag{2.16}$$

$$\sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_r, \cdots, k_s, \cdots, k_n) a_{1k_1} \cdots a_{rk_r} \cdots a_{sk_s} \cdots a_{nk_n}$$

$$= \sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_s, \cdots, k_r, \cdots, k_n) a_{1k_1} \cdots a_{rk_s} \cdots a_{sk_r} \cdots a_{nk_n}$$

$$= \sum_{(k_1, \cdots, k_n)} -sgn(k_1, \cdots, k_r, \cdots, k_s, \cdots, k_n) a_{1k_1} \cdots a_{rk_s} \cdots a_{sk_r} \cdots a_{nk_n}$$

$$= -\det (A(1, \cdots, s, \cdots, r, \cdots, n)). \tag{2.17}$$

Consequently,

$$\det (A(1, \cdots, s, \cdots, r, \cdots, n)) = -\det (A(1, \cdots, r, \cdots, s, \cdots, n)) = -\det (A)$$

Now letting $A(1, \cdots, s, \cdots, r, \cdots, n)$ play the role of A, and continuing in this way, we eventually arrive at the conclusion

$$\det (A(r_1, \cdots, r_n)) = (-1)^p \det (A)$$

where it took p switches to obtain$(r_1, \cdots, r_n)$ from $(1, \cdots, n)$. By Corollary 2.32 this implies

$$\det (A(r_1, \cdots, r_n)) = sgn(r_1, \cdots, r_n) \det (A)$$

and proves the proposition in the case when there are no repeated numbers in the ordered list, $(r_1, \cdots, r_n)$. However, if there is a repeat, say the r^{th} row equals the s^{th} row, then the reasoning of (2.16) -(2.17) shows that $A(r_1, \cdots, r_n) = 0$ and we also know that $sgn(r_1, \cdots, r_n) = 0$ so the formula holds in this case also.

Corollary 2.35 *We have the following formula for* $\det(A)$.

$$\det(A) = \frac{1}{n!} \sum_{(r_1,\cdots,r_n)} \sum_{(k_1,\cdots,k_n)} sgn(r_1,\cdots,r_n)\, sgn(k_1,\cdots,k_n)\, a_{r_1k_1}\cdots a_{r_nk_n}. \tag{2.18}$$

And also $\det(A^T) = \det(A)$ *where* A^T *is the transpose of* A. *Thus if* $A^T = (a_{ij}^T)$, *we have* $a_{ij}^T = a_{ji}$.

Proof: From Proposition 2.34, if the r_i are distinct,

$$\det(A) = \sum_{(k_1,\cdots,k_n)} sgn(r_1,\cdots,r_n)\, sgn(k_1,\cdots,k_n)\, a_{r_1k_1}\cdots a_{r_nk_n}.$$

Summing over all ordered lists, $(r_1,\cdots,r_n)$ where the r_i are distinct, (If the r_i are not distinct, we know $sgn(r_1,\cdots,r_n) = 0$ and so there is no contribution to the sum.) we obtain

$$n!\det(A) = \sum_{(r_1,\cdots,r_n)} \sum_{(k_1,\cdots,k_n)} sgn(r_1,\cdots,r_n)\, sgn(k_1,\cdots,k_n)\, a_{r_1k_1}\cdots a_{r_nk_n}.$$

This proves the corollary.

Corollary 2.36 *If we switch two rows or two columns in an* $n\times n$ *matrix,* A, *the determinant of the resulting matrix equals* (-1) *times the determinant of the original matrix. If* A *is an* $n\times n$ *matrix in which two rows are equal or two columns are equal then* $\det(A) = 0$.

Proof: By Proposition 2.34 when we switch two rows the determinant of the resulting matrix is (-1) times the determinant of the original matrix. By Corollary 2.35 the same holds for columns because the columns of the matrix equal the rows of the transposed matrix. Thus if A_1 is the matrix obtained from A by switching two columns, then

$$\det(A) = \det(A^T) = -\det(A_1^T) = -\det(A_1).$$

If A has two equal columns or two equal rows, then switching them results in the same matrix. Therefore, $\det(A) = -\det(A)$ and so $\det(A) = 0$.

Definition 2.37 *If* A *and* B *are* $n\times n$ *matrices,* $A = (a_{ij})$ *and* $B = (b_{ij})$, *we form the product,* $AB = (c_{ij})$ *by defining*

$$c_{ij} \equiv \sum_{k=1}^{n} a_{ik}b_{kj}.$$

This is just the usual rule for matrix multiplication.

One of the most important rules about determinants is that the determinant of a product equals the product of the determinants.

Theorem 2.38 *Let* A *and* B *be* $n\times n$ *matrices. Then* $\det(AB) = \det(A)\det(B)$.

Proof: We will denote by c_{ij} the ij^{th} entry of AB. Thus by Proposition 2.34,

$$\begin{aligned}
\det(AB) &= \sum_{(k_1,\cdots,k_n)} sgn(k_1,\cdots,k_n)\, c_{1k_1}\cdots c_{nk_n} \\
&= \sum_{(k_1,\cdots,k_n)} sgn(k_1,\cdots,k_n) \left(\sum_{r_1} a_{1r_1}b_{r_1k_1}\right)\cdots\left(\sum_{r_n} a_{nr_n}b_{r_nk_n}\right) \\
&= \sum_{(r_1\cdots,r_n)} \sum_{(k_1,\cdots,k_n)} sgn(k_1,\cdots,k_n)\, b_{r_1k_1}\cdots b_{r_nk_n}\,(a_{1r_1}\cdots a_{nr_n}) \\
&= \sum_{(r_1\cdots,r_n)} sgn(r_1\cdots r_n)\, a_{1r_1}\cdots a_{nr_n}\det(B) = \det(A)\det(B).
\end{aligned}$$

This proves the theorem.

In terms of the theory of determinants, arguably the most important idea is that of Laplace expansion along a row or a column.

Definition 2.39 *Let $A = (a_{ij})$ be an $n \times n$ matrix. Then we define a new matrix, $cof\,(A)$ by $cof\,(A) = (c_{ij})$ where to obtain c_{ij} we delete the i^{th} row and the j^{th} column of A, take the determinant of the $n-1 \times n-1$ matrix which results and then multiply this number by $(-1)^{i+j}$. The determinant of the $n-1 \times n-1$ matrix just described is called the ij^{th} minor of A. To make the formulas easier to remember, we shall write $cof\,(A)_{ij}$ for the ij^{th} entry of the cofactor matrix.*

The main result is the following monumentally important theorem. It states that you can expand an $n \times n$ matrix along any row or column. This is often taken as a definition in elementary courses but how anyone in their right mind could believe without a proof that you always get the same answer by expanding along any row or column is totally beyond my powers of comprehension.

Theorem 2.40 *Let A be an $n \times n$ matrix. Then*

$$\det(A) = \sum_{j=1}^{n} a_{ij} cof\,(A)_{ij} = \sum_{i=1}^{n} a_{ij} cof\,(A)_{ij}\,. \tag{2.19}$$

The first formula consists of expanding the determinant along the i^{th} row and the second expands the determinant along the j^{th} column.

Proof: We will prove this by using the definition and then doing a computation and verifying that we have what we want.

$$\det(A) = \sum_{(k_1,\cdots,k_n)} sgn\,(k_1,\cdots,k_r,\cdots,k_n)\, a_{1k_1}\cdots a_{rk_r}\cdots a_{nk_n}$$

$$= \sum_{k_r=1}^{n}\left(\sum_{(k_1,\cdots,k_r,\cdots,k_n)} sgn\,(k_1,\cdots,k_r,\cdots,k_n)\, a_{1k_1}\cdots a_{(r-1)k_{(r-1)}} a_{(r+1)k_{(r+1)}} a_{nk_n}\right) a_{rk_r}$$

$$= \sum_{j=1}^{n} (-1)^{r-1}\cdot$$

$$\left(\sum_{(k_1,\cdots,j,\cdots,k_n)} sgn\,(j,k_1,\cdots,k_{r-1},k_{r+1}\cdots,k_n)\, a_{1k_1}\cdots a_{(r-1)k_{(r-1)}} a_{(r+1)k_{(r+1)}} a_{nk_n}\right) a_{rj}. \tag{2.20}$$

We need to consider for fixed j the term

$$\sum_{(k_1,\cdots,j,\cdots,k_n)} sgn\,(j,k_1,\cdots,k_{r-1},k_{r+1}\cdots,k_n)\, a_{1k_1}\cdots a_{(r-1)k_{(r-1)}} a_{(r+1)k_{(r+1)}} a_{nk_n}. \tag{2.21}$$

We may assume all the indices in $(k_1,\cdots,j,\cdots,k_n)$ are distinct. We define $(l_1,\cdots,l_{n-1})$ as follows. If $k_\alpha < j$, then $l_\alpha \equiv k_\alpha$. If $k_\alpha > j$, then $l_\alpha \equiv k_\alpha - 1$. Thus every choice of the ordered list, $(k_1,\cdots,j,\cdots,k_n)$, corresponds to an ordered list, $(l_1,\cdots,l_{n-1})$ of indices from $\{1,\cdots,n-1\}$. Now define

$$b_{\alpha l_\alpha} \equiv \begin{cases} a_{\alpha k_\alpha} \text{ if } \alpha < r, \\ a_{(\alpha+1)k_\alpha} \text{ if } n-1 \geq \alpha > r \end{cases}$$

where here k_α corresponds to l_α as just described. Thus $(b_{\alpha\beta})$ is the $n-1 \times n-1$ matrix which results from deleting the r^{th} row and the j^{th} column. In computing

$$\pi(j, k_1, \cdots, k_{r-1}, k_{r+1} \cdots, k_n),$$

we note there are exactly $j-1$ of the k_i which are less than j. Therefore,

$$sgn(k_1, \cdots, k_{r-1}, k_{r+1} \cdots, k_n)(-1)^{j-1} = sgn(j, k_1, \cdots, k_{r-1}, k_{r+1} \cdots, k_n).$$

But it also follows from the definition that

$$sgn(k_1, \cdots, k_{r-1}, k_{r+1} \cdots, k_n) = sgn(l_1 \cdots, l_{n-1})$$

and so the term in (2.21) equals

$$(-1)^{j-1} \sum_{(l_1, \cdots, l_{n-1})} sgn(l_1, \cdots, l_{n-1}) b_{1l_1} \cdots b_{(n-1)l_{(n-1)}}$$

Using this in (2.20) we see

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} (-1)^{r-1} (-1)^{j-1} \left(\sum_{(l_1, \cdots, l_{n-1})} sgn(l_1, \cdots, l_{n-1}) b_{1l_1} \cdots b_{(n-1)l_{(n-1)}} \right) a_{rj} \\ &= \sum_{j=1}^{n} (-1)^{r+j} \left(\sum_{(l_1, \cdots, l_{n-1})} sgn(l_1, \cdots, l_{n-1}) b_{1l_1} \cdots b_{(n-1)l_{(n-1)}} \right) a_{rj} \\ &= \sum_{j=1}^{n} a_{rj} cof(A)_{rj} \end{aligned}$$

as claimed. Now to get the second half of (2.19), we can apply the first part to A^T and write for $A^T = \left(a_{ij}^T\right)$

$$\begin{aligned} \det(A) &= \det\left(A^T\right) = \sum_{j=1}^{n} a_{ij}^T cof\left(A^T\right)_{ij} \\ &= \sum_{j=1}^{n} a_{ji} cof(A)_{ji} = \sum_{i=1}^{n} a_{ij} cof(A)_{ij}. \end{aligned}$$

This proves the theorem. We leave it as an exercise to show that $cof\left(A^T\right)_{ij} = cof(A)_{ji}$.

Note that this gives us an easy way to write a formula for the inverse of an $n \times n$ matrix.

Definition 2.41 *We say an $n \times n$ matrix, A has an inverse, A^{-1} if and only if $AA^{-1} = A^{-1}A = I$ where $I = (\delta_{ij})$ for*

$$\delta_{ij} \equiv \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Theorem 2.42 *A^{-1} exists if and only if $\det(A) \neq 0$. If $\det(A) \neq 0$, then $A^{-1} = \left(a_{ij}^{-1}\right)$ where*

$$a_{ij}^{-1} = \det(A)^{-1} C_{ji}$$

for C_{ij} the ij^{th} cofactor of A.

Proof: By Theorem 2.40 and letting $(a_{ir}) = A$, if we assume $\det(A) \neq 0$,

$$\sum_{i=1}^{n} a_{ir} C_{ir} \det(A)^{-1} = \det(A)\det(A)^{-1} = 1.$$

Now we consider

$$\sum_{i=1}^{n} a_{ir} C_{ik} \det(A)^{-1}$$

when $k \neq r$. We replace the k^{th} column with the r^{th} column to obtain a matrix, B_k whose determinant equals zero by Corollary 2.36. However, expanding this matrix along the k^{th} column yields

$$0 = \det(B_k)\det(A)^{-1} = \sum_{i=1}^{n} a_{ir} C_{ik} \det(A)^{-1}$$

Summarizing,

$$\sum_{i=1}^{n} a_{ir} C_{ik} \det(A)^{-1} = \delta_{rk}.$$

Using the other formula in Theorem 2.40, we can also write using similar reasoning,

$$\sum_{j=1}^{n} a_{rj} C_{kj} \det(A)^{-1} = \delta_{rk}$$

This proves that if $\det(A) \neq 0$, then A^{-1} exists and if $A^{-1} = \left(a_{ij}^{-1}\right)$,

$$a_{ij}^{-1} = C_{ji} \det(A)^{-1}.$$

Now suppose A^{-1} exists. Then by Theorem 2.38,

$$1 = \det(I) = \det\left(AA^{-1}\right) = \det(A)\det\left(A^{-1}\right)$$

so $\det(A) \neq 0$. This proves the theorem.

This theorem says that to find the inverse, we can take the transpose of the cofactor matrix and divide by the determinant. The transpose of the cofactor matrix is called the adjugate or sometimes the classical adjoint of the matrix A. It is an abomination to call it the adjoint. The term, adjoint, should be reserved for something much more interesting which will be discussed later. In words, A^{-1} is equal to one over the determinant of A times the adjugate matrix of A.

In case we are solving a system of equations,

$$A\mathbf{x} = \mathbf{y}$$

for $\mathbf{x}$, it follows that if A^{-1} exists, we can write

$$\mathbf{x} = \left(A^{-1}A\right)\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{y}$$

thus solving the system. Now in the case that A^{-1} exists, we just presented a formula for A^{-1}. Using this formula, we see

$$x_i = \sum_{j=1}^{n} a_{ij}^{-1} y_j = \sum_{j=1}^{n} \frac{1}{\det(A)} cof(A)_{ji}\, y_j.$$

By the formula for the expansion of a determinant along a column,

$$x_i = \frac{1}{\det(A)} \det \begin{pmatrix} * & \cdots & y_1 & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & y_n & \cdots & * \end{pmatrix},$$

where here we have replaced the i^{th} column of A with the column vector, $(y_1 \cdots\cdot, y_n)^T$, taken its determinant and divided by $\det(A)$. This formula is known as Cramer's rule.

Definition 2.43 *We say a matrix M, is upper triangular if $M_{ij} = 0$ whenever $i > j$. Thus such a matrix equals zero below the main diagonal, the entries of the form M_{ii} as shown.*

$$\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * \end{pmatrix}$$

A lower triangular matrix is defined similarly as a matrix for which all entries above the main diagonal are equal to zero.

With this definition, we give the following corollary of Theorem 2.40.

Corollary 2.44 *Let M be an upper (lower) triangular matrix. Then $\det(M)$ is obtained by taking the product of the entries on the main diagonal.*

2.6 The characteristic polynomial

Definition 2.45 *Let A be an $n \times n$ matrix. The characteristic polynomial is defined as*

$$p_A(t) \equiv \det(tI - A).$$

A principal submatrix of A is one lying in the same set of k rows and columns and a principal minor is the determinant of a principal submatrix. There are $\binom{n}{k}$ principal minors of A. How do we get a typical principal submatrix? We pick k rows, say $r_1, \cdots, r_k$ and consider the $k \times k$ matrix which results from using exactly those entries of these k rows which are also in one of the $r_1, \cdots, r_k$ columns. We denote by $E_k(A)$ the sum of the principal $k \times k$ minors of A.

We write a formula for the characteristic polynomial in terms of the $E_k(A)$.

$$p_A(t) = \sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_n)(t\delta_{1k_1} - a_{1k_1}) \cdots (t\delta_{1k_n} - a_{1k_n})$$

Consider the terms which are multiplied by t^r. A typical such term would be

$$t^r(-1)^{n-r} \sum_{(k_1, \cdots, k_n)} sgn(k_1, \cdots, k_n) \delta_{m_1 k_{m_1}} \cdots \delta_{m_r k_{m_r}} a_{s_1 k_{s_1}} \cdots a_{s_{(n-r)} k_{s_{(n-r)}}} \tag{2.22}$$

where $\{m_1, \cdots, m_r, s_1, \cdots, s_{n-r}\} = \{1, \cdots, n\}$. From the definition of determinant, the sum in the above expression is the determinant of a matrix like

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

where the starred rows are simply the original rows of the matrix, A. Using the row operation which involves replacing a row with a multiple of another row added to itself, we can use the ones to zero out everything above them and below them, obtaining a modified matrix which has the same determinant (See Problem 2). In the given example this would result in a matrix of the form

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

and so the sum in (2.22) is just the principal minor corresponding to the subset $\{m_1, \cdots, m_r\}$ of $\{1, \cdots, n\}$. For each of the $\binom{n}{r}$ such choices, there is such a term equal to the principal minor determined in this way and so the sum of these equals the coefficient of the t^r term. Therefore, the coefficient of t^r equals $(-1)^{n-r} E_{n-r}(A)$. It follows

$$\begin{aligned} p_A(t) &= \sum_{r=0}^{n} t^r (-1)^{n-r} E_{n-r}(A) \\ &= (-1)^n E_n(A) + (-1)^{n-1} t E_{n-1}(A) + \cdots + (-1) t^{n-1} E_1(A) + t^n. \end{aligned}$$

Definition 2.46 *The solutions to $p_A(t) = 0$ are called the eigenvalues of A.*

We know also that

$$p_A(t) = \prod_{k=1}^{n} (t - \lambda_k)$$

where λ_k are the roots of the equation, $p_A(t) = 0$. (Note these might be complex numbers.) Therefore, expanding the above polynomial,

$$E_k(A) = S_k(\lambda_1, \cdots, \lambda_n)$$

where $S_k(\lambda_1, \cdots, \lambda_n)$, called the k^{th} elementary symmetric function of the numbers $\lambda_1, \cdots, \lambda_n$, is defined as the sum of all possible products of k elements of $\{\lambda_1, \cdots, \lambda_n\}$. Therefore,

$$p_A(t) = t^n - S_1(\lambda_1, \cdots, \lambda_n) t^{n-1} + S_2(\lambda_1, \cdots, \lambda_n) t^{n-2} + \cdots \pm S_n(\lambda_1, \cdots, \lambda_n).$$

A remarkable and profound theorem is the Cayley Hamilton theorem which states that every matrix satisfies its characteristic equation. We give a simple proof of this theorem using the following lemma.

Lemma 2.47 *Suppose for all $|\lambda|$ large enough, we have*

$$A_0 + A_1\lambda + \cdots + A_m\lambda^m = 0,$$

where the A_i are $n \times n$ matrices. Then each $A_i = 0$.

Proof: Multiply by λ^{-m} to obtain

$$A_0\lambda^{-m} + A_1\lambda^{-m+1} + \cdots + A_{m-1}\lambda^{-1} + A_m = 0.$$

Now let $|\lambda| \to \infty$ to obtain $A_m = 0$. With this, multiply by λ to obtain

$$A_0\lambda^{-m+1} + A_1\lambda^{-m+2} + \cdots + A_{m-1} = 0.$$

Now let $|\lambda| \to \infty$ to obtain $A_{m-1} = 0$. Continue multiplying by λ and letting $\lambda \to \infty$ to obtain that all the $A_i = 0$. This proves the lemma.

With the lemma, we have the following simple corollary.

Corollary 2.48 *Let A_i and B_i be $n \times n$ matrices and suppose*

$$A_0 + A_1\lambda + \cdots + A_m\lambda^m = B_0 + B_1\lambda + \cdots + B_m\lambda^m$$

for all $|\lambda|$ large enough. Then $A_i = B_i$ for all i.

Proof: Subtract and use the result of the lemma.
With this preparation, we can now give an easy proof of the Cayley Hamilton theorem.

Theorem 2.49 *Let A be an $n \times n$ matrix and let $p(\lambda) \equiv \det(\lambda I - A)$ be the characteristic polynomial. Then $p(A) = 0$.*

Proof: Let $C(\lambda)$ equal the transpose of the cofactor matrix of $(\lambda I - A)$ for $|\lambda|$ large. (If $|\lambda|$ is large enough, then λ cannot be in the finite list of eigenvalues of A and so for such λ, $(\lambda I - A)^{-1}$ exists.) Therefore, by Theorem 2.42 we may write

$$C(\lambda) = p(\lambda)(\lambda I - A)^{-1}.$$

Note that each entry in $C(\lambda)$ is a polynomial in λ having degree no more than $n-1$. Therefore, collecting the terms, we may write

$$C(\lambda) = C_0 + C_1\lambda + \cdots + C_{n-1}\lambda^{n-1}$$

for C_j some $n \times n$ matrix. It follows that for all $|\lambda|$ large enough,

$$(A - \lambda I)\left(C_0 + C_1\lambda + \cdots + C_{n-1}\lambda^{n-1}\right) = p(\lambda) I$$

and so we are in the situation of Corollary 2.48. It follows the matrix coefficients corresponding to equal powers of λ are equal on both sides of this equation.Therefore, we may replace λ with A and the two will be equal. Thus

$$0 = (A - A)\left(C_0 + C_1 A + \cdots + C_{n-1}A^{n-1}\right) = p(A) I = p(A).$$

This proves the Cayley Hamilton theorem.

2.7 The rank of a matrix

Definition 2.50 *Let A be an $m \times n$ matrix. Then the row rank is the dimension of the span of the rows in $\mathbb{F}^n$, the column rank is the dimension of the span of the columns, and the determinant rank equals r where r is the largest number such that some $r \times r$ submatrix of A has a non zero determinant.Note the column rank of A is nothing more than the dimension of $A(\mathbb{F}^n)$.*

Theorem 2.51 *The determinant rank, row rank, and column rank coincide.*

Proof: Suppose the determinant rank of $A = (a_{ij})$ equals r. First note that if rows and columns are interchanged, the row, column, and determinant ranks of the modified matrix are unchanged. Thus we may assume without loss of generality that there is an $r \times r$ matrix in the upper left corner of the matrix which has non zero determinant. Consider the matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1p} \\ \vdots & & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{rp} \\ a_{l1} & \cdots & a_{lr} & a_{lp} \end{pmatrix}$$

where we will denote by C the $r \times r$ matrix which has non zero determinant. The above matrix has determinant equal to zero. There are two cases to consider in verifying this claim. First, suppose $p > r$. Then the claim follows from the assumption that A has determinant rank r. On the other hand, if $p < r$, then the determinant is zero because there are two identical columns. Expand the determinant along the last column and divide by $\det(C)$ to obtain

$$a_{lp} = -\sum_{i=1}^{r} \frac{C_{ip}}{\det(C)} a_{ip},$$

where C_{ip} is the cofactor of a_{ip}. Now note that C_{ip} does not depend on p. Therefore the above sum is of the form

$$a_{lp} = \sum_{i=1}^{r} m_i a_{ip}$$

which shows the lth row is a linear combination of the first r rows of A. Thus the first r rows of A are linearly independent and span the row space so the row rank equals r. It follows from this that

$$\text{column rank of } A = \text{ row rank of } A^T =$$

$$= \text{ determinant rank of } A^T \;=\; \text{determinant rank of } A = \text{ row rank of } A.$$

This proves the theorem.

2.8 Exercises

1. Let $A \in \mathcal{L}(V, V)$ where V is a vector space of dimension n. Show using the fundamental theorem of algebra which states that every non constant polynomial has a zero in the complex numbers, that A has an eigenvalue and eigenvector. Recall that $(\lambda, \mathbf{v})$ is an eigen pair if $\mathbf{v} \neq \mathbf{0}$ and $(A - \lambda I)(\mathbf{v}) = \mathbf{0}$.

2. Show that if we replace a row (column) of an $n \times n$ matrix A with itself added to some multiple of another row (column) then the new matrix has the same determinant as the original one.

3. Let A be an $n \times n$ matrix and let

$$A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, \ |\mathbf{v}_1| = 1.$$

   Show there exists an orthonormal basis, $\{\mathbf{v}_1, \cdot\cdot\cdot, \mathbf{v}_n\}$ for $\mathbb{F}^n$. Let Q_0 be a matrix whose i^{th} column is $\mathbf{v}_i$. Show $Q_0^* A Q_0$ is of the form

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

   where A_1 is an $n-1 \times n-1$ matrix.

4. Using the result of problem 3, show there exists an orthogonal matrix $\widetilde{Q}_1$ such that $\widetilde{Q}_1^* A_1 \, \widetilde{Q}_1$ is of the form

$$\begin{pmatrix} \lambda_2 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix}.$$

Now let Q_1 be the $n \times n$ matrix of the form

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widetilde{Q}_1 \end{pmatrix}.$$

Show $Q_1^* Q_0^* A Q_0 Q_1$ is of the form

$$\begin{pmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & & & \\ \vdots & \vdots & & A_2 & \\ 0 & 0 & & & \end{pmatrix}$$

where A_2 is an $n-2 \times n-2$ matrix. Continuing in this way, show there exists an orthogonal matrix Q such that

$$Q^* A Q = T$$

where T is upper triangular. This is called the Schur form of the matrix.

5. Let A be an $m \times n$ matrix. Show the column rank of A equals the column rank of A^*A. Next verify column rank of A^*A is no larger than column rank of A^*. Next justify the following inequality to conclude the column rank of A equals the column rank of A^*.

$$\text{rank } (A) = \text{rank } (A^*A) \leq \text{rank } (A^*) \leq$$

$$= \text{rank } (AA^*) \leq \text{rank } (A).$$

Hint: Start with an orthonormal basis, $\{A\mathbf{x}_j\}_{j=1}^r$ of $A(\mathbb{F}^n)$ and verify $\{A^*A\mathbf{x}_j\}_{j=1}^r$ is a basis for $A^*A(\mathbb{F}^n)$.

6. Show the λ_i on the main diagonal of T in problem 4 are the eigenvalues of A.

7. We say A is normal if

$$A^*A = AA^*.$$

Show that if $A^* = A$, then A is normal. Show that if A is normal and Q is an orthogonal matrix, then Q^*AQ is also normal. Show that if T is upper triangular and normal, then T is a diagonal matrix. Conclude the Shur form of every normal matrix is diagonal.

8. If A is such that there exists an orthogonal matrix, Q such that

$$Q^*AQ = \text{ diagonal matrix},$$

is it necessary that A be normal? (We know from problem 7 that if A is normal, such an orthogonal matrix exists.)

General topology

This chapter is a brief introduction to general topology. Topological spaces consist of a set and a subset of the set of all subsets of this set called the open sets or topology which satisfy certain axioms. Like other areas in mathematics the abstraction inherent in this approach is an attempt to unify many different useful examples into one general theory.

For example, consider $\mathbb{R}^n$ with the usual norm given by

$$|\mathbf{x}| \equiv \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}.$$

We say a set U in $\mathbb{R}^n$ is an open set if every point of U is an "interior" point which means that if $\mathbf{x} \in U$, there exists $\delta > 0$ such that if $|\mathbf{y} - \mathbf{x}| < \delta$, then $\mathbf{y} \in U$. It is easy to see that with this definition of open sets, the axioms (3.1) - (3.2) given below are satisfied if τ is the collection of open sets as just described. There are many other sets of interest besides $\mathbb{R}^n$ however, and the appropriate definition of "open set" may be very different and yet the collection of open sets may still satisfy these axioms. By abstracting the concept of open sets, we can unify many different examples. Here is the definition of a general topological space.

Let X be a set and let τ be a collection of subsets of X satisfying

$$\emptyset \in \tau, \ X \in \tau, \tag{3.1}$$

$$\text{If } \mathcal{C} \subseteq \tau, \text{ then } \cup \mathcal{C} \in \tau$$

$$\text{If } A, B \in \tau, \text{ then } A \cap B \in \tau. \tag{3.2}$$

Definition 3.1 *A set X together with such a collection of its subsets satisfying (3.1)-(3.2) is called a topological space. τ is called the topology or set of open sets of X. Note $\tau \subseteq \mathcal{P}(X)$, the set of all subsets of X, also called the power set.*

Definition 3.2 *A subset $\mathcal{B}$ of τ is called a basis for τ if whenever $p \in U \in \tau$, there exists a set $B \in \mathcal{B}$ such that $p \in B \subseteq U$. The elements of $\mathcal{B}$ are called basic open sets.*

The preceding definition implies that every open set (element of τ) may be written as a union of basic open sets (elements of $\mathcal{B}$). This brings up an interesting and important question. If a collection of subsets $\mathcal{B}$ of a set X is specified, does there exist a topology τ for X satisfying (3.1)-(3.2) such that $\mathcal{B}$ is a basis for τ?

Theorem 3.3 *Let X be a set and let $\mathcal{B}$ be a set of subsets of X. Then $\mathcal{B}$ is a basis for a topology τ if and only if whenever $p \in B \cap C$ for $B, C \in \mathcal{B}$, there exists $D \in \mathcal{B}$ such that $p \in D \subseteq C \cap B$ and $\cup \mathcal{B} = X$. In this case τ consists of all unions of subsets of $\mathcal{B}$.*

Proof: The only if part is left to the reader. Let τ consist of all unions of sets of $\mathcal{B}$ and suppose $\mathcal{B}$ satisfies the conditions of the proposition. Then $\emptyset \in \tau$ because $\emptyset \subseteq \mathcal{B}$. $X \in \tau$ because $\cup\mathcal{B} = X$ by assumption. If $\mathcal{C} \subseteq \tau$ then clearly $\cup\mathcal{C} \in \tau$. Now suppose $A, B \in \tau$, $A = \cup\mathcal{S}$, $B = \cup\mathcal{R}$, $\mathcal{S}, \mathcal{R} \subseteq \mathcal{B}$. We need to show $A \cap B \in \tau$. If $A \cap B = \emptyset$, we are done. Suppose $p \in A \cap B$. Then $p \in S \cap R$ where $S \in \mathcal{S}$, $R \in \mathcal{R}$. Hence there exists $U \in \mathcal{B}$ such that $p \in U \subseteq S \cap R$. It follows, since $p \in A \cap B$ was arbitrary, that $A \cap B =$ union of sets of $\mathcal{B}$. Thus $A \cap B \in \tau$. Hence τ satisfies (3.1)-(3.2).

Definition 3.4 *A topological space is said to be Hausdorff if whenever p and q are distinct points of X, there exist disjoint open sets U, V such that $p \in U, q \in V$.*

Hausdorff

Definition 3.5 *A subset of a topological space is said to be closed if its complement is open. Let p be a point of X and let $E \subseteq X$. Then p is said to be a limit point of E if every open set containing p contains a point of E distinct from p.*

Theorem 3.6 *A subset, E, of X is closed if and only if it contains all its limit points.*

Proof: Suppose first that E is closed and let x be a limit point of E. We need to show $x \in E$. If $x \notin E$, then E^C is an open set containing x which contains no points of E, a contradiction. Thus $x \in E$. Now suppose E contains all its limit points. We need to show the complement of E is open. But if $x \in E^C$, then x is not a limit point of E and so there exists an open set, U containing x such that U contains no point of E other than x. Since $x \notin E$, it follows that $x \in U \subseteq E^C$ which implies E^C is an open set.

Theorem 3.7 *If (X, τ) is a Hausdorff space and if $p \in X$, then $\{p\}$ is a closed set.*

Proof: If $x \neq p$, there exist open sets U and V such that $x \in U, p \in V$ and $U \cap V = \emptyset$. Therefore, $\{p\}^C$ is an open set so $\{p\}$ is closed.

Note that the Hausdorff axiom was stronger than needed in order to draw the conclusion of the last theorem. In fact it would have been enough to assume that if $x \neq y$, then there exists an open set containing x which does not intersect y.

Definition 3.8 *A topological space (X, τ) is said to be regular if whenever C is a closed set and p is a point not in C, then there exist disjoint open sets U and V such that $p \in U$, $C \subseteq V$. The topological space, (X, τ) is said to be normal if whenever C and K are disjoint closed sets, there exist disjoint open sets U and V such that $C \subseteq U$, $K \subseteq V$.*

Regular

Normal

Definition 3.9 *Let E be a subset of X. $\bar{E}$ is defined to be the smallest closed set containing E. Note that this is well defined since X is closed and the intersection of any collection of closed sets is closed.*

Theorem 3.10 $\overline{E} = E \cup \{$*limit points of* $E\}$.

Proof: Let $x \in \overline{E}$ and suppose that $x \notin E$. If x is not a limit point either, then there exists an open set, U,containing x which does not intersect E. But then U^C is a closed set which contains E which does not contain x, contrary to the definition that $\overline{E}$ is the intersection of all closed sets containing E. Therefore, x must be a limit point of E after all.

Now $E \subseteq \overline{E}$ so suppose x is a limit point of E. We need to show $x \in \overline{E}$. If H is a closed set containing E, which does not contain x, then H^C is an open set containing x which contains no points of E other than x negating the assumption that x is a limit point of E.

Definition 3.11 *Let X be a set and let $d : X \times X \to [0, \infty)$ satisfy*

$$d(x, y) = d(y, x), \tag{3.3}$$

$$d(x, y) + d(y, z) \geq d(x, z), \ (\textit{triangle inequality})$$

$$d(x, y) = 0 \textit{ if and only if } x = y. \tag{3.4}$$

Such a function is called a metric. For $r \in [0, \infty)$ and $x \in X$, define

$$B(x, r) = \{y \in X : d(x, y) < r\}$$

This may also be denoted by $N(x, r)$.

Definition 3.12 *A topological space (X, τ) is called a metric space if there exists a metric, d, such that the sets $\{B(x, r), x \in X, \ r > 0\}$ form a basis for τ. We write (X, d) for the metric space.*

Theorem 3.13 *Suppose X is a set and d satisfies (3.3)-(3.4). Then the sets $\{B(x, r) : r > 0, \ x \in X\}$ form a basis for a topology on X.*

Proof: We observe that the union of these balls includes the whole space, X. We need to verify the condition concerning the intersection of two basic sets. Let $p \in B(x, r_1) \cap B(z, r_2)$. Consider

$$r \equiv \min(r_1 - d(x, p), r_2 - d(z, p))$$

and suppose $y \in B(p, r)$. Then

$$d(y, x) \leq d(y, p) + d(p, x) < r_1 - d(x, p) + d(x, p) = r_1$$

and so $B(p, r) \subseteq B(x, r_1)$. By similar reasoning, $B(p, r) \subseteq B(z, r_2)$. This verifies the conditions for this set of balls to be the basis for some topology.

Theorem 3.14 *If (X, τ) is a metric space, then (X, τ) is Hausdorff, regular, and normal.*

Proof: It is obvious that any metric space is Hausdorff. Since each point is a closed set, it suffices to verify any metric space is normal. Let H and K be two disjoint closed nonempty sets. For each $h \in H$, there exists $r_h > 0$ such that $B(h, r_h) \cap K = \emptyset$ because K is closed. Similarly, for each $k \in K$ there exists $r_k > 0$ such that $B(k, r_k) \cap H = \emptyset$. Now let

$$U \equiv \cup \{B(h, r_h/2) : h \in H\}, \ V \equiv \cup \{B(k, r_k/2) : k \in K\}.$$

then these open sets contain H and K respectively and have empty intersection for if $x \in U \cap V$, then $x \in B(h, r_h/2) \cap B(k, r_k/2)$ for some $h \in H$ and $k \in K$. Suppose $r_h \geq r_k$. Then

$$d(h, k) \leq d(h, x) + d(x, k) < r_h,$$

a contradiction to $B(h, r_h) \cap K = \emptyset$. If $r_k \geq r_h$, the argument is similar. This proves the theorem.

Definition 3.15 *A metric space is said to be separable if there is a countable dense subset of the space. This means there exists $D = \{p_i\}_{i=1}^{\infty}$ such that for all x and $r > 0$, $B(x,r) \cap D \neq \emptyset$.*

Definition 3.16 *A topological space is said to be completely separable if it has a countable basis for the topology.*

Theorem 3.17 *A metric space is separable if and only if it is completely separable.*

Proof: If the metric space has a countable basis for the topology, pick a point from each of the basic open sets to get a countable dense subset of the metric space.

Now suppose the metric space, (X, d), has a countable dense subset, D. Let $\mathcal{B}$ denote all balls having centers in D which have positive rational radii. We will show this is a basis for the topology. It is clear it is a countable set. Let U be any open set and let $z \in U$. Then there exists $r > 0$ such that $B(z,r) \subseteq U$. In $B(z, r/3)$ pick a point from D, x. Now let r_1 be a positive rational number in the interval $(r/3, 2r/3)$ and consider the set from $\mathcal{B}$, $B(x, r_1)$. If $y \in B(x, r_1)$ then

$$d(y,z) \leq d(y,x) + d(x,z) < r_1 + r/3 < 2r/3 + r/3 = r.$$

Thus $B(x, r_1)$ contains z and is contained in U. This shows, since z is an arbitrary point of U that U is the union of a subset of $\mathcal{B}$.

We already discussed Cauchy sequences in the context of $\mathbb{R}^p$ but the concept makes perfectly good sense in any metric space.

Definition 3.18 *A sequence $\{p_n\}_{n=1}^{\infty}$ in a metric space is called a Cauchy sequence if for every $\varepsilon > 0$ there exists N such that $d(p_n, p_m) < \varepsilon$ whenever $n, m > N$. A metric space is called complete if every Cauchy sequence converges to some element of the metric space.*

Example 3.19 $\mathbb{R}^n$ *and* $\mathbb{C}^n$ *are complete metric spaces for the metric defined by* $d(\mathbf{x},\mathbf{y}) \equiv |\mathbf{x}-\mathbf{y}| \equiv (\sum_{i=1}^{n} |x_i - y_i|^2)^{1/2}$.

Not all topological spaces are metric spaces and so the traditional $\epsilon - \delta$ definition of continuity must be modified for more general settings. The following definition does this for general topological spaces.

Definition 3.20 *Let (X, τ) and (Y, η) be two topological spaces and let $f : X \to Y$. We say f is continuous at $x \in X$ if whenever V is an open set of Y containing $f(x)$, there exists an open set $U \in \tau$ such that $x \in U$ and $f(U) \subseteq V$. We say that f is continuous if $f^{-1}(V) \in \tau$ whenever $V \in \eta$.*

Definition 3.21 *Let (X, τ) and (Y, η) be two topological spaces. $X \times Y$ is the Cartesian product. ($X \times Y = \{(x,y) : x \in X,\ y \in Y\}$). We can define a product topology as follows. Let $\mathcal{B} = \{(A \times B) : A \in \tau,\ B \in \eta\}$. $\mathcal{B}$ is a basis for the product topology.*

Theorem 3.22 *$\mathcal{B}$ defined above is a basis satisfying the conditions of Theorem 3.3.*

More generally we have the following definition which considers any finite Cartesian product of topological spaces.

Definition 3.23 *If (X_i, τ_i) is a topological space, we make $\prod_{i=1}^{n} X_i$ into a topological space by letting a basis be $\prod_{i=1}^{n} A_i$ where $A_i \in \tau_i$.*

Theorem 3.24 *Definition 3.23 yields a basis for a topology.*

The proof of this theorem is almost immediate from the definition and is left for the reader.

The definition of compactness is also considered for a general topological space. This is given next.

Definition 3.25 *A subset, E, of a topological space (X, τ) is said to be compact if whenever $\mathcal{C} \subseteq \tau$ and $E \subseteq \cup \mathcal{C}$, there exists a finite subset of $\mathcal{C}, \{U_1 \cdots U_n\}$, such that $E \subseteq \cup_{i=1}^n U_i$. (Every open covering admits a finite subcovering.) We say E is precompact if $\overline{E}$ is compact. A topological space is called locally compact if it has a basis $\mathcal{B}$, with the property that $\overline{B}$ is compact for each $B \in \mathcal{B}$. Thus the topological space is locally compact if it has a basis of precompact open sets.*

In general topological spaces there may be no concept of "bounded". Even if there is, closed and bounded is not necessarily the same as compactness. However, we can say that in any Hausdorff space every compact set must be a closed set.

Theorem 3.26 *If (X, τ) is a Hausdorff space, then every compact subset must also be a closed set.*

Proof: Suppose $p \notin K$. For each $x \in X$, there exist open sets, U_x and V_x such that

$$x \in U_x, \ p \in V_x,$$

and

$$U_x \cap V_x = \emptyset.$$

Since K is assumed to be compact, there are finitely many of these sets, $U_{x_1}, \cdots, U_{x_m}$ which cover K. Then let $V \equiv \cap_{i=1}^m V_{x_i}$. It follows that V is an open set containing p which has empty intersection with each of the U_{x_i}. Consequently, V contains no points of K and is therefore not a limit point. This proves the theorem.

Lemma 3.27 *Let (X, τ) be a topological space and let $\mathcal{B}$ be a basis for τ. Then K is compact if and only if every open cover of basic open sets admits a finite subcover.*

The proof follows directly from the definition and is left to the reader. A very important property enjoyed by a collection of compact sets is the property that if it can be shown that any finite intersection of this collection has non empty intersection, then it can be concluded that the intersection of the whole collection has non empty intersection.

Definition 3.28 *If every finite subset of a collection of sets has nonempty intersection, we say the collection has the finite intersection property.*

Theorem 3.29 *Let $\mathcal{K}$ be a set whose elements are compact subsets of a Hausdorff topological space, (X, τ). Suppose $\mathcal{K}$ has the finite intersection property. Then $\emptyset \neq \cap \mathcal{K}$.*

Proof: Suppose to the contrary that $\emptyset = \cap \mathcal{K}$. Then consider

$$\mathcal{C} \equiv \left\{K^C : K \in \mathcal{K}\right\}.$$

It follows $\mathcal{C}$ is an open cover of K_0 where K_0 is any particular element of $\mathcal{K}$. But then there are finitely many $K \in \mathcal{K}$, $K_1, \cdots, K_r$ such that $K_0 \subseteq \cup_{i=1}^r K_i^C$ implying that $\cap_{i=0}^r K_i = \emptyset$, contradicting the finite intersection property.

It is sometimes important to consider the Cartesian product of compact sets. The following is a simple example of the sort of theorem which holds when this is done.

Theorem 3.30 *Let X and Y be topological spaces, and K_1, K_2 be compact sets in X and Y respectively. Then $K_1 \times K_2$ is compact in the topological space $X \times Y$.*

Proof: Let $\mathcal{C}$ be an open cover of $K_1 \times K_2$ of sets $A \times B$ where A and B are open sets. Thus $\mathcal{C}$ is a open cover of basic open sets. For $y \in Y$, define

$$\mathcal{C}_y = \{A \times B \in \mathcal{C} : y \in B\}, \ \mathcal{D}_y = \{A : A \times B \in \mathcal{C}_y\}$$

Claim: $\mathcal{D}_y$ covers K_1.

Proof: Let $x \in K_1$. Then $(x, y) \in K_1 \times K_2$ so $(x, y) \in A \times B \in \mathcal{C}$. Therefore $A \times B \in \mathcal{C}_y$ and so $x \in A \in \mathcal{D}_y$.

Since K_1 is compact,

$$\{A_1, \cdots, A_{n(y)}\} \subseteq \mathcal{D}_y$$

covers K_1. Let

$$B_y = \cap_{i=1}^{n(y)} B_i$$

Thus $\{A_1, \cdots, A_{n(y)}\}$ covers K_1 and $A_i \times B_y \subseteq A_i \times B_i \in \mathcal{C}_y$.

Since K_2 is compact, there is a finite list of elements of K_2, $y_1, \cdots, y_r$ such that

$$\{B_{y_1}, \cdots, B_{y_r}\}$$

covers K_2. Consider

$$\{A_i \times B_{y_l}\}_{i=1}^{n(y_l)}{}_{l=1}^{r}.$$

If $(x, y) \in K_1 \times K_2$, then $y \in B_{y_j}$ for some $j \in \{1, \cdots, r\}$. Then $x \in A_i$ for some $i \in \{1, \cdots, n(y_j)\}$. Hence $(x, y) \in A_i \times B_{y_j}$. Each of the sets $A_i \times B_{y_j}$ is contained in some set of $\mathcal{C}$ and so this proves the theorem.

Another topic which is of considerable interest in general topology and turns out to be a very useful concept in analysis as well is the concept of a subbasis.

Definition 3.31 *$\mathcal{S} \subseteq \tau$ is called a subbasis for the topology τ if the set $\mathcal{B}$ of finite intersections of sets of $\mathcal{S}$ is a basis for the topology, τ.*

Recall that the compact sets in $\mathbb{R}^n$ with the usual topology are exactly those that are closed and bounded. We will have use of the following simple result in the following chapters.

Theorem 3.32 *Let U be an open set in $\mathbb{R}^n$. Then there exists a sequence of open sets, $\{U_i\}$ satisfying*

$$\cdots U_i \subseteq \overline{U_i} \subseteq U_{i+1} \cdots$$

and

$$U = \cup_{i=1}^{\infty} U_i.$$

Proof: The following lemma will be interesting for its own sake and in addition to this, is exactly what is needed for the proof of this theorem.

Lemma 3.33 *Let S be any nonempty subset of a metric space, (X, d) and define*

$$dist\,(x,S) \equiv \inf\{d\,(x, s) : s \in S\}.$$

Then the mapping, $x \to dist\,(x, S)$ satisfies

$$|dist\,(y, S) - dist\,(x, S)| \leq d\,(x, y).$$

Proof of the lemma: One of $dist\,(y, S)\,, dist\,(x, S)$ is larger than or equal to the other. Assume without loss of generality that it is $dist\,(y, S)$. Choose $s_1 \in S$ such that

$$dist\,(x, S) + \epsilon > d\,(x, s_1)$$

Then

$$|dist\,(y,S) - dist\,(x,S)| = dist\,(y,S) - dist\,(x,S) \leq$$

$$d\,(y,s_1) - d\,(x,s_1) + \epsilon \leq d\,(x,y) + d\,(x,s_1) - d\,(x,s_1) + \epsilon$$

$$= d\,(x,y) + \epsilon.$$

Since ϵ is arbitrary, this proves the lemma.

If $U = \mathbb{R}^n$ it is clear that $U = \cup_{i=1}^{\infty} B\,(\mathbf{0}, i)$ and so, letting $U_i = B\,(\mathbf{0}, i)$,

$$B\,(\mathbf{0}, i) = \{\mathbf{x} \in \mathbb{R}^n : dist\,(\mathbf{x}, \{\mathbf{0}\}) < i\}$$

and by continuity of $dist\,(\cdot, \{\mathbf{0}\})$,

$$\overline{B\,(\mathbf{0}, i)} = \{\mathbf{x} \in \mathbb{R}^n : dist\,(\mathbf{x}, \{\mathbf{0}\}) \leq i\}\,.$$

Therefore, the Heine Borel theorem applies and we see the theorem is true in this case.

Now we use this lemma to finish the proof in the case where U is not all of $\mathbb{R}^n$. Since $\mathbf{x} \to dist\left(\mathbf{x}, U^C\right)$ is continuous, the set,

$$U_i \equiv \left\{\mathbf{x} \in U : dist\left(\mathbf{x}, U^C\right) > \frac{1}{i} \text{ and } |\mathbf{x}| < i\right\},$$

is an open set. Also $U = \cup_{i=1}^{\infty} U_i$ and these sets are increasing. By the lemma,

$$\overline{U_i} = \left\{\mathbf{x} \in U : dist\left(\mathbf{x}, U^C\right) \geq \frac{1}{i} \text{ and } |\mathbf{x}| \leq i\right\},$$

a compact set by the Heine Borel theorem and also, $\cdots U_i \subseteq \overline{U_i} \subseteq U_{i+1} \cdots$.

3.1 Compactness in metric space

Many existence theorems in analysis depend on some set being compact. Therefore, it is important to be able to identify compact sets. The purpose of this section is to describe compact sets in a metric space.

Definition 3.34 *In any metric space, we say a set E is totally bounded if for every $\epsilon > 0$ there exists a finite set of points $\{x_1, \cdot \cdot \cdot, x_n\}$ such that*

$$E \subseteq \cup_{i=1}^{n} B\,(x_i, \epsilon).$$

This finite set of points is called an ϵ net.

The following proposition tells which sets in a metric space are compact.

Proposition 3.35 *Let (X, d) be a metric space. Then the following are equivalent.*

$$(X, d) \text{ is compact,} \tag{3.5}$$

$$(X, d) \text{ is sequentially compact,} \tag{3.6}$$

$$(X, d) \text{ is complete and totally bounded.} \tag{3.7}$$

Recall that X is "sequentially compact" means every sequence has a convergent subsequence converging so an element of X.

Proof: Suppose (3.5) and let $\{x_k\}$ be a sequence. Suppose $\{x_k\}$ has no convergent subsequence. If this is so, then $\{x_k\}$ has no limit point and no value of the sequence is repeated more than finitely many times. Thus the set

$$C_n = \cup \{x_k : k \geq n\}$$

is a closed set and if

$$U_n = C_n^C,$$

then

$$X = \cup_{n=1}^{\infty} U_n$$

but there is no finite subcovering, contradicting compactness of (X, d).

Now suppose (3.6) and let $\{x_n\}$ be a Cauchy sequence. Then $x_{n_k} \to x$ for some subsequence. Let $\epsilon > 0$ be given. Let n_0 be such that if $m, n \geq n_0$, then $d(x_n, x_m) < \frac{\epsilon}{2}$ and let l be such that if $k \geq l$ then $d(x_{n_k}, x) < \frac{\epsilon}{2}$. Let $n_1 > \max(n_l, n_0)$. If $n > n_1$, let $k > l$ and $n_k > n_0$.

$$\begin{aligned} d(x_n, x) &\leq d(x_n, x_{n_k}) + d(x_{n_k}, x) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Thus $\{x_n\}$ converges to x and this shows (X, d) is complete. If (X, d) is not totally bounded, then there exists $\epsilon > 0$ for which there is no ϵ net. Hence there exists a sequence $\{x_k\}$ with $d(x_k, x_l) \geq \epsilon$ for all $l \neq k$. This contradicts (3.6) because this is a sequence having no convergent subsequence. This shows (3.6) implies (3.7).

Now suppose (3.7). We show this implies (3.6). Let $\{p_n\}$ be a sequence and let $\{x_i^n\}_{i=1}^{m_n}$ be a 2^{-n} net for $n = 1, 2, \cdots$. Let

$$B_n \equiv B\left(x_{i_n}^n, 2^{-n}\right)$$

be such that B_n contains p_k for infinitely many values of k and $B_n \cap B_{n+1} \neq \emptyset$. Let p_{n_k} be a subsequence having

$$p_{n_k} \in B_k.$$

Then if $k \geq l$,

$$\begin{aligned} d(p_{n_k}, p_{n_l}) &\leq \sum_{i=l}^{k-1} d\left(p_{n_{i+1}}, p_{n_i}\right) \\ &< \sum_{i=l}^{k-1} 2^{-(i-1)} < 2^{-(l-2)}. \end{aligned}$$

Consequently $\{p_{n_k}\}$ is a Cauchy sequence. Hence it converges. This proves (3.6).

Now suppose (3.6) and (3.7). Let D_n be a n^{-1} net for $n = 1, 2, \cdots$ and let

$$D = \cup_{n=1}^{\infty} D_n.$$

Thus D is a countable dense subset of (X, d). The set of balls

$$\mathcal{B} = \{B(q, r) : q \in D,\ r \in Q \cap (0, \infty)\}$$

is a countable basis for (X, d). To see this, let $p \in B(x, \epsilon)$ and choose $r \in Q \cap (0, \infty)$ such that

$$\epsilon - d(p, x) > 2r.$$

Let $q \in B(p, r) \cap D$. If $y \in B(q, r)$, then

$$\begin{aligned} d(y, x) &\leq d(y, q) + d(q, p) + d(p, x) \\ &< r + r + \epsilon - 2r = \epsilon. \end{aligned}$$

Hence $p \in B(q, r) \subseteq B(x, \epsilon)$ and this shows each ball is the union of balls of $\mathcal{B}$. Now suppose $\mathcal{C}$ is any open cover of X. Let $\widetilde{\mathcal{B}}$ denote the balls of $\mathcal{B}$ which are contained in some set of $\mathcal{C}$. Thus

$$\cup \widetilde{\mathcal{B}} = X.$$

For each $B \in \widetilde{\mathcal{B}}$, pick $U \in \mathcal{C}$ such that $U \supseteq B$. Let $\widetilde{\mathcal{C}}$ be the resulting countable collection of sets. Then $\widetilde{\mathcal{C}}$ is a countable open cover of X. Say $\widetilde{\mathcal{C}} = \{U_n\}_{n=1}^{\infty}$. If $\mathcal{C}$ admits no finite subcover, then neither does $\widetilde{\mathcal{C}}$ and we can pick $p_n \in X \setminus \cup_{k=1}^{n} U_k$. Then since X is sequentially compact, there is a subsequence $\{p_{n_k}\}$ such that $\{p_{n_k}\}$ converges. Say

$$p = \lim_{k \to \infty} p_{n_k}.$$

All but finitely many points of $\{p_{n_k}\}$ are in $X \setminus \cup_{k=1}^{n} U_k$. Therefore $p \in X \setminus \cup_{k=1}^{n} U_k$ for each n. Hence

$$p \notin \cup_{k=1}^{\infty} U_k$$

contradicting the construction of $\{U_n\}_{n=1}^{\infty}$. Hence X is compact. This proves the proposition.

Next we apply this very general result to a familiar example, $\mathbb{R}^n$. In this setting totally bounded and bounded are the same. This will yield another proof of the Heine Borel theorem.

Lemma 3.36 *A subset of $\mathbb{R}^n$ is totally bounded if and only if it is bounded.*

Proof: Let A be totally bounded. We need to show it is bounded. Let $\mathbf{x}_1, \cdots, \mathbf{x}_p$ be a 1 net for A. Now consider the ball $B(\mathbf{0}, r+1)$ where $r > \max(||\mathbf{x}_i|| : i = 1, \cdots, p)$. If $\mathbf{z} \in A$, then $\mathbf{z} \in B(\mathbf{x}_j, 1)$ for some j and so by the triangle inequality,

$$||\mathbf{z} - \mathbf{0}|| \leq ||\mathbf{z} - \mathbf{x}_j|| + ||\mathbf{x}_j|| < 1 + r.$$

Thus $A \subseteq B(\mathbf{0}, r+1)$ and so A is bounded.

Now suppose A is bounded and suppose A is not totally bounded. Then there exists $\epsilon > 0$ such that there is no ϵ net for A. Therefore, there exists a sequence of points $\{a_i\}$ with $||a_i - a_j|| \geq \epsilon$ if $i \neq j$. Since A is bounded, there exists $r > 0$ such that

$$A \subseteq [-r, r)^n.$$

($\mathbf{x} \in [-r, r)^n$ means $x_i \in [-r, r)$ for each i.) Now define $\mathcal{S}$ to be all cubes of the form

$$\prod_{k=1}^{n} [a_k, b_k)$$

where

$$a_k = -r + i 2^{-p} r, \; b_k = -r + (i+1) 2^{-p} r,$$

for $i \in \{0, 1, \cdots, 2^{p+1} - 1\}$. Thus $\mathcal{S}$ is a collection of $\left(2^{p+1}\right)^n$ nonoverlapping cubes whose union equals $[-r, r)^n$ and whose diameters are all equal to $2^{-p} r \sqrt{n}$. Now choose p large enough that the diameter of these cubes is less than ϵ. This yields a contradiction because one of the cubes must contain infinitely many points of $\{a_i\}$. This proves the lemma.

The next theorem is called the Heine Borel theorem and it characterizes the compact sets in $\mathbb{R}^n$.

Theorem 3.37 *A subset of* $\mathbb{R}^n$ *is compact if and only if it is closed and bounded.*

Proof: Since a set in $\mathbb{R}^n$ is totally bounded if and only if it is bounded, this theorem follows from Proposition 3.35 and the observation that a subset of $\mathbb{R}^n$ is closed if and only if it is complete. This proves the theorem.

The following corollary is an important existence theorem which depends on compactness.

Corollary 3.38 *Let* (X, τ) *be a compact topological space and let* $f : X \to \mathbb{R}$ *be continuous. Then* $\max \{f(x) : x \in X\}$ *and* $\min \{f(x) : x \in X\}$ *both exist.*

Proof: Since f is continuous, it follows that $f(X)$ is compact. From Theorem 3.37 $f(X)$ is closed and bounded. This implies it has a largest and a smallest value. This proves the corollary.

3.2 Connected sets

Stated informally, connected sets are those which are in one piece. More precisely, we give the following definition.

Definition 3.39 *We say a set,* S *in a general topological space is separated if there exist sets,* A, B *such that*

$$S = A \cup B, \ A, B \neq \emptyset, \text{ and } \overline{A} \cap B = \overline{B} \cap A = \emptyset.$$

In this case, the sets A *and* B *are said to separate* S*. We say a set is connected if it is not separated.*

One of the most important theorems about connected sets is the following.

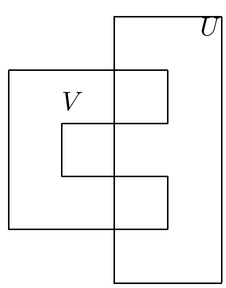
Theorem 3.40 *Suppose* U *and* V *are connected sets having nonempty intersection. Then* $U \cup V$ *is also connected.*

Proof: Suppose $U \cup V = A \cup B$ where $\overline{A} \cap B = \overline{B} \cap A = \emptyset$. Consider the sets, $A \cap U$ and $B \cup U$. Since

$$\overline{(A \cap U)} \cap (B \cap U) = (A \cap U) \cap \left(\overline{B \cap U}\right) = \emptyset,$$

It follows one of these sets must be empty since otherwise, U would be separated. It follows that U is contained in either A or B. Similarly, V must be contained in either A or B. Since U and V have nonempty intersection, it follows that both V and U are contained in one of the sets, A, B. Therefore, the other must be empty and this shows $U \cup V$ cannot be separated and is therefore, connected.

The intersection of connected sets is not necessarily connected as is shown by the following picture.



Theorem 3.41 *Let* $f : X \to Y$ *be continuous where* X *and* Y *are topological spaces and* X *is connected. Then* $f(X)$ *is also connected.*

Proof: We show $f(X)$ is not separated. Suppose to the contrary that $f(X) = A \cup B$ where A and B separate $f(X)$. Then consider the sets, $f^{-1}(A)$ and $f^{-1}(B)$. If $z \in f^{-1}(B)$, then $f(z) \in B$ and so $f(z)$ is not a limit point of A. Therefore, there exists an open set, U containing $f(z)$ such that $U \cap A = \emptyset$. But then, the continuity of f implies that $f^{-1}(U)$ is an open set containing z such that $f^{-1}(U) \cap f^{-1}(A) = \emptyset$. Therefore, $f^{-1}(B)$ contains no limit points of $f^{-1}(A)$. Similar reasoning implies $f^{-1}(A)$ contains no limit points of $f^{-1}(B)$. It follows that X is separated by $f^{-1}(A)$ and $f^{-1}(B)$, contradicting the assumption that X was connected.

An arbitrary set can be written as a union of maximal connected sets called connected components. This is the concept of the next definition.

Definition 3.42 *Let S be a set and let $p \in S$. Denote by C_p the union of all connected subsets of S which contain p. This is called the connected component determined by p.*

Theorem 3.43 *Let C_p be a connected component of a set S in a general topological space. Then C_p is a connected set and if $C_p \cap C_q \neq \emptyset$, then $C_p = C_q$.*

Proof: Let $\mathcal{C}$ denote the connected subsets of S which contain p. If $C_p = A \cup B$ where

$$\overline{A} \cap B = \overline{B} \cap A = \emptyset,$$

then p is in one of A or B. Suppose without loss of generality $p \in A$. Then every set of $\mathcal{C}$ must also be contained in A also since otherwise, as in Theorem 3.40, the set would be separated. But this implies B is empty. Therefore, C_p is connected. From this, and Theorem 3.40, the second assertion of the theorem is proved.

This shows the connected components of a set are equivalence classes and partition the set.

A set, I is an interval in $\mathbb{R}$ if and only if whenever $x, y \in I$ then $(x, y) \subseteq I$. The following theorem is about the connected sets in $\mathbb{R}$.

Theorem 3.44 *A set, C in $\mathbb{R}$ is connected if and only if C is an interval.*

Proof: Let C be connected. If C consists of a single point, p, there is nothing to prove. The interval is just $[p, p]$. Suppose $p < q$ and $p, q \in C$. We need to show $(p, q) \subseteq C$. If

$$x \in (p, q) \setminus C$$

let $C \cap (-\infty, x) \equiv A$, and $C \cap (x, \infty) \equiv B$. Then $C = A \cup B$ and the sets, A and B separate C contrary to the assumption that C is connected.

Conversely, let I be an interval. Suppose I is separated by A and B. Pick $x \in A$ and $y \in B$. Suppose without loss of generality that $x < y$. Now define the set,

$$S \equiv \{t \in [x, y] : [x, t] \subseteq A\}$$

and let l be the least upper bound of S. Then $l \in \overline{A}$ so $l \notin B$ which implies $l \in A$. But if $l \notin \overline{B}$, then for some $\delta > 0$,

$$(l, l + \delta) \cap B = \emptyset$$

contradicting the definition of l as an upper bound for S. Therefore, $l \in \overline{B}$ which implies $l \notin A$ after all, a contradiction. It follows I must be connected.

The following theorem is a very useful description of the open sets in $\mathbb{R}$.

Theorem 3.45 *Let U be an open set in $\mathbb{R}$. Then there exist countably many disjoint open sets, $\{(a_i, b_i)\}_{i=1}^{\infty}$ such that $U = \cup_{i=1}^{\infty} (a_i, b_i)$.*

Proof: Let $p \in U$ and let $z \in C_p$, the connected component determined by p. Since U is open, there exists, $\delta > 0$ such that $(z - \delta, z + \delta) \subseteq U$. It follows from Theorem 3.40 that

$$(z - \delta, z + \delta) \subseteq C_p.$$

This shows C_p is open. By Theorem 3.44, this shows C_p is an open interval, (a, b) where $a, b \in [-\infty, \infty]$. There are therefore at most countably many of these connected components because each must contain a rational number and the rational numbers are countable. Denote by $\{(a_i, b_i)\}_{i=1}^{\infty}$ the set of these connected components. This proves the theorem.

Definition 3.46 *We say a topological space, E is arcwise connected if for any two points, $p, q \in E$, there exists a closed interval, $[a, b]$ and a continuous function, $\gamma : [a, b] \to E$ such that $\gamma(a) = p$ and $\gamma(b) = q$. We say E is locally connected if it has a basis of connected open sets. We say E is locally arcwise connected if it has a basis of arcwise connected open sets.*

An example of an arcwise connected topological space would be the any subset of $\mathbb{R}^n$ which is the continuous image of an interval. Locally connected is not the same as connected. A well known example is the following.

$$\left\{\left(x, \sin\frac{1}{x}\right) : x \in (0, 1]\right\} \cup \{(0, y) : y \in [-1, 1]\} \tag{3.8}$$

We leave it as an exercise to verify that this set of points considered as a metric space with the metric from $\mathbb{R}^2$ is not locally connected or arcwise connected but is connected.

Proposition 3.47 *If a topological space is arcwise connected, then it is connected.*

Proof: Let X be an arcwise connected space and suppose it is separated. Then $X = A \cup B$ where A, B are two separated sets. Pick $p \in A$ and $q \in B$. Since X is given to be arcwise connected, there must exist a continuous function $\gamma : [a, b] \to X$ such that $\gamma(a) = p$ and $\gamma(b) = q$. But then we would have $\gamma([a, b]) = (\gamma([a, b]) \cap A) \cup (\gamma([a, b]) \cap B)$ and the two sets, $\gamma([a, b]) \cap A$ and $\gamma([a, b]) \cap B$ are separated thus showing that $\gamma([a, b])$ is separated and contradicting Theorem 3.44 and Theorem 3.41. It follows that X must be connected as claimed.

Theorem 3.48 *Let U be an open subset of a locally arcwise connected topological space, X. Then U is arcwise connected if and only if U if connected. Also the connected components of an open set in such a space are open sets, hence arcwise connected.*

Proof: By Proposition 3.47 we only need to verify that if U is connected and open in the context of this theorem, then U is arcwise connected. Pick $p \in U$. We will say $x \in U$ satisfies $\mathcal{P}$ if there exists a continuous function, $\gamma : [a, b] \to U$ such that $\gamma(a) = p$ and $\gamma(b) = x$.

$$A \equiv \{x \in U \text{ such that } x \text{ satisfies } \mathcal{P}.\}$$

If $x \in A$, there exists, according to the assumption that X is locally arcwise connected, an open set, V, containing x and contained in U which is arcwise connected. Thus letting $y \in V$, there exist intervals, $[a, b]$ and $[c, d]$ and continuous functions having values in U, γ, η such that $\gamma(a) = p, \gamma(b) = x, \eta(c) = x$, and $\eta(d) = y$. Then let $\gamma_1 : [a, b + d - c] \to U$ be defined as

$$\gamma_1(t) \equiv \begin{cases} \gamma(t) & \text{if } t \in [a, b] \\ \eta(t) & \text{if } t \in [b, b + d - c] \end{cases}$$

Then it is clear that γ_1 is a continuous function mapping p to y and showing that $V \subseteq A$. Therefore, A is open. We also know that $A \neq \emptyset$ because there is an open set, V containing p which is contained in U and is arcwise connected.

Now consider $B \equiv U \setminus A$. We will verify that this is also open. If B is not open, there exists a point $z \in B$ such that every open set conaining z is not contained in B. Therefore, letting V be one of the basic open sets chosen such that $z \in V \subseteq U$, we must have points of A contained in V. But then, a repeat of the above argument shows $z \in A$ also. Hence B is open and so if $B \neq \emptyset$, then $U = B \cup A$ and so U is separated by the two sets, B and A contradicting the assumption that U is connected.

We need to verify the connected components are open. Let $z \in C_p$ where C_p is the connected component determined by p. Then picking V an arcwise connected open set which contains z and is contained in U, $C_p \cup V$ is connected and contained in U and so it must also be contained in C_p. This proves the theorem.

3.3 The Tychonoff theorem

This section might be omitted on a first reading of the book. It is on the very important theorem about products of compact topological spaces. In order to prove this theorem we need to use a fundamental result about partially ordered sets which we describe next.

Definition 3.49 *Let $\mathcal{F}$ be a nonempty set. $\mathcal{F}$ is called a partially ordered set if there is a relation, denoted here by $\leq$, such that*

$$x \leq x \text{ for all } x \in \mathcal{F}.$$

$$\text{If } x \leq y \text{ and } y \leq z \text{ then } x \leq z.$$

$\mathcal{C} \subseteq \mathcal{F}$ is said to be a chain if every two elements of $\mathcal{C}$ are related. By this we mean that if $x, y \in \mathcal{C}$, then either $x \leq y$ or $y \leq x$. Sometimes we call a chain a totally ordered set. $\mathcal{C}$ is said to be a maximal chain if whenever $\mathcal{D}$ is a chain containing $\mathcal{C}$, $\mathcal{D} = \mathcal{C}$.

The most common example of a partially ordered set is the power set of a given set with $\subseteq$ being the relation. The following theorem is equivalent to the axiom of choice. For a discussion of this, see the appendix on the subject.

Theorem 3.50 *(Hausdorff Maximal Principle) Let $\mathcal{F}$ be a nonempty partially ordered set. Then there exists a maximal chain.*

The main tool in the study of products of compact topological spaces is the Alexander subbasis theorem which we present next.

Theorem 3.51 *Let (X, τ) be a topological space and let $\mathcal{S} \subseteq \tau$ be a subbasis for τ. (Recall this means that finite intersections of sets of $\mathcal{S}$ form a basis for τ.) Then if $H \subseteq X$, H is compact if and only if every open cover of H consisting entirely of sets of $\mathcal{S}$ admits a finite subcover.*

Proof: The only if part is obvious. To prove the other implication, first note that if every basic open cover, an open cover composed of basic open sets, admits a finite subcover, then H is compact. Now suppose that every subbasic open cover of H admits a finite subcover but H is not compact. This implies that there exists a basic open cover of H, $\mathcal{O}$, which admits no finite subcover. Let $\mathcal{F}$ be defined as

$$\{\mathcal{O} : \mathcal{O} \text{ is a basic open cover of } H \text{ which admits no finite subcover}\}.$$

Partially order $\mathcal{F}$ by set inclusion and use the Hausdorff maximal principle to obtain a maximal chain, $\mathcal{C}$, of such open covers. Let

$$\mathcal{D} = \cup \mathcal{C}.$$

Then it follows that $\mathcal{D}$ is an open cover of H which is maximal with respect to the property of being a basic open cover having no finite subcover of H. (If $\mathcal{D}$ admits a finite subcover, then since $\mathcal{C}$ is a chain and the finite subcover has only finitely many sets, some element of $\mathcal{C}$ would also admit a finite subcover, contrary to the definition of $\mathcal{F}$.) Thus if $\mathcal{D}' \supsetneqq \mathcal{D}$ and $\mathcal{D}'$ is a basic open cover of H, then $\mathcal{D}'$ has a finite subcover of H. One of the sets of $\mathcal{D}$, U, has the property that

$$U = \cap_{i=1}^{m} B_i,\ B_i \in \mathcal{S}$$

and no B_i is in $\mathcal{D}$. If not, we could replace each set in $\mathcal{D}$ with a subbasic set also in $\mathcal{D}$ containing it and thereby obtain a subbasic cover which would, by assumption, admit a finite subcover, contrary to the properties of $\mathcal{D}$. Thus $\mathcal{D} \cup \{\mathcal{B}_i\}$ admits a finite subcover,

$$V_1^i, \cdots, V_{m_i}^i, B_i$$

for each $i = 1, \cdots, m$. Consider

$$\{U, V_j^i,\ j = 1, \cdots, m_i,\ i = 1, \cdots, m\}.$$

If $p \in H \setminus \cup\{V_j^i\}$, then $p \in B_i$ for each i and so $p \in U$. This is therefore a finite subcover of $\mathcal{D}$ contradicting the properties of $\mathcal{D}$. This proves the theorem.

Let I be a set and suppose for each $i \in I$, (X_i, τ_i) is a nonempty topological space. The Cartesian product of the X_i, denoted by

$$\prod_{i \in I} X_i,$$

consists of the set of all choice functions defined on I which select a single element of each X_i. Thus

$$f \in \prod_{i \in I} X_i$$

means for every $i \in I$, $f(i) \in X_i$. The axiom of choice says $\prod_{i \in I} X_i$ is nonempty. Let

$$P_j(A) = \prod_{i \in I} B_i$$

where $B_i = X_i$ if $i \neq j$ and $B_j = A$. A subbasis for a topology on the product space consists of all sets $P_j(A)$ where $A \in \tau_j$. (These sets have an open set in the j^{th} slot and the whole space in the other slots.) Thus a basis consists of finite intersections of these sets. It is easy to see that finite intersections do form a basis for a topology. This topology is called the product topology and we will denote it by $\prod \tau_i$. Next we use the Alexander subbasis theorem to prove the Tychonoff theorem.

Theorem 3.52 *If $(X_i \tau_i)$ is compact, then so is $(\prod_{i \in I} X_i, \prod \tau_i)$.*

Proof: By the Alexander subbasis theorem, we will establish compactness of the product space if we show every subbasic open cover admits a finite subcover. Therefore, let $\mathcal{O}$ be a subbasic open cover of $\prod_{i \in I} X_i$. Let

$$\mathcal{O}_j = \{Q \in \mathcal{O} : Q = P_j(A) \text{ for some } A \in \tau_j\}.$$

Let

$$\pi_j \mathcal{O}_j = \{A : P_j(A) \in \mathcal{O}_j\}.$$

If no $\pi_j \mathcal{O}_j$ covers X_j, then pick

$$f \in \prod_{i \in I} X_i \setminus \cup \pi_i \mathcal{O}_i$$

so $f(j) \notin \cup \pi_j \mathcal{O}_j$ and so $f \notin \cup \mathcal{O}$ contradicting $\mathcal{O}$ is an open cover. Hence, for some j,

$$X_j = \cup \pi_j \mathcal{O}_j$$

and so there exist $A_1, \cdots, A_m$, sets in τ_j such that

$$X_j \subseteq \cup_{i=1}^m A_i$$

and $P_j(A_i) \in \mathcal{O}$. Therefore, $\{P_j(A_i)\}_{i=1}^m$ covers $\prod_{i \in I} X_i$.

3.4 Exercises

1. Prove the definition of distance in $\mathbb{R}^n$ or $\mathbb{C}^n$ satisfies (3.3)-(3.4). In addition to this, prove that $||\cdot||$ given by $||\mathbf{x}|| = (\sum_{i=1}^n |x_i|^2)^{1/2}$ is a norm. This means it satisfies the following.

$$||\mathbf{x}|| \geq 0,\ ||\mathbf{x}|| = 0 \text{ if and only if } \mathbf{x} = \mathbf{0}.$$

$$||\alpha \mathbf{x}|| = |\alpha| ||\mathbf{x}|| \text{ for } \alpha \text{ a number.}$$

$$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||.$$

2. Completeness of $\mathbb{R}$ is an axiom. Using this, show $\mathbb{R}^n$ and $\mathbb{C}^n$ are complete metric spaces with respect to the distance given by the usual norm.

3. Prove Urysohn's lemma. A Hausdorff space, X, is normal if and only if whenever K and H are disjoint nonempty closed sets, there exists a continuous function $f : X \to [0, 1]$ such that $f(k) = 0$ for all $k \in K$ and $f(h) = 1$ for all $h \in H$.

4. Prove that $f : X \to Y$ is continuous if and only if f is continuous at every point of X.

5. Suppose (X, d), and (Y, ρ) are metric spaces and let $f : X \to Y$. Show f is continuous at $x \in X$ if and only if whenever $x_n \to x$, $f(x_n) \to f(x)$. (Recall that $x_n \to x$ means that for all $\epsilon > 0$, there exists n_ϵ such that $d(x_n, x) < \epsilon$ whenever $n > n_\epsilon$.)

6. If (X, d) is a metric space, give an easy proof independent of Problem 3 that whenever K, H are disjoint non empty closed sets, there exists $f : X \to [0, 1]$ such that f is continuous, $f(K) = \{0\}$, and $f(H) = \{1\}$.

7. Let (X, τ) (Y, η)be topological spaces with (X, τ) compact and let $f : X \to Y$ be continuous. Show $f(X)$ is compact.

8. (An example) Let $X = [-\infty, \infty]$ and consider $\mathcal{B}$ defined by sets of the form (a, b), $[-\infty, b)$, and $(a, \infty]$. Show $\mathcal{B}$ is the basis for a topology on X.

9. $\uparrow$ Show (X, τ) defined in Problem 8 is a compact Hausdorff space.

10. $\uparrow$ Show (X, τ) defined in Problem 8 is completely separable.

11. $\uparrow$ In Problem 8, show sets of the form $[-\infty, b)$ and $(a, \infty]$ form a subbasis for the topology described in Problem 8.

12. Let (X,τ) and (Y,η) be topological spaces and let $f : X \rightarrow Y$. Also let $\mathcal{S}$ be a subbasis for η. Show f is continuous if and only if $f^{-1}(V) \in \tau$ for all $V \in \mathcal{S}$. Thus, it suffices to check inverse images of subbasic sets in checking for continuity.

13. Show the usual topology of $\mathbb{R}^n$ is the same as the product topology of
$$\prod_{i=1}^{n} \mathbb{R} \equiv \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}.$$
Do the same for $\mathbb{C}^n$.

14. If M is a separable metric space and $T \subseteq M$, then T is separable also.

15. Prove the Heine Borel theorem as follows. First show $[a,b]$ is compact in $\mathbb{R}$. Next use Theorem 3.30 to show that $\prod_{i=1}^{n} [a_i, b_i]$ is compact. Use this to verify that compact sets are exactly those which are closed and bounded.

16. Show the rational numbers, $\mathbb{Q}$, are countable.

17. Verify that the set of (3.8) is connected but not locally connected or arcwise connected.

18. Let α be an n dimensional multi-index. This means
$$\alpha = (\alpha_1, \cdot \cdot \cdot, \alpha_n)$$
where each α_i is a natural number or zero. Also, we let
$$|\alpha| \equiv \sum_{i=1}^{n} |\alpha_i|$$
When we write $\mathbf{x}^{\alpha}$, we mean
$$\mathbf{x}^{\alpha} \equiv x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_3^{\alpha_n}.$$
An n dimensional polynomial of degree m is a function of the form
$$\sum_{|\alpha| \leq m} d_{\alpha} \mathbf{x}^{\alpha}.$$
Let $\mathcal{R}$ be all n dimensional polynomials whose coefficients d_{α} come from the rational numbers, $\mathbb{Q}$. Show $\mathcal{R}$ is countable.

19. Let (X,d) be a metric space where d is a bounded metric. Let $\mathcal{C}$ denote the collection of closed subsets of X. For $A, B \in \mathcal{C}$, define
$$\rho(A,B) \equiv \inf \{\delta > 0 : A_{\delta} \supseteq B \text{ and } B_{\delta} \supseteq A\}$$
where for a set S,
$$S_{\delta} \equiv \{x : dist(x,S) \equiv \inf \{d(x,s) : s \in S\} \leq \delta\}.$$
Show S_{δ} is a closed set containing S. Also show that ρ is a metric on $\mathcal{C}$. This is called the Hausdorff metric.

20. Using 19, suppose (X, d) is a compact metric space. Show $(\mathcal{C}, \rho)$ is a complete metric space. **Hint:** Show first that if $W_n \downarrow W$ where W_n is closed, then $\rho(W_n, W) \to 0$. Now let $\{A_n\}$ be a Cauchy sequence in $\mathcal{C}$. Then if $\epsilon > 0$ there exists N such that when $m, n \geq N$, then $\rho(A_n, A_m) < \epsilon$. Therefore, for each $n \geq N$,
$$(A_n)_\epsilon \supseteq \overline{\cup_{k=n}^{\infty} A_k}.$$
Let $A \equiv \cap_{n=1}^{\infty} \overline{\cup_{k=n}^{\infty} A_k}$. By the first part, there exists $N_1 > N$ such that for $n \geq N_1$,
$$\rho\left(\overline{\cup_{k=n}^{\infty} A_k}, A\right) < \epsilon, \text{ and } (A_n)_\epsilon \supseteq \overline{\cup_{k=n}^{\infty} A_k}.$$
Therefore, for such n, $A_\epsilon \supseteq W_n \supseteq A_n$ and $(W_n)_\epsilon \supseteq (A_n)_\epsilon \supseteq A$ because
$$(A_n)_\epsilon \supseteq \overline{\cup_{k=n}^{\infty} A_k} \supseteq A.$$

21. In the situation of the last two problems, let X be a compact metric space. Show $(\mathcal{C}, \rho)$ is compact. **Hint:** Let $\mathcal{D}_n$ be a 2^{-n} net for X. Let $\mathcal{K}_n$ denote finite unions of sets of the form $\overline{B(p, 2^{-n})}$ where $p \in \mathcal{D}_n$. Show $\mathcal{K}_n$ is a $2^{-(n-1)}$ net for $(\mathcal{C}, \rho)$.

Spaces of Continuous Functions

This chapter deals with vector spaces whose vectors are continuous functions.

4.1 Compactness in spaces of continuous functions

Let (X, τ) be a compact space and let $C(X; \mathbb{R}^n)$ denote the space of continuous $\mathbb{R}^n$ valued functions. For $f \in C(X; \mathbb{R}^n)$ let

$$||f||_\infty \equiv \sup\{|f(x)| : x \in X\}$$

where the norm in the parenthesis refers to the usual norm in $\mathbb{R}^n$.

The following proposition shows that $C(X; \mathbb{R}^n)$ is an example of a Banach space.

Proposition 4.1 *$(C(X; \mathbb{R}^n), || \, ||_\infty)$ is a Banach space.*

Proof: It is obvious $|| \, ||_\infty$ is a norm because (X, τ) is compact. Also it is clear that $C(X; \mathbb{R}^n)$ is a linear space. Suppose $\{f_r\}$ is a Cauchy sequence in $C(X; \mathbb{R}^n)$. Then for each $x \in X$, $\{f_r(x)\}$ is a Cauchy sequence in $\mathbb{R}^n$. Let

$$f(x) \equiv \lim_{k\to\infty} f_k(x).$$

Therefore,

$$\sup_{x\in X} |f(x) - f_k(x)| = \sup_{x\in X} \lim_{m\to\infty} |f_m(x) - f_k(x)|$$

$$\leq \lim \sup_{m\to\infty} ||f_m - f_k||_\infty < \epsilon$$

for all k large enough. Thus,

$$\lim_{k\to\infty} \sup_{x\in X} |f(x) - f_k(x)| = 0.$$

It only remains to show that f is continuous. Let

$$\sup_{x\in X} |f(x) - f_k(x)| < \epsilon/3$$

whenever $k \geq k_0$ and pick $k \geq k_0$.

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\ &< 2\epsilon/3 + |f_k(x) - f_k(y)| \end{aligned}$$

Now f_k is continuous and so there exists U an open set containing x such that if $y \in U$, then

$$|f_k(x) - f_k(y)| < \epsilon/3.$$

Thus, for all $y \in U$, $|f(x) - f(y)| < \epsilon$ and this shows that f is continuous and proves the proposition.

This space is a normed linear space and so it is a metric space with the distance given by $d(f,g) \equiv ||f-g||_\infty$. The next task is to find the compact subsets of this metric space. We know these are the subsets which are complete and totally bounded by Proposition 3.35, but which sets are those? We need another way to identify them which is more convenient. This is the extremely important Ascoli Arzela theorem which is the next big theorem.

Definition 4.2 *We say $\mathcal{F} \subseteq C(X;\mathbb{R}^n)$ is equicontinuous at x_0 if for all $\epsilon > 0$ there exists $U \in \tau$, $x_0 \in U$, such that if $x \in U$, then for all $f \in \mathcal{F}$,*

$$|f(x) - f(x_0)| < \epsilon.$$

If $\mathcal{F}$ is equicontinuous at every point of X, we say $\mathcal{F}$ is equicontinuous. We say $\mathcal{F}$ is bounded if there exists a constant, M, such that $||f||_\infty < M$ for all $f \in \mathcal{F}$.

Lemma 4.3 *Let $\mathcal{F} \subseteq C(X;\mathbb{R}^n)$ be equicontinuous and bounded and let $\epsilon > 0$ be given. Then if $\{f_r\} \subseteq \mathcal{F}$, there exists a subsequence $\{g_k\}$, depending on ϵ, such that*

$$||g_k - g_m||_\infty < \epsilon$$

whenever k, m are large enough.

Proof: If $x \in X$ there exists an open set U_x containing x such that for all $f \in \mathcal{F}$ and $y \in U_x$,

$$|f(x) - f(y)| < \epsilon/4. \tag{4.1}$$

Since X is compact, finitely many of these sets, $U_{x_1}, \cdot\cdot\cdot, U_{x_p}$, cover X. Let $\{f_{1k}\}$ be a subsequence of $\{f_k\}$ such that $\{f_{1k}(x_1)\}$ converges. Such a subsequence exists because $\mathcal{F}$ is bounded. Let $\{f_{2k}\}$ be a subsequence of $\{f_{1k}\}$ such that $\{f_{2k}(x_i)\}$ converges for $i = 1, 2$. Continue in this way and let $\{g_k\} = \{f_{pk}\}$. Thus $\{g_k(x_i)\}$ converges for each x_i. Therefore, if $\epsilon > 0$ is given, there exists m_ϵ such that for $k, m > m_\epsilon$,

$$\max\{|g_k(x_i) - g_m(x_i)| : i = 1, \cdot\cdot\cdot, p\} < \frac{\epsilon}{2}.$$

Now if $y \in X$, then $y \in U_{x_i}$ for some x_i. Denote this x_i by x_y. Now let $y \in X$ and $k, m > m_\epsilon$. Then by (4.1),

$$|g_k(y) - g_m(y)| \leq |g_k(y) - g_k(x_y)| + |g_k(x_y) - g_m(x_y)| + |g_m(x_y) - g_m(y)|$$

$$< \frac{\epsilon}{4} + \max\{|g_k(x_i) - g_m(x_i)| : i = 1, \cdot\cdot\cdot, p\} + \frac{\epsilon}{4} < \varepsilon.$$

It follows that for such k, m,

$$||g_k - g_m||_\infty < \epsilon$$

and this proves the lemma.

Theorem 4.4 *(Ascoli Arzela) Let $\mathcal{F} \subseteq C(X;\mathbb{R}^n)$. Then $\mathcal{F}$ is compact if and only if $\mathcal{F}$ is closed, bounded, and equicontinuous.*

Proof: Suppose $\mathcal{F}$ is closed, bounded, and equicontinuous. We will show this implies $\mathcal{F}$ is totally bounded. Then since $\mathcal{F}$ is closed, it follows that $\mathcal{F}$ is complete and will therefore be compact by Proposition 3.35. Suppose $\mathcal{F}$ is not totally bounded. Then there exists $\epsilon > 0$ such that there is no ϵ net. Hence there exists a sequence $\{f_k\} \subseteq \mathcal{F}$ such that

$$||f_k - f_l|| \geq \epsilon$$

for all $k \neq l$. This contradicts Lemma 4.3. Thus $\mathcal{F}$ must be totally bounded and this proves half of the theorem.

Now suppose $\mathcal{F}$ is compact. Then it must be closed and totally bounded. This implies $\mathcal{F}$ is bounded. It remains to show $\mathcal{F}$ is equicontinuous. Suppose not. Then there exists $x \in X$ such that $\mathcal{F}$ is not equicontinuous at x. Thus there exists $\epsilon > 0$ such that for every open U containing x, there exists $f \in \mathcal{F}$ such that $|f(x) - f(y)| \geq \epsilon$ for some $y \in U$.

Let $\{h_1, \cdots, h_p\}$ be an $\epsilon/4$ net for $\mathcal{F}$. For each z, let U_z be an open set containing z such that for all $y \in U_z$,

$$|h_i(z) - h_i(y)| < \epsilon/8$$

for all $i = 1, \cdots, p$. Let $U_{x_1}, \cdots, U_{x_m}$ cover X. Then $x \in U_{x_i}$ for some x_i and so, for some $y \in U_{x_i}$,there exists $f \in \mathcal{F}$ such that $|f(x) - f(y)| \geq \epsilon$. Since $\{h_1, \cdots, h_p\}$ is an $\epsilon/4$ net, it follows that for some j, $||f - h_j||_\infty < \frac{\epsilon}{4}$ and so

$$\epsilon \leq |f(x) - f(y)| \leq |f(x) - h_j(x)| + |h_j(x) - h_j(y)| +$$

$$|h_i(y) - f(y)| \leq \epsilon/2 + |h_j(x) - h_j(y)| \leq \epsilon/2 +$$

$$|h_j(x) - h_j(x_i)| + |h_j(x_i) - h_j(y)| \leq 3\epsilon/4,$$

a contradiction. This proves the theorem.

4.2 Stone Weierstrass theorem

In this section we give a proof of the important approximation theorem of Weierstrass and its generalization by Stone. This theorem is about approximating an arbitrary continuous function uniformly by a polynomial or some other such function.

Definition 4.5 *We say $\mathcal{A}$ is an algebra of functions if $\mathcal{A}$ is a vector space and if whenever $f, g \in \mathcal{A}$ then $fg \in \mathcal{A}$.*

We will assume that the field of scalars is $\mathbb{R}$ in this section unless otherwise indicated. The approach to the Stone Weierstrass depends on the following estimate which may look familiar to someone who has taken a probability class. The left side of the following estimate is the variance of a binomial distribution. However, it is not necessary to know anything about probability to follow the proof below although what is being done is an application of the moment generating function technique to find the variance.

Lemma 4.6 *The following estimate holds for $x \in [0, 1]$.*

$$\sum_{k=0}^{n} \binom{n}{k} (k - nx)^2 x^k (1 - x)^{n-k} \leq 2n$$

Proof: By the Binomial theorem,

$$\sum_{k=0}^{n}\binom{n}{k}\left(e^{t}x\right)^{k}(1-x)^{n-k}=\left(1-x+e^{t}x\right)^{n}.$$

Differentiating both sides with respect to t and then evaluating at $t=0$ yields

$$\sum_{k=0}^{n}\binom{n}{k}kx^{k}(1-x)^{n-k}=nx.$$

Now doing two derivatives with respect to t yields

$$\sum_{k=0}^{n}\binom{n}{k}k^{2}\left(e^{t}x\right)^{k}(1-x)^{n-k}=n(n-1)\left(1-x+e^{t}x\right)^{n-2}e^{2t}x^{2}$$

$$+n\left(1-x+e^{t}x\right)^{n-1}xe^{t}.$$

Evaluating this at $t=0$,

$$\sum_{k=0}^{n}\binom{n}{k}k^{2}(x)^{k}(1-x)^{n-k}=n(n-1)x^{2}+nx.$$

Therefore,

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}(k-nx)^{2}x^{k}(1-x)^{n-k} &= n(n-1)x^{2}+nx-2n^{2}x^{2}+n^{2}x^{2}\\ &= n\left(x-x^{2}\right)\leq 2n.\end{aligned}$$

This proves the lemma.

Definition 4.7 *Let $f\in C([0,1])$. Then the following polynomials are known as the Bernstein polynomials.*

$$p_{n}(x)\equiv\sum_{k=0}^{n}\binom{n}{k}f\left(\frac{k}{n}\right)x^{k}(1-x)^{n-k}.$$

Theorem 4.8 *Let $f\in C([0,1])$ and let p_n be given in Definition 4.7. Then*

$$\lim_{n\to\infty}||f-p_{n}||_{\infty}=0.$$

Proof: Since f is continuous on the compact $[0,1]$, it follows f is uniformly continuous there and so if $\epsilon>0$ is given, there exists $\delta>0$ such that if

$$|y-x|\leq\delta,$$

then

$$|f(x)-f(y)|<\epsilon/2.$$

By the Binomial theorem,

$$f(x)=\sum_{k=0}^{n}\binom{n}{k}f(x)x^{k}(1-x)^{n-k}$$

and so

$$|p_n(x) - f(x)| \leq \sum_{k=0}^{n} \binom{n}{k} \left| f\left(\frac{k}{n}\right) - f(x) \right| x^k (1-x)^{n-k}$$

$$\leq \sum_{|k/n-x|>\delta} \binom{n}{k} \left| f\left(\frac{k}{n}\right) - f(x) \right| x^k (1-x)^{n-k} +$$

$$\sum_{|k/n-x|\leq\delta} \binom{n}{k} \left| f\left(\frac{k}{n}\right) - f(x) \right| x^k (1-x)^{n-k}$$

$$< \quad \epsilon/2 + 2\,||f||_\infty \sum_{(k-nx)^2>n^2\delta^2} \binom{n}{k} x^k (1-x)^{n-k}$$

$$\leq \quad \frac{2\,||f||_\infty}{n^2\delta^2} \sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^k (1-x)^{n-k} + \epsilon/2.$$

By the lemma,

$$\leq \frac{4\,||f||_\infty}{\delta^2 n} + \epsilon/2 < \epsilon$$

whenever n is large enough. This proves the theorem.

The next corollary is called the Weierstrass approximation theorem.

Corollary 4.9 *The polynomials are dense in $C([a,b])$.*

Proof: Let $f \in C([a,b])$ and let $h : [0,1] \to [a,b]$ be linear and onto. Then $f \circ h$ is a continuous function defined on $[0,1]$ and so there exists a polynomial, p_n such that

$$|f(h(t)) - p_n(t)| < \epsilon$$

for all $t \in [0,1]$. Therefore for all $x \in [a,b]$,

$$\left| f(x) - p_n\left(h^{-1}(x)\right) \right| < \epsilon.$$

Since h is linear $p_n \circ h^{-1}$ is a polynomial. This proves the theorem.

The next result is the key to the profound generalization of the Weierstrass theorem due to Stone in which an interval will be replaced by a compact or locally compact set and polynomials will be replaced with elements of an algebra satisfying certain axioms.

Corollary 4.10 *On the interval $[-M,M]$, there exist polynomials p_n such that*

$$p_n(0) = 0$$

and

$$\lim_{n\to\infty} ||p_n - |\cdot|||_\infty = 0.$$

Proof: Let $\tilde{p}_n \to |\cdot|$ uniformly and let

$$p_n \equiv \tilde{p}_n - \tilde{p}_n(0).$$

This proves the corollary.

The following generalization is known as the Stone Weierstrass approximation theorem. First, we say an algebra of functions, $\mathcal{A}$ defined on A, annihilates no point of A if for all $x \in A$, there exists $g \in \mathcal{A}$ such that $g(x) \neq 0$. We say the algebra separates points if whenever $x_1 \neq x_2$, then there exists $g \in \mathcal{A}$ such that $g(x_1) \neq g(x_2)$.

Theorem 4.11 *Let A be a compact topological space and let $\mathcal{A} \subseteq C(A;\mathbb{R})$ be an algebra of functions which separates points and annihilates no point. Then $\mathcal{A}$ is dense in $C(A;\mathbb{R})$.*

Proof: We begin by proving a simple lemma.

Lemma 4.12 *Let c_1 and c_2 be two real numbers and let $x_1 \neq x_2$ be two points of A. Then there exists a function $f_{x_1x_2}$ such that*

$$f_{x_1x_2}(x_1) = c_1,\ f_{x_1x_2}(x_2) = c_2.$$

Proof of the lemma: Let $g \in \mathcal{A}$ satisfy

$$g(x_1) \neq g(x_2).$$

Such a g exists because the algebra separates points. Since the algebra annihilates no point, there exist functions h and k such that

$$h(x_1) \neq 0,\ k(x_2) \neq 0.$$

Then let

$$u \equiv gh - g(x_2)h,\ v \equiv gk - g(x_1)k.$$

It follows that $u(x_1) \neq 0$ and $u(x_2) = 0$ while $v(x_2) \neq 0$ and $v(x_1) = 0$. Let

$$f_{x_1x_2} \equiv \frac{c_1 u}{u(x_1)} + \frac{c_2 v}{v(x_2)}.$$

This proves the lemma. Now we continue with the proof of the theorem.

First note that $\overline{\mathcal{A}}$ satisfies the same axioms as $\mathcal{A}$ but in addition to these axioms, $\overline{\mathcal{A}}$ is closed. Suppose $f \in \overline{\mathcal{A}}$ and suppose M is large enough that

$$||f||_\infty < M.$$

Using Corollary 4.10, let p_n be a sequence of polynomials such that

$$||p_n - |\cdot|||_\infty \to 0,\ p_n(0) = 0.$$

It follows that $p_n \circ f \in \overline{\mathcal{A}}$ and so $|f| \in \overline{\mathcal{A}}$ whenever $f \in \overline{\mathcal{A}}$. Also note that

$$\max(f,g) = \frac{|f-g| + (f+g)}{2}$$

$$\min(f,g) = \frac{(f+g) - |f-g|}{2}.$$

Therefore, this shows that if $f, g \in \overline{\mathcal{A}}$ then

$$\max(f, g), \ \min(f, g) \in \overline{\mathcal{A}}.$$

By induction, if $f_i, i = 1, 2, \cdots, m$ are in $\overline{\mathcal{A}}$ then

$$\max(f_i, i = 1, 2, \cdots, m), \ \min(f_i, i = 1, 2, \cdots, m) \in \overline{\mathcal{A}}.$$

Now let $h \in C(A; \mathbb{R})$ and use Lemma 4.12 to obtain f_{xy}, a function of $\overline{\mathcal{A}}$ which agrees with h at x and y. Let $\epsilon > 0$ and let $x \in A$. Then there exists an open set $U(y)$ containing y such that

$$f_{xy}(z) > h(z) - \epsilon \text{ if } z \in U(y).$$

Since A is compact, let $U(y_1), \cdots, U(y_l)$ cover A. Let

$$f_x \equiv \max(f_{xy_1}, f_{xy_2}, \cdots, f_{xy_l}).$$

Then $f_x \in \overline{\mathcal{A}}$ and

$$f_x(z) > h(z) - \epsilon$$

for all $z \in A$ and $f_x(x) = h(x)$. Then for each $x \in A$ there exists an open set $V(x)$ containing x such that for $z \in V(x)$,

$$f_x(z) < h(z) + \epsilon.$$

Let $V(x_1), \cdots, V(x_m)$ cover A and let

$$f \equiv \min(f_{x_1}, \cdots, f_{x_m}).$$

Therefore,

$$f(z) < h(z) + \epsilon$$

for all $z \in A$ and since each

$$f_x(z) > h(z) - \epsilon,$$

it follows

$$f(z) > h(z) - \epsilon$$

also and so

$$|f(z) - h(z)| < \epsilon$$

for all z. Since ϵ is arbitrary, this shows $h \in \overline{\mathcal{A}}$ and proves $\overline{\mathcal{A}} = C(A; \mathbb{R})$. This proves the theorem.

4.3 Exercises

1. Let $(X, \tau), (Y, \eta)$ be topological spaces and let $A \subseteq X$ be compact. Then if $f : X \to Y$ is continuous, show that $f(A)$ is also compact.

2. $\uparrow$ In the context of Problem 1, suppose $\mathbb{R} = Y$ where the usual topology is placed on $\mathbb{R}$. Show f achieves its maximum and minimum on A.

3. Let V be an open set in $\mathbb{R}^n$. Show there is an increasing sequence of compact sets, K_m, such that $V = \cup_{m=1}^{\infty} K_m$. **Hint:** Let

$$C_m \equiv \left\{ \mathbf{x} \in \mathbb{R}^n : dist\left(\mathbf{x}, V^C\right) \geq \frac{1}{m} \right\}$$

where

$$dist\,(\mathbf{x}, S) \equiv \inf\{|\mathbf{y} - \mathbf{x}| \text{ such that } \mathbf{y} \in S\}.$$

Consider $K_m \equiv C_m \cap \overline{B\,(\mathbf{0}, m)}$.

4. Let $B\,(X; \mathbb{R}^n)$ be the space of functions $\mathbf{f}$, mapping X to $\mathbb{R}^n$ such that

$$\sup\{|\mathbf{f}\,(\mathbf{x})| : \mathbf{x} \in X\} < \infty.$$

Show $B\,(X; \mathbb{R}^n)$ is a complete normed linear space if

$$||\mathbf{f}|| \equiv \sup\{|\mathbf{f}\,(\mathbf{x})| : \mathbf{x} \in X\}.$$

5. Let $\alpha \in [0, 1]$. We define, for X a compact subset of $\mathbb{R}^p$,

$$C^{\alpha}\,(X; \mathbb{R}^n) \equiv \{\mathbf{f} \in C\,(X; \mathbb{R}^n) : \rho_{\alpha}\,(\mathbf{f}) + ||\mathbf{f}|| \equiv ||\mathbf{f}||_{\alpha} < \infty\}$$

where

$$||\mathbf{f}|| \equiv \sup\{|\mathbf{f}\,(\mathbf{x})| : \mathbf{x} \in X\}$$

and

$$\rho_{\alpha}\,(\mathbf{f}) \equiv \sup\{\frac{|\mathbf{f}\,(\mathbf{x}) - \mathbf{f}\,(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^{\alpha}} : \mathbf{x}, \mathbf{y} \in X,\ \mathbf{x} \neq \mathbf{y}\}.$$

Show that $(C^{\alpha}\,(X; \mathbb{R}^n), ||\cdot||_{\alpha})$ is a complete normed linear space.

6. Let $\{\mathbf{f}_n\}_{n=1}^{\infty} \subseteq C^{\alpha}\,(X; \mathbb{R}^n)$ where X is a compact subset of $\mathbb{R}^p$ and suppose

$$||\mathbf{f}_n||_{\alpha} \leq M$$

for all n. Show there exists a subsequence, n_k, such that $\mathbf{f}_{n_k}$ converges in $C\,(X; \mathbb{R}^n)$. We say the given sequence is precompact when this happens. (This also shows the embedding of $C^{\alpha}\,(X; \mathbb{R}^n)$ into $C\,(X; \mathbb{R}^n)$ is a compact embedding.)

7. Let $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be continuous and bounded and let $\mathbf{x}_0 \in \mathbb{R}^n$. If

$$\mathbf{x} : [0, T] \to \mathbb{R}^n$$

and $h > 0$, let

$$\tau_h \mathbf{x}\,(s) \equiv \begin{cases} \mathbf{x}_0 \text{ if } s \leq h, \\ \mathbf{x}\,(s - h)\,, \text{ if } s > h. \end{cases}$$

For $t \in [0, T]$, let

$$\mathbf{x}_h\,(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}\,(s, \tau_h \mathbf{x}_h\,(s))\, ds.$$

Show using the Ascoli Arzela theorem that there exists a sequence $h \to 0$ such that

$$\mathbf{x}_h \to \mathbf{x}$$

in $C([0,T];\mathbb{R}^n)$. Next argue

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t \mathbf{f}(s, \mathbf{x}(s))\, ds$$

and conclude the following theorem. If $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and bounded, and if $\mathbf{x}_0 \in \mathbb{R}^n$ is given, there exists a solution to the following initial value problem.

$$\begin{aligned} \mathbf{x}' &= \mathbf{f}(t, \mathbf{x}), \quad t \in [0,T] \\ \mathbf{x}(0) &= \mathbf{x}_0. \end{aligned}$$

This is the Peano existence theorem for ordinary differential equations.

8. Show the set of polynomials $\mathcal{R}$ described in Problem 18 of Chapter 3 is dense in the space $C(A;\mathbb{R})$ when A is a compact subset of $\mathbb{R}^n$. Conclude from this other problem that $C(A;\mathbb{R})$ is separable.

9. Let H and K be disjoint closed sets in a metric space, (X,d), and let

$$g(x) \equiv \frac{2}{3}h(x) - \frac{1}{3}$$

where

$$h(x) \equiv \frac{dist(x,H)}{dist(x,H) + dist(x,K)}.$$

Show $g(x) \in \left[-\frac{1}{3}, \frac{1}{3}\right]$ for all $x \in X$, g is continuous, and g equals $\frac{-1}{3}$ on H while g equals $\frac{1}{3}$ on K. Is it necessary to be in a metric space to do this?

10. $\uparrow$ Suppose M is a closed set in X where X is the metric space of problem 9 and suppose $f : M \to [-1,1]$ is continuous. Show there exists $g : X \to [-1,1]$ such that g is continuous and $g = f$ on M. **Hint:** Show there exists

$$g_1 \in C(X),\ g_1(x) \in \left[\frac{-1}{3}, \frac{1}{3}\right],$$

and $|f(x) - g_1(x)| \leq \frac{2}{3}$ for all $x \in H$. To do this, consider the disjoint closed sets

$$H \equiv f^{-1}\left(\left[-1, \frac{-1}{3}\right]\right),\ K \equiv f^{-1}\left(\left[\frac{1}{3}, 1\right]\right)$$

and use Problem 9 if the two sets are nonempty. When this has been done, let

$$\frac{3}{2}(f(x) - g_1(x))$$

play the role of f and let g_2 be like g_1. Obtain

$$\left| f(x) - \sum_{i=1}^{n} \left(\frac{2}{3}\right)^{i-1} g_i(x) \right| \leq \left(\frac{2}{3}\right)^n$$

and consider

$$g(x) \equiv \sum_{i=1}^{\infty} \left(\frac{2}{3}\right)^{i-1} g_i(x).$$

Is it necessary to be in a metric space to do this?

11. ↑ Let M be a closed set in a metric space (X, d) and suppose $f \in C(M)$. Show there exists $g \in C(X)$ such that $g(x) = f(x)$ for all $x \in M$ and if $f(M) \subseteq [a, b]$, then $g(X) \subseteq [a, b]$. This is a version of the Tietze extension theorem. Is it necessary to be in a metric space for this to work?

12. Let X be a compact topological space and suppose $\{f_n\}$ is a sequence of functions continuous on X having values in $\mathbb{R}^n$. Show there exists a countable dense subset of X, $\{x_i\}$ and a subsequence of $\{f_n\}$, $\{f_{n_k}\}$, such that $\{f_{n_k}(x_i)\}$ converges for each x_i. **Hint:** First get a subsequence which converges at x_1, then a subsequence of this subsequence which converges at x_2 and a subsequence of this one which converges at x_3 and so forth. Thus the second of these subsequences converges at both x_1 and x_2 while the third converges at these two points and also at x_3 and so forth. List them so the second is under the first and the third is under the second and so forth thus obtaining an infinite matrix of entries. Now consider the diagonal sequence and argue it is ultimately a subsequence of every one of these subsequences described earlier and so it must converge at each x_i. This procedure is called the Cantor diagonal process.

13. ↑ Use the Cantor diagonal process to give a different proof of the Ascoli Arzela theorem than that presented in this chapter. **Hint:** Start with a sequence of functions in $C(X; \mathbb{R}^n)$ and use the Cantor diagonal process to produce a subsequence which converges at each point of a countable dense subset of X. Then show this sequence is a Cauchy sequence in $C(X; \mathbb{R}^n)$.

14. What about the case where $C_0(X)$ consists of complex valued functions and the field of scalars is $\mathbb{C}$ rather than $\mathbb{R}$? In this case, suppose $\mathcal{A}$ is an algebra of functions in $C_0(X)$ which separates the points, annihilates no point, and has the property that if $f \in \mathcal{A}$, then $\overline{f} \in \mathcal{A}$. Show that $\mathcal{A}$ is dense in $C_0(X)$. **Hint:** Let $\operatorname{Re}\mathcal{A} \equiv \{\operatorname{Re} f : f \in \mathcal{A}\}$, $\operatorname{Im}\mathcal{A} \equiv \{\operatorname{Im} f : f \in \mathcal{A}\}$. Show $\mathcal{A} = \operatorname{Re}\mathcal{A} + i \operatorname{Im}\mathcal{A} = \operatorname{Im}\mathcal{A} + i \operatorname{Re}\mathcal{A}$. Then argue that both $\operatorname{Re}\mathcal{A}$ and $\operatorname{Im}\mathcal{A}$ are real algebras which annihilate no point of X and separate the points of X. Apply the Stone Weierstrass theorem to approximate $\operatorname{Re} f$ and $\operatorname{Im} f$ with functions from these real algebras.

15. Let (X, d) be a metric space where d is a bounded metric. Let $\mathcal{C}$ denote the collection of closed subsets of X. For $A, B \in \mathcal{C}$, define
$$\rho(A, B) \equiv \inf \{\delta > 0 : A_\delta \supseteq B \text{ and } B_\delta \supseteq A\}$$
where for a set S,
$$S_\delta \equiv \{x : dist(x, S) \equiv \inf \{d(x, s) : s \in S\} \leq \delta\}.$$
Show $x \to dist(x, S)$ is continuous and that therefore, S_δ is a closed set containing S. Also show that ρ is a metric on $\mathcal{C}$. This is called the Hausdorff metric.

16. ↑Suppose (X, d) is a compact metric space. Show $(\mathcal{C}, \rho)$ is a complete metric space. **Hint:** Show first that if $W_n \downarrow W$ where W_n is closed, then $\rho(W_n, W) \to 0$. Now let $\{A_n\}$ be a Cauchy sequence in $\mathcal{C}$. Then if $\epsilon > 0$ there exists N such that when $m, n \geq N$, then $\rho(A_n, A_m) < \epsilon$. Therefore, for each $n \geq N$,
$$(A_n)_\epsilon \supseteq \overline{\cup_{k=n}^\infty A_k}.$$
Let $A \equiv \cap_{n=1}^\infty \overline{\cup_{k=n}^\infty A_k}$. By the first part, there exists $N_1 > N$ such that for $n \geq N_1$,
$$\rho\left(\overline{\cup_{k=n}^\infty A_k}, A\right) < \epsilon, \text{ and } (A_n)_\epsilon \supseteq \overline{\cup_{k=n}^\infty A_k}.$$
Therefore, for such n, $A_\epsilon \supseteq W_n \supseteq A_n$ and $(W_n)_\epsilon \supseteq (A_n)_\epsilon \supseteq A$ because
$$(A_n)_\epsilon \supseteq \overline{\cup_{k=n}^\infty A_k} \supseteq A.$$

17. ↑ Let X be a compact metric space. Show $(\mathcal{C}, \rho)$ is compact. **Hint:** Let $\mathcal{D}_n$ be a 2^{-n} net for X. Let $\mathcal{K}_n$ denote finite unions of sets of the form $\overline{B(p, 2^{-n})}$ where $p \in \mathcal{D}_n$. Show $\mathcal{K}_n$ is a $2^{-(n-1)}$ net for $(\mathcal{C}, \rho)$.

Abstract measure and Integration

5.1 σ Algebras

This chapter is on the basics of measure theory and integration. A measure is a real valued mapping from some subset of the power set of a given set which has values in $[0, \infty]$. We will see that many apparently different things can be considered as measures and also that whenever we are in such a measure space defined below, there is an integral defined. By discussing this in terms of axioms and in a very abstract setting, we unify many topics into one general theory. For example, it will turn out that sums are included as an integral of this sort. So is the usual integral as well as things which are often thought of as being in between sums and integrals.

Let Ω be a set and let $\mathcal{F}$ be a collection of subsets of Ω satisfying

$$\emptyset \in \mathcal{F},\ \Omega \in \mathcal{F}, \tag{5.1}$$

$$E \in \mathcal{F} \text{ implies } E^C \equiv \Omega \setminus E \in \mathcal{F},$$

$$\text{If } \{E_n\}_{n=1}^{\infty} \subseteq \mathcal{F}, \text{ then } \cup_{n=1}^{\infty} E_n \in \mathcal{F}. \tag{5.2}$$

Definition 5.1 *A collection of subsets of a set, Ω, satisfying Formulas (5.1)-(5.2) is called a σ algebra.*

As an example, let Ω be any set and let $\mathcal{F} = \mathcal{P}(\Omega)$, the set of all subsets of Ω (power set). This obviously satisfies Formulas (5.1)-(5.2).

Lemma 5.2 *Let $\mathcal{C}$ be a set whose elements are σ algebras of subsets of Ω. Then $\cap\mathcal{C}$ is a σ algebra also.*

Example 5.3 *Let τ denote the collection of all open sets in $\mathbb{R}^n$ and let $\sigma(\tau) \equiv$ intersection of all σ algebras that contain τ. $\sigma(\tau)$ is called the σ algebra of Borel sets .*

This is a very important σ algebra and it will be referred to frequently as the Borel sets. Attempts to describe a typical Borel set are more trouble than they are worth and it is not easy to do so. Rather, one uses the definition just given in the example. Note, however, that all countable intersections of open sets and countable unions of closed sets are Borel sets. Such sets are called G_δ and F_σ respectively.

5.2 Monotone classes and algebras

Definition 5.4 *$\mathcal{A}$ is said to be an algebra of subsets of a set, Z if $Z \in \mathcal{A}, \emptyset \in \mathcal{A}$, and when $E, F \in \mathcal{A}$, $E \cup F$ and $E \setminus F$ are both in $\mathcal{A}$.*

It is important to note that if $\mathcal{A}$ is an algebra, then it is also closed under finite intersections. Thus, $E \cap F = (E^C \cup F^C)^C \in \mathcal{A}$ because $E^C = Z \setminus E \in \mathcal{A}$ and similarly $F^C \in \mathcal{A}$.

Definition 5.5 *$\mathcal{M} \subseteq \mathcal{P}(Z)$ is called a monotone class if*
a.) $\cdots E_n \supseteq E_{n+1} \cdots$, $E = \cap_{n=1}^{\infty} E_n$, and $E_n \in \mathcal{M}$, then $E \in \mathcal{M}$.
b.) $\cdots E_n \subseteq E_{n+1} \cdots$, $E = \cup_{n=1}^{\infty} E_n$, and $E_n \in \mathcal{M}$, then $E \in \mathcal{M}$.
(In simpler notation, $E_n \downarrow E$ and $E_n \in \mathcal{M}$ implies $E \in \mathcal{M}$. $E_n \uparrow E$ and $E_n \in \mathcal{M}$ implies $E \in \mathcal{M}$.)

How can we easily identify algebras? The following lemma is useful for this problem.

Lemma 5.6 *Suppose that $\mathcal{R}$ and $\mathcal{E}$ are subsets of $\mathcal{P}(Z)$ such that $\mathcal{E}$ is defined as the set of all finite disjoint unions of sets of $\mathcal{R}$. Suppose also that*

$$\emptyset, Z \in \mathcal{R}$$

$$A \cap B \in \mathcal{R} \text{ whenever } A, B \in \mathcal{R},$$

$$A \setminus B \in \mathcal{E} \text{ whenever } A, B \in \mathcal{R}.$$

Then $\mathcal{E}$ is an algebra of sets of Z.

Proof: Note first that if $A \in \mathcal{R}$, then $A^C \in \mathcal{E}$ because $A^C = Z \setminus A$. Now suppose that $E_1$and E_2 are in $\mathcal{E}$,

$$E_1 = \cup_{i=1}^{m} R_i, \;\; E_2 = \cup_{j=1}^{n} R_j$$

where the R_i are disjoint sets in $\mathcal{R}$ and the R_j are disjoint sets in $\mathcal{R}$. Then

$$E_1 \cap E_2 = \cup_{i=1}^{m} \cup_{j=1}^{n} R_i \cap R_j$$

which is clearly an element of $\mathcal{E}$ because no two of the sets in the union can intersect and by assumption they are all in $\mathcal{R}$. Thus finite intersections of sets of $\mathcal{E}$ are in $\mathcal{E}$. If $E = \cup_{i=1}^{n} R_i$

$$E^C = \cap_{i=1}^{n} R_i^C = \text{ finite intersection of sets of } \mathcal{E}$$

which was just shown to be in $\mathcal{E}$. Thus if $E_1, E_2 \in \mathcal{E}$,

$$E_1 \setminus E_2 = E_1 \cap E_2^C \in \mathcal{E}$$

and

$$E_1 \cup E_2 = (E_1 \setminus E_2) \cup E_2 \in \mathcal{E}$$

from the definition of $\mathcal{E}$. This proves the lemma.

Corollary 5.7 *Let $(Z_1, \mathcal{R}_1, \mathcal{E}_1)$ and $(Z_2, \mathcal{R}_2, \mathcal{E}_2)$ be as described in Lemma 5.6. Then $(Z_1 \times Z_2, \mathcal{R}, \mathcal{E})$ also satisfies the conditions of Lemma 5.6 if $\mathcal{R}$ is defined as*

$$\mathcal{R} \equiv \{R_1 \times R_2 : R_i \in \mathcal{R}_i\}$$

and

$$\mathcal{E} \equiv \{ \text{ finite disjoint unions of sets of } \mathcal{R}\}.$$

Consequently, $\mathcal{E}$ is an algebra of sets.

Proof: It is clear $\emptyset, Z_1 \times Z_2 \in \mathcal{R}$. Let $R_1^1 \times R_2^1$ and $R_1^2 \times R_2^2$ be two elements of $\mathcal{R}$.

$$R_1^1 \times R_2^1 \cap R_1^2 \times R_2^2 = R_1^1 \cap R_1^2 \times R_2^1 \cap R_2^2 \in \mathcal{R}$$

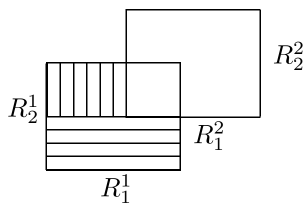
by assumption.

$$R_1^1 \times R_2^1 \setminus \left(R_1^2 \times R_2^2\right) =$$

$$R_1^1 \times \left(R_2^1 \setminus R_2^2\right) \cup \left(R_1^1 \setminus R_1^2\right) \times \left(R_2^2 \cap R_2^1\right)$$

$$= R_1^1 \times A_2 \cup A_1 \times R_2$$

where $A_2 \in \mathcal{E}_2$, $A_1 \in \mathcal{E}_1$, and $R_2 \in \mathcal{R}_2$.



Since the two sets in the above expression on the right do not intersect, and each A_i is a finite union of disjoint elements of $\mathcal{R}_i$, it follows the above expression is in $\mathcal{E}$. This proves the corollary. The following example will be referred to frequently.

Example 5.8 *Consider for $\mathcal{R}$, sets of the form $I = (a, b] \cap (-\infty, \infty)$ where $a \in [-\infty, \infty]$ and $b \in [-\infty, \infty]$. Then, clearly, $\emptyset, (-\infty, \infty) \in \mathcal{R}$ and it is not hard to see that all conditions for Corollary 5.7 are satisfied. Applying this corollary repeatedly, we find that for*

$$\mathcal{R} \equiv \left\{ \prod_{i=1}^{n} I_i : I_i = (a_i, b_i] \cap (-\infty, \infty) \right\}$$

and $\mathcal{E}$ is defined as finite disjoint unions of sets of $\mathcal{R}$,

$$(\mathbb{R}^n, \mathcal{R}, \mathcal{E})$$

satisfies the conditions of Corollary 5.7 and in particular $\mathcal{E}$ is an algebra of sets of $\mathbb{R}^n$. It is clear that the same would hold if I were of the form $[a, b) \cap (-\infty, \infty)$.

Theorem 5.9 *(Monotone Class theorem) Let $\mathcal{A}$ be an algebra of subsets of Z and let $\mathcal{M}$ be a monotone class containing $\mathcal{A}$. Then $\mathcal{M} \supseteq \sigma(\mathcal{A})$, the smallest σ-algebra containing $\mathcal{A}$.*

Proof: We may assume $\mathcal{M}$ is the smallest monotone class containing $\mathcal{A}$. Such a smallest monotone class exists because the intersection of monotone classes containing $\mathcal{A}$ is a monotone class containing $\mathcal{A}$. We show that $\mathcal{M}$ is a σ-algebra. It will then follow $\mathcal{M} \supseteq \sigma(\mathcal{A})$. For $A \in \mathcal{A}$, define

$$\mathcal{M}_A \equiv \{B \in \mathcal{M} \text{ such that } A \cup B \in \mathcal{M}\}.$$

Clearly $\mathcal{M}_A$ is a monotone class containing $\mathcal{A}$. Hence $\mathcal{M}_A = \mathcal{M}$ because $\mathcal{M}$ is the smallest such monotone class. This shows that $A \cup B \in \mathcal{M}$ whenever $A \in \mathcal{A}$ and $B \in \mathcal{M}$. Now pick $B \in \mathcal{M}$ and define

$$\mathcal{M}_B \equiv \{D \in \mathcal{M} \text{ such that } D \cup B \in \mathcal{M}\}.$$

We just showed $\mathcal{A} \subseteq \mathcal{M}_B$. It is clear that $\mathcal{M}_B$ is a monotone class. Thus $\mathcal{M}_B = \mathcal{M}$ and it follows that $D \cup B \in \mathcal{M}$ whenever $D \in \mathcal{M}$ and $B \in \mathcal{M}$.

A similar argument shows that $D \setminus B \in \mathcal{M}$ whenever $D, B \in \mathcal{M}$. (For $A \in \mathcal{A}$, let

$$\mathcal{M}_A = \{B \in \mathcal{M} \text{ such that } B \setminus A \text{ and } A \setminus B \in \mathcal{M}\}.$$

Argue $\mathcal{M}_A$ is a monotone class containing $\mathcal{A}$, etc.)

Thus $\mathcal{M}$ is both a monotone class and an algebra. Hence, if $E \in \mathcal{M}$ then $Z \setminus E \in \mathcal{M}$. We want to show $\mathcal{M}$ is a σ-algebra. But if $E_i \in \mathcal{M}$ and $F_n = \cup_{i=1}^n E_i$, then $F_n \in \mathcal{M}$ and $F_n \uparrow \cup_{i=1}^\infty E_i$. Since $\mathcal{M}$ is a monotone class, $\cup_{i=1}^\infty E_i \in \mathcal{M}$ and so $\mathcal{M}$ is a σ-algebra. This proves the theorem.

Definition 5.10 *Let $\mathcal{F}$ be a σ algebra of sets of Ω and let $\mu : \mathcal{F} \to [0, \infty]$. We call μ a measure if*

$$\mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i) \tag{5.3}$$

whenever the E_i are disjoint sets of $\mathcal{F}$. The triple, $(\Omega, \mathcal{F}, \mu)$ is called a measure space and the elements of $\mathcal{F}$ are called the measurable sets. We say $(\Omega, \mathcal{F}, \mu)$ is a finite measure space when $\mu(\Omega) < \infty$.

Theorem 5.11 *Let $\{E_m\}_{m=1}^\infty$ be a sequence of measurable sets in a measure space $(\Omega, \mathcal{F}, \mu)$. Then if $\cdots E_n \subseteq E_{n+1} \subseteq E_{n+2} \subseteq \cdots$,*

$$\mu(\cup_{i=1}^\infty E_i) = \lim_{n\to\infty} \mu(E_n) \tag{5.4}$$

and if $\cdots E_n \supseteq E_{n+1} \supseteq E_{n+2} \supseteq \cdots$ and $\mu(E_1) < \infty$, then

$$\mu(\cap_{i=1}^\infty E_i) = \lim_{n\to\infty} \mu(E_n). \tag{5.5}$$

Proof: First note that $\cap_{i=1}^\infty E_i = (\cup_{i=1}^\infty E_i^C)^C \in \mathcal{F}$ so $\cap_{i=1}^\infty E_i$ is measurable. To show (5.4), note that (5.4) is obviously true if $\mu(E_k) = \infty$ for any k. Therefore, assume $\mu(E_k) < \infty$ for all k. Thus

$$\mu(E_{k+1} \setminus E_k) = \mu(E_{k+1}) - \mu(E_k).$$

Hence by (5.3),

$$\mu(\cup_{i=1}^\infty E_i) = \mu(E_1) + \sum_{k=1}^\infty \mu(E_{k+1} \setminus E_k) = \mu(E_1)$$

$$+ \sum_{k=1}^\infty \mu(E_{k+1}) - \mu(E_k)$$

$$= \mu(E_1) + \lim_{n\to\infty} \sum_{k=1}^n \mu(E_{k+1}) - \mu(E_k) = \lim_{n\to\infty} \mu(E_{n+1}).$$

This shows part (5.4). To verify (5.5), since $\mu(E_1) < \infty$,

$$\mu(E_1) - \mu(\cap_{i=1}^\infty E_i) = \mu(E_1 \setminus \cap_{i=1}^\infty E_i) = \lim_{n\to\infty} \mu(E_1 \setminus \cap_{i=1}^n E_i)$$

$$= \mu(E_1) - \lim_{n\to\infty} \mu(\cap_{i=1}^n E_i) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n),$$

where the second equality follows from part (5.4). Hence

$$\lim_{n\to\infty} \mu(E_n) = \mu(\cap_{i=1}^\infty E_i).$$

This proves the theorem.

Definition 5.12 *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let (X, τ) be a topological space. A function $f : \Omega \to X$ is said to be measurable if $f^{-1}(U) \in \mathcal{F}$ whenever $U \in \tau$. (Inverse images of open sets are measurable.)*

Note the analogy with a continuous function for which inverse images of open sets are open.

Definition 5.13 *Let $\{a_n\}_{n=1}^{\infty} \subseteq X,\ (X, \tau)$ where X and τ are described above. Then*

$$\lim_{n\to\infty} a_n = a$$

means that whenever $a \in U \in \tau$, there exists n_0 such that if $n > n_0$, then $a_n \in U$. (Every open set containing a also contains a_n for all but finitely many values of n.) Note this agrees with the definition given earlier for $\mathbb{R}^p$, and $\mathbb{C}$ while also giving a definition of what is meant for convergence in general topological spaces.

Recall that (X, τ) has a countable basis if there is a countable subset of τ, $\mathcal{B}$, such that every set of τ is the union of sets of $\mathcal{B}$. We observe that for X given as either $\mathbb{R}$, $\mathbb{C}$, or $[0, \infty]$ with the definition of τ described earlier (a subbasis for the topology of $[0, \infty]$ is sets of the form $[0, b)$ and sets of the form $(b, \infty]$), the following hold.

$$(X, \tau) \text{ has a countable basis, } \mathcal{B}. \tag{5.6}$$

Whenever $U \in \mathcal{B}$, there exists a sequence of open sets, $\{V_m\}_{m=1}^{\infty}$, such that

$$\cdots V_m \subseteq \overline{V}_m \subseteq V_{m+1} \subseteq \cdots,\ U = \bigcup_{m=1}^{\infty} V_m. \tag{5.7}$$

Recall $\overline{S}$ is defined as the union of the set S with all its limit points.

Theorem 5.14 *Let f_n and f be functions mapping Ω to X where $\mathcal{F}$ is a σ algebra of measurable sets of Ω and (X, τ) is a topological space satisfying Formulas (5.6) - (5.7). Then if f_n is measurable, and $f(\omega) = \lim_{n\to\infty} f_n(\omega)$, it follows that f is also measurable. (Pointwise limits of measurable functions are measurable.)*

Proof: Let $\mathcal{B}$ be the countable basis of (5.6) and let $U \in \mathcal{B}$. Let $\{V_m\}$ be the sequence of (5.7). Since f is the pointwise limit of f_n,

$$f^{-1}(V_m) \subseteq \{\omega : f_k(\omega) \in V_m \text{ for all } k \text{ large enough}\} \subseteq f^{-1}(\overline{V}_m).$$

Therefore,

$$f^{-1}(U) = \cup_{m=1}^{\infty} f^{-1}(V_m) \subseteq \cup_{m=1}^{\infty} \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} f_k^{-1}(V_m)$$

$$\subseteq \cup_{m=1}^{\infty} f^{-1}(\bar{V}_m) = f^{-1}(U).$$

It follows $f^{-1}(U) \in \mathcal{F}$ because it equals the expression in the middle which is measurable. Now let $W \in \tau$. Since $\mathcal{B}$ is countable, $W = \cup_{n=1}^{\infty} U_n$ for some sets $U_n \in \mathcal{B}$. Hence

$$f^{-1}(W) = \cup_{n=1}^{\infty} f^{-1}(U_n) \in \mathcal{F}.$$

This proves the theorem.

Example 5.15 *Let $X = [-\infty, \infty]$ and let a basis for a topology, τ, be sets of the form $[-\infty, a)$, (a, b), and $(a, \infty]$. Then it is clear that (X, τ) satisfies Formulas (5.6) - (5.7) with a countable basis, $\mathcal{B}$, given by sets of this form but with a and b rational.*

Definition 5.16 *Let* $f_n : \Omega \to [-\infty, \infty]$.

$$\limsup_{n\to\infty} f_n(\omega) = \lim_{n\to\infty} (\sup\{f_k(\omega) : k \geq n\}). \tag{5.8}$$

$$\liminf_{n\to\infty} f_n(\omega) = \lim_{n\to\infty} (\inf\{f_k(\omega) : k \geq n\}). \tag{5.9}$$

Note that in $[-\infty, \infty]$ with the topology just described, every increasing sequence converges and every decreasing sequence converges. This follows from Definition 5.13. Also, if

$$A_n(\omega) = \inf\{f_k(\omega) : k \geq n\},\ B_n(\omega) = \sup\{f_k(\omega) : k \geq n\}.$$

It is clear that $B_n(\omega)$ is decreasing while $A_n(\omega)$ is increasing. Therefore, Formulas (5.8) and (5.9) always make sense unlike the limit.

Lemma 5.17 *Let* $f : \Omega \to [-\infty, \infty]$ *where* $\mathcal{F}$ *is a* σ *algebra of subsets of* Ω. *Then* f *is measurable if any of the following hold.*

$$f^{-1}((d, \infty]) \in \mathcal{F} \text{ for all finite } d,$$

$$f^{-1}([-\infty, c)) \in \mathcal{F} \text{ for all finite } c,$$

$$f^{-1}([d, \infty]) \in \mathcal{F} \text{ for all finite } d,$$

$$f^{-1}([-\infty, c]) \in \mathcal{F} \text{ for all finite } c.$$

Proof: First note that the first and the third are equivalent. To see this, note

$$f^{-1}([d, \infty]) = \cap_{n=1}^{\infty} f^{-1}((d - 1/n, \infty]),$$

$$f^{-1}((d, \infty]) = \cup_{n=1}^{\infty} f^{-1}([d + 1/n, \infty]).$$

Similarly, the second and fourth conditions are equivalent.

$$f^{-1}([-\infty, c]) = (f^{-1}((c, \infty]))^C$$

so the first and fourth conditions are equivalent. Thus all four conditions are equivalent and if any of them hold,

$$f^{-1}((a, b)) = f^{-1}([-\infty, b)) \cap f^{-1}((a, \infty]) \in \mathcal{F}.$$

Thus $f^{-1}(B) \in \mathcal{F}$ whenever B is a basic open set described in Example 5.15. Since every open set can be obtained as a countable union of these basic open sets, it follows that if any of the four conditions hold, then f is measurable. This proves the lemma.

Theorem 5.18 *Let* $f_n : \Omega \to [-\infty, \infty]$ *be measurable with respect to a* σ *algebra,* $\mathcal{F}$*, of subsets of* Ω. *Then* $\limsup_{n\to\infty} f_n$ *and* $\liminf_{n\to\infty} f_n$ *are measurable.*

Proof: Let $g_n(\omega) = \sup\{f_k(\omega) : k \geq n\}$. Then

$$g_n^{-1}((c, \infty]) = \cup_{k=n}^{\infty} f_k^{-1}((c, \infty]) \in \mathcal{F}.$$

Therefore g_n is measurable.

$$\limsup_{n\to\infty} f_n(\omega) = \lim_{n\to\infty} g_n(\omega)$$

and so by Theorem 5.14 $\limsup_{n\to\infty} f_n$ is measurable. Similar reasoning shows $\liminf_{n\to\infty} f_n$ is measurable.

Theorem 5.19 *Let $f_i, i = 1, \cdots, n$ be a measurable function mapping Ω to the topological space (X, τ) and suppose that τ has a countable basis, $\mathcal{B}$. Then $\mathbf{f} = (f_1 \cdots f_n)^T$ is a measurable function from Ω to $\prod_{i=1}^n X$. (Here it is understood that the topology of $\prod_{i=1}^n X$ is the standard product topology and that $\mathcal{F}$ is the σ algebra of measurable subsets of Ω.)*

Proof: First we observe that sets of the form $\prod_{i=1}^n B_i,\ B_i \in \mathcal{B}$ form a countable basis for the product topology. Now

$$\mathbf{f}^{-1}(\prod_{i=1}^n B_i) = \cap_{i=1}^n f_i^{-1}(B_i) \in \mathcal{F}.$$

Since every open set is a countable union of these sets, it follows $\mathbf{f}^{-1}(U) \in \mathcal{F}$ for all open U.

Theorem 5.20 *Let $(\Omega, \mathcal{F})$ be a measure space and let $f_i, i = 1, \cdots, n$ be measurable functions mapping Ω to (X, τ), a topological space with a countable basis. Let $g : \prod_{i=1}^n X \to X$ be continuous and let $\mathbf{f} = (f_1 \cdots f_n)^T$. Then $g \circ \mathbf{f}$ is a measurable function.*

Proof: Let U be open.

$$(g \circ \mathbf{f})^{-1}(U) = \mathbf{f}^{-1}(g^{-1}(U)) = \mathbf{f}^{-1}(\text{open set}) \in \mathcal{F}$$

by Theorem 5.19.

Example 5.21 *Let $X = (-\infty, \infty]$ with a basis for the topology given by sets of the form (a, b) and $(c, \infty]$, a, b, c rational numbers. Let $+ : X \times X \to X$ be given by $+(x, y) = x + y$. Then $+$ is continuous; so if f, g are measurable functions mapping Ω to X, we may conclude by Theorem 5.20 that $f + g$ is also measurable. Also, if a, b are positive real numbers and $l(x, y) = ax + by$, then $l : X \times X \to X$ is continuous and so $l(f, g) = af + bg$ is measurable.*

Note that the basis given in this example provides the usual notions of convergence in $(-\infty, \infty]$. Theorems 5.19 and 5.20 imply that under appropriate conditions, sums, products, and, more generally, continuous functions of measurable functions are measurable. The following is also interesting.

Theorem 5.22 *Let $f : \Omega \to X$ be measurable. Then $f^{-1}(B) \in \mathcal{F}$ for every Borel set, B, of (X, τ).*

Proof: Let $\mathcal{S} \equiv \{B \subseteq X \text{ such that } f^{-1}(B) \in \mathcal{F}\}$. $\mathcal{S}$ contains all open sets. It is also clear that $\mathcal{S}$ is a σalgebra. Hence $\mathcal{S}$ contains the Borel sets because the Borel sets are defined as the intersection of all σalgebras containing the open sets.

The following theorem is often very useful when dealing with sequences of measurable functions.

Theorem 5.23 *(Egoroff) Let $(\Omega, \mathcal{F}, \mu)$ be a finite measure space*

$$(\mu(\Omega) < \infty)$$

and let f_n, f be complex valued measurable functions such that

$$\lim_{n \to \infty} f_n(\omega) = f(\omega)$$

for all $\omega \notin E$ where $\mu(E) = 0$. Then for every $\varepsilon > 0$, there exists a set,

$$F \supseteq E,\ \mu(F) < \varepsilon,$$

such that f_n converges uniformly to f on F^C.

Proof: Let $E_{km} = \{\omega \in E^C : |f_n(\omega) - f(\omega)| \geq 1/m$ for some $n > k\}$. By Theorems 5.19 and 5.20,

$$\{\omega \in E^C : |f_n(\omega) - f(\omega)| \geq 1/m\}$$

is measurable. Hence E_{km} is measurable because

$$E_{km} = \cup_{n=k+1}^{\infty} \{\omega \in E^C : |f_n(\omega) - f(\omega)| \geq 1/m\}.$$

For fixed $m, \cap_{k=1}^{\infty} E_{km} = \emptyset$ and so it has measure 0. Note also that

$$E_{km} \supseteq E_{(k+1)m}.$$

Since $\mu(E_{1m}) < \infty$,

$$0 = \mu(\cap_{k=1}^{\infty} E_{km}) = \lim_{k \to \infty} \mu(E_{km})$$

by Theorem 5.11. Let $k(m)$ be chosen such that $\mu(E_{k(m)m}) < \varepsilon 2^{-m}$. Let

$$F = E \cup \bigcup_{m=1}^{\infty} E_{k(m)m}.$$

Then $\mu(F) < \varepsilon$ because

$$\mu(F) \leq \mu(E) + \sum_{m=1}^{\infty} \mu\left(E_{k(m)m}\right).$$

Now let $\eta > 0$ be given and pick m_0 such that $m_0^{-1} < \eta$. If $\omega \in F^C$, then

$$\omega \in \bigcap_{m=1}^{\infty} E_{k(m)m}^C.$$

Hence $\omega \in E_{k(m_0)m_0}^C$ so

$$|f_n(\omega) - f(\omega)| < 1/m_0 < \eta$$

for all $n > k(m_0)$. This holds for all $\omega \in F^C$and so f_n converges uniformly to f on F^C. This proves the theorem.

We conclude this chapter with a comment about notation. We say that something happens for μ a.e. ω and say μ almost everywhere if there exists a set E with $\mu(E) = 0$ and the thing takes place for all $\omega \notin E$. Thus $f(\omega) = g(\omega)$ a.e. if $f(\omega) = g(\omega)$ for all $\omega \notin E$ where $\mu(E) = 0$.

We also say a measure space, $(\Omega, \mathcal{F}, \mu)$ is σ finite if there exist measurable sets, Ω_n such that $\mu(\Omega_n) < \infty$ and $\Omega = \cup_{n=1}^{\infty} \Omega_n$.

5.3 Exercises

1. Let $\Omega = \mathbb{N} = \{1, 2, \cdots\}$. Let $\mathcal{F} = \mathcal{P}(\mathbb{N})$ and let $\mu(S) =$ number of elements in S. Thus $\mu(\{1\}) = 1 = \mu(\{2\})$, $\mu(\{1,2\}) = 2$, etc. Show $(\Omega, \mathcal{F}, \mu)$ is a measure space. It is called counting measure.

2. Let Ω be any uncountable set and let $\mathcal{F} = \{A \subseteq \Omega :$ either A or A^C is countable$\}$. Let $\mu(A) = 1$ if A is uncountable and $\mu(A) = 0$ if A is countable. Show $(\Omega, \mathcal{F}, \mu)$ is a measure space.

3. Let $\mathcal{F}$ be a σ algebra of subsets of Ω and suppose $\mathcal{F}$ has infinitely many elements. Show that $\mathcal{F}$ is uncountable.

4. Prove Lemma 5.2.

5. We say g is Borel measurable if whenever U is open, $g^{-1}(U)$ is Borel. Let $f : \Omega \to X$ and let $g : X \to Y$ where X, Y are topological spaces and $\mathcal{F}$ is a σ algebra of sets of Ω. Suppose f is measurable and g is Borel measurable. Show $g \circ f$ is measurable.

6. Let $(\Omega, \mathcal{F})$ be a measure space and suppose $f : \Omega \to \mathbb{C}$. Show f is measurable if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable real-valued functions.

7. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Define $\overline{\mu} : \mathcal{P}(\Omega) \to [0, \infty]$ by
$$\overline{\mu}(A) = \inf\{\mu(B) : B \supseteq A,\ B \in \mathcal{F}\}.$$
Show $\overline{\mu}$ satisfies
$$\overline{\mu}(\emptyset) = 0, \text{ if } A \subseteq B,\ \overline{\mu}(A) \leq \overline{\mu}(B),\ \overline{\mu}(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \overline{\mu}(A_i).$$
If $\overline{\mu}$ satisfies these conditions, it is called an outer measure. This shows every measure determines an outer measure on the power set.

8. Let $\{E_i\}$ be a sequence of measurable sets with the property that
$$\sum_{i=1}^{\infty} \mu(E_i) < \infty.$$
Let $S = \{\omega \in \Omega \text{ such that } \omega \in E_i \text{ for infinitely many values of } i\}$. Show $\mu(S) = 0$ and S is measurable. This is part of the Borel Cantelli lemma.

9. $\uparrow$ Let f_n, f be measurable functions with values in $\mathbb{C}$. We say that f_n converges in measure if
$$\lim_{n\to\infty} \mu(x \in \Omega : |f(x) - f_n(x)| \geq \varepsilon) = 0$$
for each fixed $\varepsilon > 0$. Prove the theorem of F. Riesz. If f_n converges to f in measure, then there exists a subsequence $\{f_{n_k}\}$ which converges to f a.e. **Hint:** Choose n_1 such that
$$\mu(x : |f(x) - f_{n_1}(x)| \geq 1) < 1/2.$$
Choose $n_2 > n_1$ such that
$$\mu(x : |f(x) - f_{n_2}(x)| \geq 1/2) < 1/2^2,$$
$n_3 > n_2$ such that
$$\mu(x : |f(x) - f_{n_3}(x)| \geq 1/3) < 1/2^3,$$
etc. Now consider what it means for $f_{n_k}(x)$ to fail to converge to $f(x)$. Then remember Problem 8.

10. Let $\mathcal{C} \equiv \{E_i\}_{i=1}^{\infty}$ be a countable collection of sets and let $\Omega_1 \equiv \cup_{i=1}^{\infty} E_i$. Show there exists an algebra of sets, $\mathcal{A}$, such that $\mathcal{A} \supseteq \mathcal{C}$ and $\mathcal{A}$ is countable. **Hint:** Let $\mathcal{C}_1$ denote all finite unions of sets of $\mathcal{C}$ and Ω_1. Thus $\mathcal{C}_1$ is countable. Now let $\mathcal{B}_1$ denote all complements with respect to Ω_1 of sets of $\mathcal{C}_1$. Let $\mathcal{C}_2$ denote all finite unions of sets of $\mathcal{B}_1 \cup \mathcal{C}_1$. Continue in this way, obtaining an increasing sequence $\mathcal{C}_n$, each of which is countable. Let
$$\mathcal{A} \equiv \cup_{i=1}^{\infty} \mathcal{C}_i.$$

5.4 The Abstract Lebesgue Integral

In this section we develop the Lebesgue integral and present some of its most important properties. In all that follows μ will be a measure defined on a σ algebra $\mathcal{F}$ of subsets of Ω. We always define $0 \cdot \infty = 0$. This may seem somewhat arbitrary and this is so. However, a little thought will soon demonstrate that this is the right definition for this meaningless expression in the context of measure theory. To see this, consider the zero function defined on $\mathbb{R}$. What do we want the integral of this function to be? Obviously, by an analogy with the Riemann integral, we would want this to equal zero. Formally, it is zero times the length of the set or infinity. The following notation will be used.

For a set E,

$$\mathcal{X}_E(\omega) = \begin{cases} 1 \text{ if } \omega \in E, \\ 0 \text{ if } \omega \notin E. \end{cases}$$

This is called the characteristic function of E.

Definition 5.24 *A function, s, is called simple if it is measurable and has only finitely many values. These values will never be $\pm\infty$.*

Definition 5.25 *If $s(x) \geq 0$ and s is simple,*

$$\int s \equiv \sum_{i=1}^{m} a_i \mu(A_i)$$

where $A_i = \{\omega : s(x) = a_i\}$ and $a_1, \cdots, a_m$ are the distinct values of s.

Note that $\int s$ could equal $+\infty$ if $\mu(A_k) = \infty$ and $a_k > 0$ for some k, but $\int s$ is well defined because $s \geq 0$ and we use the convention that $0 \cdot \infty = 0$.

Lemma 5.26 *If $a, b \geq 0$ and if s and t are nonnegative simple functions, then*

$$\int as + bt \equiv a \int s + b \int t.$$

Proof: Let

$$s(\omega) = \sum_{i=1}^{n} \alpha_i \mathcal{X}_{A_i}(\omega),\ t(\omega) = \sum_{i=1}^{m} \beta_j \mathcal{X}_{B_j}(\omega)$$

where α_i are the distinct values of s and the β_j are the distinct values of t. Clearly $as + bt$ is a nonnegative simple function. Also,

$$(as + bt)(\omega) = \sum_{j=1}^{m} \sum_{i=1}^{n} (a\alpha_i + b\beta_j) \mathcal{X}_{A_i \cap B_j}(\omega)$$

where the sets $A_i \cap B_j$ are disjoint. Now we don't know that all the values $a\alpha_i + b\beta_j$ are distinct, but we note that if $E_1, \cdots, E_r$ are disjoint measurable sets whose union is E, then $\alpha\mu(E) = \alpha \sum_{i=1}^{r} \mu(E_i)$. Thus

$$\begin{aligned} \int as + bt &= \sum_{j=1}^{m} \sum_{i=1}^{n} (a\alpha_i + b\beta_j) \mu(A_i \cap B_j) \\ &= a \sum_{i=1}^{n} \alpha_i \mu(A_i) + b \sum_{j=1}^{m} \beta_j \mu(B_j) \\ &= a \int s + b \int t. \end{aligned}$$

This proves the lemma.

Corollary 5.27 *Let $s = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}$ where $a_i \geq 0$ and the E_i are not necessarily disjoint. Then*

$$\int s = \sum_{i=1}^{n} a_i \mu(E_i).$$

Proof: $\int a\mathcal{X}_{E_i} = a\mu(E_i)$ so this follows from Lemma 5.26.

Now we are ready to define the Lebesgue integral of a nonnegative measurable function.

Definition 5.28 *Let $f : \Omega \to [0, \infty]$ be measurable. Then*

$$\int f d\mu \equiv \sup\{\int s : 0 \leq s \leq f, \ s \ \textit{simple}\}.$$

Lemma 5.29 *If $s \geq 0$ is a nonnegative simple function, $\int s d\mu = \int s$. Moreover, if $f \geq 0$, then $\int f d\mu \geq 0$.*

Proof: The second claim is obvious. To verify the first, suppose $0 \leq t \leq s$ and t is simple. Then clearly $\int t \leq \int s$ and so

$$\int s d\mu = \sup\{\int t : 0 \leq t \leq s, \ t \text{ simple}\} \leq \int s.$$

But $s \leq s$ and s is simple so $\int s d\mu \geq \int s$.

The next theorem is one of the big results that justifies the use of the Lebesgue integral.

Theorem 5.30 *(Monotone Convergence theorem) Let $f \geq 0$ and suppose $\{f_n\}$ is a sequence of nonnegative measurable functions satisfying*

$$\lim_{n\to\infty} f_n(\omega) = f(\omega) \textit{ for each } \omega.$$

$$\cdots f_n(\omega) \leq f_{n+1}(\omega) \cdots \tag{5.10}$$

Then f is measurable and

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

Proof: First note that f is measurable by Theorem 5.14 since it is the limit of measurable functions. It is also clear from (5.10) that $\lim_{n\to\infty} \int f_n d\mu$ exists because $\{\int f_n d\mu\}$ forms an increasing sequence. This limit may be $+\infty$ but in any case,

$$\lim_{n\to\infty} \int f_n d\mu \leq \int f d\mu$$

because $\int f_n d\mu \leq \int f d\mu$.

Let $\delta \in (0, 1)$ and let s be a simple function with

$$0 \leq s(\omega) \leq f(\omega), \quad s(\omega) = \sum_{i=1}^{r} \alpha_i \mathcal{X}_{A_i}(\omega).$$

Then $(1 - \delta)s(\omega) \leq f(\omega)$ for all ω with strict inequality holding whenever $f(\omega) > 0$. Let

$$E_n = \{\omega : f_n(\omega) \geq (1 - \delta)s(\omega)\} \tag{5.11}$$

Then

$$\cdots E_n \subseteq E_{n+1} \cdots, \text{ and } \cup_{n=1}^{\infty} E_n = \Omega.$$

Therefore

$$\lim_{n\to\infty} \int s\mathcal{X}_{E_n} = \int s.$$

This follows from Theorem 5.11 which implies that $\alpha_i \mu(E_n \cap A_i) \to \alpha_i \mu(A_i)$. Thus, from (5.11)

$$\int f d\mu \geq \int f_n d\mu \geq \int f_n \mathcal{X}_{E_n} d\mu \geq (\int s\mathcal{X}_{E_n} d\mu)(1-\delta). \tag{5.12}$$

Letting $n \to \infty$ in (5.12) we see that

$$\int f d\mu \geq \lim_{n\to\infty} \int f_n d\mu \geq (1-\delta) \int s. \tag{5.13}$$

Now let $\delta \downarrow 0$ in (5.13) to obtain

$$\int f d\mu \geq \lim_{n\to\infty} \int f_n d\mu \geq \int s.$$

Now s was an arbitrary simple function less than or equal to f. Hence,

$$\int f d\mu \geq \lim_{n\to\infty} \int f_n d\mu \geq \sup\{\int s : 0 \leq s \leq f, \ s \text{ simple}\} \equiv \int f d\mu.$$

This proves the theorem.

The next theorem will be used frequently. It says roughly that measurable functions are pointwise limits of simple functions. This is similar to continuous functions being the limit of step functions.

Theorem 5.31 *Let $f \geq 0$ be measurable. Then there exists a sequence of simple functions $\{s_n\}$ satisfying*

$$0 \leq s_n(\omega) \tag{5.14}$$

$$\cdots s_n(\omega) \leq s_{n+1}(\omega) \cdots$$

$$f(\omega) = \lim_{n\to\infty} s_n(\omega) \text{ for all } \omega \in \Omega. \tag{5.15}$$

Before proving this, we give a definition.

Definition 5.32 *If f, g are functions having values in $[0, \infty]$,*

$$f \vee g = \max(f, g), \ f \wedge g = \min(f, g).$$

Note that if f, g have finite values,

$$f \vee g = 2^{-1}(f + g + |f - g|), \ f \wedge g = 2^{-1}(f + g - |f - g|).$$

From this observation, the following lemma is obvious.

Lemma 5.33 *If s, t are nonnegative simple functions, then*

$$s \vee t, \ s \wedge t$$

are also simple functions. (Recall $+\infty$ is not a value of either s or t.)

Proof of Theorem 5.31: Let

$$I = \{x : f(x) = +\infty\}.$$

Let $E_{nk} = f^{-1}([\frac{k}{n}, \frac{k+1}{n}))$. Let

$$t_n(\omega) = \sum_{k=0}^{2^n} \frac{k}{n} \mathcal{X}_{E_{nk}}(\omega) + n\mathcal{X}_I(\omega).$$

Then $t_n(\omega) \leq f(\omega)$ for all ω and $\lim_{n\to\infty} t_n(\omega) = f(\omega)$ for all ω. This is because $t_n(\omega) = n$ for $\omega \in I$ and if $f(\omega) \in [0, \frac{2^n+1}{n})$, then

$$0 \leq f_n(\omega) - t_n(\omega) \leq \frac{1}{n}.$$

Thus whenever $\omega \notin I$, the above inequality will hold for all n large enough. Let

$$s_1 = t_1,\ s_2 = t_1 \vee t_2,\ s_3 = t_1 \vee t_2 \vee t_3, \cdots.$$

Then the sequence $\{s_n\}$ satisfies Formulas (5.14)-(5.15) and this proves the theorem.

Next we show that the integral is linear on nonnegative functions. Roughly speaking, it shows the integral is trying to be linear and is only prevented from being linear at this point by not yet being defined on functions which could be negative or complex valued. We will define the integral for these functions soon and then this lemma will be the key to showing the integral is linear.

Lemma 5.34 *Let $f, g \geq 0$ be measurable. Let $a, b \geq 0$ be constants. Then*

$$\int (af + bg)d\mu = a\int f d\mu + b\int g d\mu.$$

Proof: Let $\{s_n\}$ and $\{\tilde{s}_n\}$ be increasing sequences of simple functions such that

$$\lim_{n\to\infty} s_n(\omega) = f(\omega),\ \lim_{n\to\infty} \tilde{s}_n(\omega) = g(\omega).$$

Then by the monotone convergence theorem and Lemma 5.26,

$$\begin{aligned}
\int (af + bg)d\mu &= \lim_{n\to\infty} \int (as_n + b\tilde{s}_n)d\mu \\
&= \lim_{n\to\infty} \int as_n + b\tilde{s}_n = \lim_{n\to\infty} a\int s_n + b\int \tilde{s}_n \\
&= \lim_{n\to\infty} a\int s_n d\mu + b\int \tilde{s}_n d\mu = a\int f d\mu + b\int g d\mu.
\end{aligned}$$

This proves the lemma.

5.5 The space L^1

Now suppose f has complex values and is measurable. We need to define what is meant by the integral of such functions. First some theorems about measurability need to be shown.

Theorem 5.35 *Let $f = u + iv$ where u, v are real-valued functions. Then f is a measurable $\mathbb{C}$ valued function if and only if u and v are both measurable $\mathbb{R}$ valued functions.*

Proof: Suppose first that f is measurable. Let $V \subseteq \mathbb{R}$ be open.

$$u^{-1}(V) = \{\omega : u(\omega) \in V\} = \{\omega : f(\omega) \in V + i\mathbb{R}\} \in \mathcal{F},$$

$$v^{-1}(V) = \{\omega : v(\omega) \in V\} = \{\omega : f(\omega) \in \mathbb{R}+iV\} \in \mathcal{F}.$$

Now suppose u and v are real and measurable.

$$f^{-1}((a,b) + i(c,d)) = u^{-1}(a,b) \cap v^{-1}(c,d) \in \mathcal{F}.$$

Since every open set in $\mathbb{C}$ may be written as a countable union of open sets of the form $(a,b)+i(c,d)$, it follows that $f^{-1}(U) \in \mathcal{F}$ whenever U is open in $\mathbb{C}$. This proves the theorem.

Definition 5.36 *$L^1(\Omega)$ is the space of complex valued measurable functions, f, satisfying*

$$\int |f(\omega)| d\mu < \infty.$$

We also write the symbol, $||f||_{L^1}$ to denote $\int |f(\omega)| \, d\mu$.

Note that if $f : \Omega \to \mathbb{C}$ is measurable, then by Theorem 5.20, $|f| : \Omega \to \mathbb{R}$ is also measurable.

Definition 5.37 *If u is real-valued,*

$$u^+ \equiv u \vee 0, \; u^- \equiv -(u \wedge 0).$$

Thus u^+ and u^- are both nonnegative and

$$u = u^+ - u^-, \; |u| = u^+ + u^-.$$

Definition 5.38 *Let $f = u + iv$ where u, v are real-valued. Suppose $f \in L^1(\Omega)$. Then*

$$\int f d\mu \equiv \int u^+ d\mu - \int u^- d\mu + i[\int v^+ d\mu - \int v^- d\mu].$$

Note that all this is well defined because $\int |f| d\mu < \infty$ and so

$$\int u^+ d\mu, \int u^- d\mu, \int v^+ d\mu, \int v^- d\mu$$

are all finite. The next theorem shows the integral is linear on $L^1(\Omega)$.

Theorem 5.39 *$L^1(\Omega)$ is a complex vector space and if $a, b \in \mathbb{C}$ and*

$$f, g \in L^1(\Omega),$$

then

$$\int af + bg d\mu = a \int f d\mu + b \int g d\mu. \tag{5.16}$$

Proof: First suppose f, g are real-valued and in $L^1(\Omega)$. We note that

$$h^+ = 2^{-1}(h + |h|),\ h^- = 2^{-1}(|h| - h)$$

whenever h is real-valued. Consequently,

$$f^+ + g^+ - (f^- + g^-) = (f + g)^+ - (f + g)^- = f + g.$$

Hence

$$f^+ + g^+ + (f + g)^- = (f + g)^+ + f^- + g^-. \tag{5.17}$$

From Lemma 5.34,

$$\int f^+ d\mu + \int g^+ d\mu + \int (f + g)^- d\mu = \int f^- d\mu + \int g^- d\mu + \int (f + g)^+ d\mu. \tag{5.18}$$

Since all integrals are finite,

$$\begin{aligned} \int (f+g) d\mu &\equiv \int (f+g)^+ d\mu - \int (f+g)^- d\mu \qquad (5.19) \\ &= \int f^+ d\mu + \int g^+ d\mu - (\int f^- d\mu + \int g^- d\mu) \\ &\equiv \int f d\mu + \int g d\mu. \end{aligned}$$

Now suppose that c is a real constant and f is real-valued. Note

$$(cf)^- = -cf^+ \text{ if } c < 0,\ \ (cf)^- = cf^- \text{ if } c \geq 0.$$

$$(cf)^+ = -cf^- \text{ if } c < 0,\ \ (cf)^+ = cf^+ \text{ if } c \geq 0.$$

If $c < 0$, we use the above and Lemma 5.34 to write

$$\begin{aligned} \int cf d\mu &\equiv \int (cf)^+ d\mu - \int (cf)^- d\mu \\ &= -c \int f^- d\mu + c \int f^+ d\mu \equiv c \int f d\mu. \end{aligned}$$

Similarly, if $c \geq 0$,

$$\begin{aligned} \int cf d\mu &\equiv \int (cf)^+ d\mu - \int (cf)^- d\mu \\ &= c \int f^+ d\mu - c \int f^- d\mu \equiv c \int f d\mu. \end{aligned}$$

This shows (5.16) holds if f, g, a, and b are all real-valued. To conclude, let $a = \alpha + i\beta,\ f = u + iv$ and use the preceding.

$$\begin{aligned} \int af d\mu &= \int (\alpha + i\beta)(u + iv) d\mu \\ &= \int (\alpha u - \beta v) + i(\beta u + \alpha v) d\mu \\ &= \alpha \int u d\mu - \beta \int v d\mu + i\beta \int u d\mu + i\alpha \int v d\mu \\ &= (\alpha + i\beta)(\int u d\mu + i \int v d\mu) = a \int f d\mu. \end{aligned}$$

Thus (5.16) holds whenever f, g, a, and b are complex valued. It is obvious that $L^1(\Omega)$ is a vector space. This proves the theorem.

The next theorem, known as Fatou's lemma is another important theorem which justifies the use of the Lebesgue integral.

Theorem 5.40 *(Fatou's lemma) Let f_n be a nonnegative measurable function with values in $[0,\infty]$. Let $g(\omega) = \liminf_{n\to\infty} f_n(\omega)$. Then g is measurable and*

$$\int g d\mu \leq \lim \inf_{n\to\infty} \int f_n d\mu.$$

Proof: Let $g_n(\omega) = \inf\{f_k(\omega) : k \geq n\}$. Then

$$g_n^{-1}([a,\infty]) = \cap_{k=n}^{\infty} f_k^{-1}([a,\infty]) \in \mathcal{F}.$$

Thus g_n is measurable by Lemma 5.17. Also $g(\omega) = \lim_{n\to\infty} g_n(\omega)$ so g is measurable because it is the pointwise limit of measurable functions. Now the functions g_n form an increasing sequence of nonnegative measurable functions so the monotone convergence theorem applies. This yields

$$\int g d\mu = \lim_{n\to\infty} \int g_n d\mu \leq \lim \inf_{n\to\infty} \int f_n d\mu.$$

The last inequality holding because

$$\int g_n d\mu \leq \int f_n d\mu.$$

This proves the Theorem.

Theorem 5.41 *(Triangle inequality) Let $f \in L^1(\Omega)$. Then*

$$|\int f d\mu| \leq \int |f| d\mu.$$

Proof: $\int f d\mu \in \mathbb{C}$ so there exists $\alpha \in \mathbb{C}$, $|\alpha| = 1$ such that $|\int f d\mu| = \alpha \int f d\mu = \int \alpha f d\mu$. Hence

$$\begin{aligned} |\int f d\mu| &= \int \alpha f d\mu = \int (\mathrm{Re}(\alpha f) + i\,\mathrm{Im}(\alpha f)) d\mu \\ &= \int \mathrm{Re}(\alpha f) d\mu = \int (\mathrm{Re}(\alpha f))^+ d\mu - \int (\mathrm{Re}(\alpha f))^- d\mu \\ &\leq \int (\mathrm{Re}(\alpha f))^+ + (\mathrm{Re}(\alpha f))^- d\mu \leq \int |\alpha f| d\mu = \int |f| d\mu \end{aligned}$$

which proves the theorem.

Theorem 5.42 *(Dominated Convergence theorem) Let $f_n \in L^1(\Omega)$ and suppose*

$$f(\omega) = \lim_{n\to\infty} f_n(\omega),$$

and there exists a measurable function g, with values in $[0,\infty]$, such that

$$|f_n(\omega)| \leq g(\omega) \text{ and } \int g(\omega) d\mu < \infty.$$

Then $f \in L^1(\Omega)$ and

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

Proof: f is measurable by Theorem 5.14. Since $|f| \leq g$, it follows that

$$f \in L^1(\Omega) \text{ and } |f - f_n| \leq 2g.$$

By Fatou's lemma (Theorem 5.40),

$$\begin{aligned} \int 2g d\mu &\leq \liminf_{n\to\infty} \int 2g - |f - f_n| d\mu \\ &= \int 2g d\mu - \limsup_{n\to\infty} \int |f - f_n| d\mu. \end{aligned}$$

Subtracting $\int 2g d\mu$,

$$0 \leq -\limsup_{n\to\infty} \int |f - f_n| d\mu.$$

Hence

$$0 \geq \limsup_{n\to\infty} \left(\int |f - f_n| d\mu \right) \geq \limsup_{n\to\infty} \left| \int f d\mu - \int f_n d\mu \right|$$

which proves the theorem.

Definition 5.43 *Let E be a measurable subset of Ω.*

$$\int_E f d\mu \equiv \int f \mathcal{X}_E d\mu.$$

Also we may refer to $L^1(E)$. The σ algebra in this case is just

$$\{E \cap A : A \in \mathcal{F}\}$$

and the measure is μ restricted to this smaller σ algebra. Clearly, if $f \in L^1(\Omega)$, then

$$f \mathcal{X}_E \in L^1(E)$$

and if $f \in L^1(E)$, then letting $\tilde{f}$ be the 0 extension of f off of E, we see that $\tilde{f} \in L^1(\Omega)$.

5.6 Double sums of nonnegative terms

The definition of the Lebesgue integral and the monotone convergence theorem imply that the order of summation of a double sum of nonnegative terms can be interchanged and in fact the terms can be added in any order. To see this, let $\Omega = \mathbb{N} \times \mathbb{N}$ and let μ be counting measure defined on the set of all subsets of $\mathbb{N} \times \mathbb{N}$. Thus, $\mu(E) =$ the number of elements of E. Then $(\Omega, \mu, \mathcal{P}(\Omega))$ is a measure space and if $a : \Omega \to [0, \infty]$, then a is a measurable function. Following the usual notation, $a_{ij} \equiv a(i, j)$.

Theorem 5.44 *Let $a : \Omega \to [0, \infty]$. Then*

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} = \int a d\mu = \sum_{k=1}^{\infty} a(\theta(k))$$

where θ is any one to one and onto map from $\mathbb{N}$ to Ω.

Proof: By the definition of the integral,

$$\sum_{j=1}^{n}\sum_{i=1}^{l}a_{ij}\leq\int ad\mu$$

for any n,l. Therefore, by the definition of what is meant by an infinite sum,

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}\leq\int ad\mu.$$

Now let $s\leq a$ and s is a nonnegative simple function. If $s\left(i,j\right)>0$ for infinitely many values of $(i,j)\in\Omega$, then

$$\int s=\infty=\int ad\mu=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}.$$

Therefore, it suffices to assume $s\left(i,j\right)>0$ for only finitely many values of $(i,j)\in\mathbb{N}\times\mathbb{N}$. Hence, for some $n>1$,

$$\int s\leq\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij}\leq\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}.$$

Since s is an arbitrary nonnegative simple function, this shows

$$\int ad\mu\leq\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}$$

and so

$$\int ad\mu=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}.$$

The same argument holds if i and j are interchanged which verifies the first two equalities in the conclusion of the theorem. The last equation in the conclusion of the theorem follows from the monotone convergence theorem.

5.7 Vitali convergence theorem

In this section we consider a remarkable convergence theorem which, in the case of finite measure spaces turns out to be better than the dominated convergence theorem.

Definition 5.45 *Let $(\Omega,\mathcal{F},\mu)$ be a measure space and let $\mathfrak{S}\subseteq L^{1}(\Omega)$. We say that $\mathfrak{S}$ is uniformly integrable if for every $\varepsilon>0$ there exists $\delta>0$ such that for all $f\in\mathfrak{S}$*

$$|\int_{E}fd\mu|<\varepsilon\text{ whenever }\mu(E)<\delta.$$

Lemma 5.46 *If $\mathfrak{S}$ is uniformly integrable, then $|\mathfrak{S}|\equiv\{|f|:f\in\mathfrak{S}\}$ is uniformly integrable. Also $\mathfrak{S}$ is uniformly integrable if $\mathfrak{S}$ is finite.*

Proof: Let $\varepsilon > 0$ be given and suppose $\mathfrak{S}$ is uniformly integrable. First suppose the functions are real valued. Let δ be such that if $\mu(E) < \delta$, then

$$\left|\int_E f d\mu\right| < \frac{\varepsilon}{2}$$

for all $f \in \mathfrak{S}$. Let $\mu(E) < \delta$. Then if $f \in \mathfrak{S}$,

$$\begin{aligned}\int_E |f|\, d\mu &\leq \int_{E\cap[f\leq 0]} (-f)\, d\mu + \int_{E\cap[f>0]} f d\mu \\ &= \left|\int_{E\cap[f\leq 0]} f d\mu\right| + \left|\int_{E\cap[f>0]} f d\mu\right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

In the above, $[f > 0]$ is short for $\{\omega \in \Omega : f(\omega) > 0\}$ with a similar definition holding for $[f \leq 0]$. In general, if $\mathfrak{S}$ is uniformly integrable, then $\operatorname{Re}\mathfrak{S} \equiv \{\operatorname{Re} f : f \in \mathfrak{S}\}$ and $\operatorname{Im}\mathfrak{S} \equiv \{\operatorname{Im} f : f \in \mathfrak{S}\}$ are easily seen to be uniformly integrable. Therefore, applying the above result for real valued functions to these sets of functions, it is routine to verify that $|\mathfrak{S}|$ is uniformly integrable also.

For the last part, is suffices to verify a single function in $L^1(\Omega)$ is uniformly integrable. To do so, note that from the dominated convergence theorem,

$$\lim_{R\to\infty} \int_{[|f|>R]} |f|\, d\mu = 0.$$

Let $\varepsilon > 0$ be given and choose R large enough that $\int_{[|f|>R]} |f|\, d\mu < \frac{\varepsilon}{2}$. Now let $\mu(E) < \frac{\varepsilon}{2R}$. Then

$$\begin{aligned}\int_E |f|\, d\mu &= \int_{E\cap[|f|\leq R]} |f|\, d\mu + \int_{E\cap[|f|>R]} |f|\, d\mu \\ &< R\mu(E) + \frac{\varepsilon}{2} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

This proves the lemma.

The following theorem is Vitali's convergence theorem.

Theorem 5.47 *Let $\{f_n\}$ be a uniformly integrable set of complex valued functions, $\mu(\Omega) < \infty$, and $f_n(x) \to f(x)$ a.e. where f is a measurable complex valued function. Then $f \in L^1(\Omega)$ and*

$$\lim_{n\to\infty} \int_\Omega |f_n - f| d\mu = 0. \tag{5.20}$$

Proof: First we show that $f \in L^1(\Omega)$. By uniform integrability, there exists $\delta > 0$ such that if $\mu(E) < \delta$, then

$$\int_E |f_n|\, d\mu < 1$$

for all n. By Egoroff's theorem, there exists a set, E of measure less than δ such that on E^C, $\{f_n\}$ converges uniformly. Therefore, if we pick p large enough, and let $n > p$,

$$\int_{E^C} |f_p - f_n|\, d\mu < 1$$

which implies

$$\int_{E^C} |f_n| \, d\mu < 1 + \int_{\Omega} |f_p| \, d\mu.$$

Then since there are only finitely many functions, f_n with $n \leq p$, we have the existence of a constant, M_1 such that for all n,

$$\int_{E^C} |f_n| \, d\mu < M_1.$$

But also, we have

$$\begin{aligned} \int_{\Omega} |f_m| \, d\mu &= \int_{E^C} |f_m| \, d\mu + \int_E |f_m| \\ &\leq M_1 + 1 \equiv M. \end{aligned}$$

Therefore, by Fatou's lemma,

$$\int_{\Omega} |f| \, d\mu \leq \lim \inf_{n \to \infty} \int |f_n| \, d\mu \leq M,$$

showing that $f \in L^1$ as hoped.

Now $\mathfrak{S} \cup \{f\}$ is uniformly integrable so there exists $\delta_1 > 0$ such that if $\mu(E) < \delta_1$, then $\int_E |g| \, d\mu < \varepsilon/3$ for all $g \in \mathfrak{S} \cup \{f\}$. Now by Egoroff's theorem, there exists a set, F with $\mu(F) < \delta_1$ such that f_n converges uniformly to f on F^C. Therefore, there exists N such that if $n > N$, then

$$\int_{F^C} |f - f_n| \, d\mu < \frac{\varepsilon}{3}.$$

It follows that for $n > N$,

$$\begin{aligned} \int_{\Omega} |f - f_n| \, d\mu &\leq \int_{F^C} |f - f_n| \, d\mu + \int_F |f| \, d\mu + \int_F |f_n| \, d\mu \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

which verifies (5.20).

5.8 The ergodic theorem

This section deals with a fundamental convergence theorem known as the ergodic theorem. It will only be used in one place later in the book so you might omit this topic on a first reading or pick it up when you need it later. I am putting it here because it seems to fit in well with the material of this chapter.

In this section $(\Omega, \mathcal{F}, \mu)$ will be a probability measure space. This means that $\mu(\Omega) = 1$. The mapping, $T : \Omega \to \Omega$ will satisfy the following condition.

$$T^{-1}(A) \in \mathcal{F} \text{ whenever } A \in \mathcal{F}, \ T \text{ is one to one.} \tag{5.21}$$

Lemma 5.48 *If T satisfies (5.21), then if f is measurable, $f \circ T$ is measurable.*

Proof: Let U be an open set. Then

$$(f \circ T)^{-1}(U) = T^{-1}\left(f^{-1}(U)\right) \in \mathcal{F}$$

by (5.21).

Now suppose that in addition to (5.21) T also satisfies

$$\mu\left(T^{-1}A\right)=\mu\left(A\right), \tag{5.22}$$

for all $A\in\mathcal{F}$. In words, T^{-1} is measure preserving. Then for T satisfying (5.21) and (5.22), we have the following simple lemma.

Lemma 5.49 *If T satisfies (5.21) and (5.22) then whenever f is nonnegative and mesurable,*

$$\int_\Omega f\left(\omega\right)d\mu=\int_\Omega f\left(T\omega\right)d\mu. \tag{5.23}$$

Also (5.23) holds whenever $f\in L^1\left(\Omega\right)$.

Proof: Let $f\geq 0$ and f is measurable. By Theorem 5.31, let s_n be an increasing sequence of simple functions converging pointwise to f. Then by (5.22) it follows

$$\int s_n\left(\omega\right)d\mu=\int s_n\left(T\omega\right)d\mu$$

and so by the Monotone convergence theorem,

$$\int_\Omega f\left(\omega\right)d\mu=\lim_{n\to\infty}\int_\Omega s_n\left(\omega\right)d\mu$$

$$=\lim_{n\to\infty}\int_\Omega s_n\left(T\omega\right)d\mu=\int_\Omega f\left(T\omega\right)d\mu.$$

Splitting $f\in L^1$ into real and imaginary parts we apply the above to the positive and negative parts of these and obtain (5.23) in this case also.

Definition 5.50 *A measurable function, f, is said to be invariant if*

$$f\left(T\omega\right)=f\left(\omega\right).$$

A set, $A\in\mathcal{F}$ is said to be invariant if $\mathcal{X}_A$ is an invariant function. Thus a set is invariant if and only if $T^{-1}A=A$.

The following theorem, the individual ergodic theorem, is the main result.

Theorem 5.51 *Let $(\Omega,\mathcal{F},\mu)$ be a probability space and let $T:\Omega\to\Omega$ satisfy (5.21) and (5.22). Then if $f\in L^1\left(\Omega\right)$ having real or complex values and*

$$S_nf\left(\omega\right)\equiv\sum_{k=1}^n f\left(T^{k-1}\omega\right),\ S_0f\left(\omega\right)\equiv 0, \tag{5.24}$$

it follows there exists a set of measure zero, N, and an invariant function g such that for all $\omega\notin N$,

$$\lim_{n\to\infty}\frac{1}{n}S_nf\left(\omega\right)=g\left(\omega\right). \tag{5.25}$$

and also

$$\lim_{n\to\infty}\frac{1}{n}S_nf=g\ \text{in}\ L^1\left(\Omega\right)$$

Proof: The proof of this theorem will make use of the following functions.

$$M_n f(\omega) \equiv \sup\{S_k f(\omega) : \ 0 \leq k \leq n\} \tag{5.26}$$

$$M_\infty f(\omega) \equiv \sup\{S_k f(\omega) : \ 0 \leq \ k\}. \tag{5.27}$$

We will also define the following for h a measurable real valued function.

$$[h > 0] \equiv \{\omega \in \Omega : h(\omega) > 0\}.$$

Now if A is an invariant set,

$$S_n(\mathcal{X}_A f)(\omega) \equiv \sum_{k=1}^{n} f\left(T^{k-1}\omega\right) \mathcal{X}_A\left(T^{k-1}\omega\right)$$

$$= \mathcal{X}_A(\omega) \sum_{k=1}^{n} f\left(T^{k-1}\omega\right) = \mathcal{X}_A(\omega) S_n f(\omega).$$

Therefore, for such an invariant set,

$$M_n(\mathcal{X}_A f)(\omega) = \mathcal{X}_A(\omega) M_n f(\omega), \ M_\infty(\mathcal{X}_A f)(\omega) = \mathcal{X}_A(\omega) M_\infty f(\omega). \tag{5.28}$$

Let $-\infty < a < b < \infty$ and define

$$N_{ab} \equiv \left\{\omega \in \Omega : -\infty < \liminf_{n\to\infty} \frac{1}{n} S_n f(\omega) < a\right.$$

$$\left. < b < \limsup_{n\to\infty} \frac{1}{n} S_n f(\omega) < \infty\right\}. \tag{5.29}$$

Observe that from the definition, if $|f(\omega)| \neq \pm\infty$,

$$\liminf_{n\to\infty} \frac{1}{n} S_n f(\omega) = \liminf_{n\to\infty} \frac{1}{n} S_n f(T\omega)$$

and

$$\limsup_{n\to\infty} \frac{1}{n} S_n f(\omega) = \limsup_{n\to\infty} \frac{1}{n} S_n f(T\omega).$$

Therefore, $TN_{ab} = N_{ab}$ so N_{ab} is an invariant set because T is one to one. Also,

$$N_{ab} \subseteq [M_\infty(f - b) > 0] \cap [M_\infty(a - f) > 0].$$

Consequently,

$$\int_{N_{ab}} (f(\omega) - b)\, d\mu = \int_{\left[\mathcal{X}_{N_{ab}} M_\infty (f-b) > 0\right]} \mathcal{X}_{N_{ab}}(\omega)(f(\omega) - b)\, d\mu$$

$$= \int_{\left[M_\infty\left(\mathcal{X}_{N_{ab}}(f-b)\right) > 0\right]} \mathcal{X}_{N_{ab}}(\omega)(f(\omega) - b)\, d\mu \tag{5.30}$$

and

$$\int_{N_{ab}} (a - f(\omega))\, d\mu = \int_{\left[\mathcal{X}_{N_{ab}} M_\infty (a-f)>0\right]} \mathcal{X}_{N_{ab}}(\omega)(a - f(\omega))\, d\mu$$

$$= \int_{\left[M_\infty\left(\mathcal{X}_{N_{ab}}(a-f)\right)>0\right]} \mathcal{X}_{N_{ab}}(\omega)(a - f(\omega))\, d\mu. \tag{5.31}$$

We will complete the proof with the aid of the following lemma which implies the last terms in (5.30) and (5.31) are nonnegative.

Lemma 5.52 *Let $f \in L^1(\mu)$. Then*

$$\int_{[M_\infty f>0]} f d\mu \geq 0.$$

We postpone the proof of this lemma till we have completed the proof of the ergodic theorem. From (5.30), (5.31), and Lemma 5.52,

$$a\mu(N_{ab}) \geq \int_{N_{ab}} f d\mu \geq b\mu(N_{ab}). \tag{5.32}$$

Since $a < b$, it follows that $\mu(N_{ab}) = 0$. Now let

$$N \equiv \cup \{N_{ab} : a < b,\ a, b \in \mathbb{Q}\}.$$

Since $f \in L^1(\Omega)$ and has complex values, it follows that $\mu(N) = 0$. Now $TN_{a,b} = N_{a,b}$ and so

$$T(N) = \cup_{a,b} T(N_{a,b}) = \cup_{a,b} N_{a,b} = N.$$

Therefore, N is measurable and has measure zero. Also, $T^n N = N$ for all $n \in \mathbb{N}$ and so

$$N \equiv \cup_{n=1}^\infty T^n N.$$

For $\omega \notin N, \lim_{n\to\infty} \frac{1}{n} S_n f(\omega)$ exists. Now let

$$g(\omega) \equiv \begin{cases} 0 \text{ if } \omega \in N \\ \lim_{n\to\infty} \frac{1}{n} S_n f(\omega) \ \text{ if } \omega \notin N \end{cases}.$$

Then it is clear g satisfies the conditions of the theorem because if $\omega \in N$, then $T\omega \in N$ also and so in this case, $g(T\omega) = g(\omega)$. On the other hand, if $\omega \notin N$, then

$$g(T\omega) = \lim_{n\to\infty} \frac{1}{n} S_n f(T\omega) = \lim_{n\to\infty} \frac{1}{n} S_n f(\omega) = g(\omega).$$

The last claim follows from the Vitali convergence theorem if we verify the sequence, $\left\{\frac{1}{n} S_n f\right\}_{n=1}^\infty$ is uniformly integrable. To see this is the case, we know $f \in L^1(\Omega)$ and so if $\epsilon > 0$ is given, there exists $\delta > 0$ such that whenever $B \in \mathcal{F}$ and $\mu(B) \leq \delta$, then $\left|\int_B f(\omega)\, d\mu\right| < \epsilon$. Now by approximating the positive and negative parts of f with simple functions we see that

$$\int_A f\left(T^{k-1}\omega\right) d\mu = \int_{T^{-(k-1)}A} f(\omega)\, d\mu.$$

Taking $\mu(A) < \delta$, it follows

$$\left|\int_A \frac{1}{n} S_n f(\omega)\, d\mu\right| \leq \left|\frac{1}{n}\sum_{k=1}^{n}\int_A f\left(T^{k-1}\omega\right) d\mu\right|$$

$$= \left|\frac{1}{n}\sum_{k=1}^{n}\int_{T^{-(k-1)}A} f(\omega)\, d\mu\right| \leq \frac{1}{n}\sum_{k=1}^{n}\left|\int_{T^{-(k-1)}A} f(\omega)\, d\mu\right| < \frac{1}{n}\sum_{k=1}^{n}\epsilon = \epsilon$$

because $\mu\left(T^{-(k-1)}A\right) = \mu(A)$ by assumption. This proves the above sequence is uniformly integrable and so, by the Vitali convergence theorem,

$$\lim_{n\to\infty}\left|\left|\frac{1}{n}S_n f - g\right|\right|_{L^1} = 0.$$

This proves the theorem.

It remains to prove the lemma.

Proof of Lemma 5.52**:** First note that $M_n f(\omega) \geq 0$ for all n and ω. This follows easily from the observation that by definition, $S_0 f(\omega) = 0$ and so $M_n f(\omega)$ is at least as large. Also note that the sets, $[M_n f > 0]$ are increasing in n and their union is $[M_\infty f > 0]$. Therefore, it suffices to show that for all $n > 0$,

$$\int_{[M_n f > 0]} f d\mu \geq 0.$$

Let $T^* h \equiv h \circ T$. Thus T^* maps measurable functions to measurable functions by Lemma 5.48. It is also clear that if $h \geq 0$, then $T^* h \geq 0$ also. Therefore,

$$S_k f(\omega) = f(\omega) + T^* S_{k-1} f(\omega) \leq f(\omega) + T^* M_n f$$

and therefore,

$$M_n f(\omega) \leq f(\omega) + T^* M_n f(\omega).$$

Now

$$\int_\Omega M_n f(\omega)\, d\mu = \int_{[M_n f > 0]} M_n f(\omega)\, d\mu$$

$$\leq \int_{[M_n f > 0]} f(\omega)\, d\mu + \int_\Omega T^* M_n f(\omega)\, d\mu$$

$$= \int_{[M_n f > 0]} f(\omega)\, d\mu + \int_\Omega M_n f(\omega)\, d\mu$$

by Lemma 5.49. This proves the lemma.

The following is a simple corollary which follows from the above theorem.

Corollary 5.53 *The conclusion of Theorem 5.51 holds if μ is only assumed to be a finite measure.*

Definition 5.54 *We say a set, $A \in \mathcal{F}$ is invariant if $T(A) = A$. We say the above mapping, T, is ergodic, if the only invariant sets have measure 0 or 1.*

If the map, T is ergodic, the following corollary holds.

Corollary 5.55 *In the situation of Theorem 5.51, if T is ergodic, then*

$$g(\omega) = \int f(\omega)\, d\mu$$

for a.e. ω.

Proof: Let g be the function of Theorem 5.51 and let R_1 be a rectangle in $\mathbb{R}^2 = \mathbb{C}$ of the form $[-a, a] \times [-a, a]$ such that $g^{-1}(R_1)$ has measure greater than 0. This set is invariant because the function, g is invariant and so it must have measure 1. Divide R_1 into four equal rectangles, R_1', R_2', R_3', R_4'. Then one of these, renamed R_2 has the property that $g^{-1}(R_2)$ has positive measure. Therefore, since the set is invariant, it must have measure 1. Continue in this way obtaining a sequence of closed rectangles, $\{R_i\}$ such that the diamter of R_i converges to zero and $g^{-1}(R_i)$ has measure 1. Then let $c = \cap_{j=1}^{\infty} R_j$. We know $\mu\left(g^{-1}(c)\right) = \lim_{n\to\infty} \mu\left(g^{-1}(R_i)\right) = 1$. It follows that $g(\omega) = c$ for a.e. ω. Now from Theorem 5.51,

$$c = \int c d\mu = \lim_{n\to\infty} \frac{1}{n} \int S_n f d\mu = \int f d\mu.$$

This proves the corollary.

5.9 Exercises

1. Let $\Omega = \mathbb{N} = \{1, 2, \cdots\}$ and $\mu(S) =$ number of elements in S. If

$$f : \Omega \to \mathbb{C}$$

   what do we mean by $\int f d\mu$? Which functions are in $L^1(\Omega)$?

2. Give an example of a measure space, $(\Omega, \mu, \mathcal{F})$, and a sequence of nonnegative measurable functions $\{f_n\}$ converging pointwise to a function f, such that inequality is obtained in Fatou's lemma.

3. Fill in all the details of the proof of Lemma 5.46.

4. Suppose (Ω, μ) is a finite measure space and $\mathfrak{S} \subseteq L^1(\Omega)$. Show $\mathfrak{S}$ is uniformly integrable and bounded in $L^1(\Omega)$ if there exists an increasing function h which satisfies

$$\lim_{t\to\infty} \frac{h(t)}{t} = \infty,\ \sup\left\{\int_\Omega h(|f|)\, d\mu : f \in \mathfrak{S}\right\} < \infty.$$

   When we say $\mathfrak{S}$ is bounded we mean there is some number, M such that

$$\int |f|\, d\mu \leq M$$

   for all $f \in \mathfrak{S}$.

5. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and suppose $f \in L^1(\Omega)$ has the property that whenever $\mu(E) > 0$,

$$\frac{1}{\mu(E)} \left| \int_E f d\mu \right| \leq C.$$

   Show $|f(\omega)| \leq C$ a.e.

6. Let $\{a_n\}, \{b_n\}$ be sequences in $[-\infty, \infty]$. Show

$$\limsup_{n\to\infty} (-a_n) = -\liminf_{n\to\infty} (a_n)$$

$$\limsup_{n\to\infty} (a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$$

provided no sum is of the form $\infty - \infty$. Also show strict inequality can hold in the inequality. State and prove corresponding statements for lim inf.

7. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and suppose $f, g : \Omega \to [-\infty, \infty]$ are measurable. Prove the sets

$$\{\omega : f(\omega) < g(\omega)\} \text{ and } \{\omega : f(\omega) = g(\omega)\}$$

are measurable.

8. Let $\{f_n\}$ be a sequence of real or complex valued measurable functions. Let

$$S = \{\omega : \{f_n(\omega)\} \text{ converges}\}.$$

Show S is measurable.

9. In the monotone convergence theorem

$$0 \le \cdots \le f_n(\omega) \le f_{n+1}(\omega) \le \cdots.$$

The sequence of functions is increasing. In what way can "increasing" be replaced by "decreasing"?

10. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and suppose f_n converges uniformly to f and that f_n is in $L^1(\Omega)$. When can we conclude that

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu?$$

11. Suppose $u_n(t)$ is a differentiable function for $t \in (a, b)$ and suppose that for $t \in (a, b)$,

$$|u_n(t)|,\ |u_n'(t)| < K_n$$

where $\sum_{n=1}^{\infty} K_n < \infty$. Show

$$(\sum_{n=1}^{\infty} u_n(t))' = \sum_{n=1}^{\infty} u_n'(t).$$

The Construction Of Measures

6.1 Outer measures

We have impressive theorems about measure spaces and the abstract Lebesgue integral but a paucity of interesting examples. In this chapter, we discuss the method of outer measures due to Caratheodory (1918). This approach shows how to obtain measure spaces starting with an outer measure. This will then be used to construct measures determined by positive linear functionals.

Definition 6.1 *Let Ω be a nonempty set and let $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ satisfy*

$$\mu(\emptyset) = 0,$$

$$\text{If } A \subseteq B, \text{ then } \mu(A) \leq \mu(B),$$

$$\mu(\cup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mu(E_i).$$

Such a function is called an outer measure. For $E \subseteq \Omega$, we say E is μ measurable if for all $S \subseteq \Omega$,

$$\mu(S) = \mu(S \setminus E) + \mu(S \cap E). \tag{6.1}$$

To help in remembering (6.1), think of a measurable set, E, as a knife which is used to divide an arbitrary set, S, into the pieces, $S \setminus E$ and $S \cap E$. If E is a sharp knife, the amount of stuff after cutting is the same as the amount you started with. The measurable sets are like sharp knives. The idea is to show that the measurable sets form a σ algebra. First we give a definition and a lemma.

Definition 6.2 $(\mu \lfloor S)(A) \equiv \mu(S \cap A)$ *for all $A \subseteq \Omega$. Thus $\mu \lfloor S$ is the name of a new outer measure.*

Lemma 6.3 *If A is μ measurable, then A is $\mu \lfloor S$ measurable.*

Proof: Suppose A is μ measurable. We need to show that for all $T \subseteq \Omega$,

$$(\mu \lfloor S)(T) = (\mu \lfloor S)(T \cap A) + (\mu \lfloor S)(T \setminus A).$$

Thus we need to show

$$\mu(S \cap T) = \mu(T \cap A \cap S) + \mu(T \cap S \cap A^C). \tag{6.2}$$

But we know (6.2) holds because A is measurable. Apply Definition 6.1 to $S \cap T$ instead of S.

The next theorem is the main result on outer measures. It is a very general result which applies whenever one has an outer measure on the power set of any set. This theorem will be referred to as Caratheodory's procedure in the rest of the book.

Theorem 6.4 *The collection of μ measurable sets, $\mathcal{S}$, forms a σ algebra and*

$$\text{If } F_i \in \mathcal{S},\ F_i \cap F_j = \emptyset,\ \text{then } \mu(\cup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} \mu(F_i). \tag{6.3}$$

If $\cdots F_n \subseteq F_{n+1} \subseteq \cdots$, then if $F = \cup_{n=1}^{\infty} F_n$ and $F_n \in \mathcal{S}$, it follows that

$$\mu(F) = \lim_{n\to\infty} \mu(F_n). \tag{6.4}$$

If $\cdots F_n \supseteq F_{n+1} \supseteq \cdots$, and if $F = \cap_{n=1}^{\infty} F_n$ for $F_n \in \mathcal{S}$ then if $\mu(F_1) < \infty$, we may conclude that

$$\mu(F) = \lim_{n\to\infty} \mu(F_n). \tag{6.5}$$

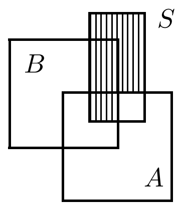
Also, $(\mathcal{S}, \mu)$ is complete. By this we mean that if $F \in \mathcal{S}$ and if $E \subseteq \Omega$ with $\mu(E \setminus F) + \mu(F \setminus E) = 0$, then $E \in \mathcal{S}$.

Proof: First note that $\emptyset$ and Ω are obviously in $\mathcal{S}$. Now suppose that $A, B \in \mathcal{S}$. We show $A \setminus B = A \cap B^C$ is in $\mathcal{S}$. Using the assumption that $B \in \mathcal{S}$ in the second equation below, in which $S \cap A$ plays the role of S in the definition for B being μ measurable,

$$\mu(S \cap (A \cap B^C)) + \mu(S \setminus (A \cap B^C)) = \mu(S \cap A \cap B^C) + \mu(S \cap (A^C \cup B))$$

$$= \mu(S \cap (A^C \cup B)) + \overbrace{\mu(S \cap A) - \mu(S \cap A \cap B)}^{=\mu(S\cap A\cap B^C)}. \tag{6.6}$$

The following picture of $S \cap (A^C \cup B)$ may be of use.



From the picture, and the measurability of A, we see that (6.6) is no larger than

$$\begin{aligned} &\leq \overbrace{\mu(S \cap A \cap B) + \mu(S \setminus A)}^{\leq \mu(S\cap(A^C\cup B))} + \overbrace{\mu(S \cap A) - \mu(S \cap A \cap B)}^{=\mu(S\cap A\cap B^C)} \\ &= \mu(S \setminus A) + \mu(S \cap A) = \mu(S). \end{aligned}$$

This has shown that if $A, B \in \mathcal{S}$, then $A \setminus B \in \mathcal{S}$. Since $\Omega \in \mathcal{S}$, this shows that $A \in \mathcal{S}$ if and only if $A^C \in \mathcal{S}$. Now if $A, B \in \mathcal{S}$, $A \cup B = (A^C \cap B^C)^C = (A^C \setminus B)^C \in \mathcal{S}$. By induction, if $A_1, \cdots, A_n \in \mathcal{S}$, then so is $\cup_{i=1}^{n} A_i$. If $A, B \in \mathcal{S}$, with $A \cap B = \emptyset$,

$$\mu(A \cup B) = \mu((A \cup B) \cap A) + \mu((A \cup B) \setminus A) = \mu(A) + \mu(B).$$

By induction, if $A_i \cap A_j = \emptyset$ and $A_i \in \mathcal{S}$, $\mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$.

Now let $A = \cup_{i=1}^{\infty} A_i$ where $A_i \cap A_j = \emptyset$ for $i \neq j$.

$$\sum_{i=1}^{\infty} \mu(A_i) \geq \mu(A) \geq \mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i).$$

Since this holds for all n, we can take the limit as $n \to \infty$ and conclude,

$$\sum_{i=1}^{\infty} \mu(A_i) = \mu(A)$$

which establishes (6.3). Part (6.4) follows from part (6.3) just as in the proof of Theorem 5.11.

In order to establish (6.5), let the F_n be as given there. Then, since $(F_1 \setminus F_n)$ increases to $(F_1 \setminus F)$, we may use part (6.4) to conclude

$$\lim_{n\to\infty} (\mu(F_1) - \mu(F_n)) = \mu(F_1 \setminus F).$$

Now $\mu(F_1 \setminus F) + \mu(F) \geq \mu(F_1)$ and so $\mu(F_1 \setminus F) \geq \mu(F_1) - \mu(F)$. Hence

$$\lim_{n\to\infty} (\mu(F_1) - \mu(F_n)) = \mu(F_1 \setminus F) \geq \mu(F_1) - \mu(F)$$

which implies

$$\lim_{n\to\infty} \mu(F_n) \leq \mu(F).$$

But since $F \subseteq F_n$, we also have

$$\mu(F) \leq \lim_{n\to\infty} \mu(F_n)$$

and this establishes (6.5).

It remains to show $\mathcal{S}$ is closed under countable unions. We already know that if $A \in \mathcal{S}$, then $A^C \in \mathcal{S}$ and $\mathcal{S}$ is closed under finite unions. Let $A_i \in \mathcal{S}$, $A = \cup_{i=1}^{\infty} A_i$, $B_n = \cup_{i=1}^{n} A_i$. Then

$$\begin{aligned} \mu(S) &= \mu(S \cap B_n) + \mu(S \setminus B_n) \\ &= (\mu\lfloor S)(B_n) + (\mu\lfloor S)(B_n^C). \end{aligned} \tag{6.7}$$

By Lemma 6.3 we know B_n is $(\mu\lfloor S)$ measurable and so is B_n^C. We want to show $\mu(S) \geq \mu(S \setminus A) + \mu(S \cap A)$. If $\mu(S) = \infty$, there is nothing to prove. Assume $\mu(S) < \infty$. Then we apply Parts (6.5) and (6.4) to (6.7) and let $n \to \infty$. Thus

$$B_n \uparrow A, \;\; B_n^C \downarrow A^C$$

and this yields $\mu(S) = (\mu\lfloor S)(A) + (\mu\lfloor S)(A^C) = \mu(S \cap A) + \mu(S \setminus A)$.

Thus $A \in \mathcal{S}$ and this proves Parts (6.3), (6.4), and (6.5).

Let $F \in \mathcal{S}$ and let $\mu(E \setminus F) + \mu(F \setminus E) = 0$. Then

$$\begin{aligned} \mu(S) &\leq \mu(S \cap E) + \mu(S \setminus E) \\ &= \mu(S \cap E \cap F) + \mu(S \cap E \cap F^C) + \mu(S \cap E^C) \\ &\leq \mu(S \cap F) + \mu(E \setminus F) + \mu(S \setminus F) + \mu(F \setminus E) \\ &= \mu(S \cap F) + \mu(S \setminus F) = \mu(S). \end{aligned}$$

Hence $\mu(S) = \mu(S \cap E) + \mu(S \setminus E)$ and so $E \in \mathcal{S}$. This shows that $(\mathcal{S}, \mu)$ is complete.

Where do outer measures come from? One way to obtain an outer measure is to start with a measure μ, defined on a σ algebra of sets, $\mathcal{S}$, and use the following definition of the outer measure induced by the measure.

Definition 6.5 *Let μ be a measure defined on a σ algebra of sets, $\mathcal{S} \subseteq \mathcal{P}(\Omega)$. Then the outer measure induced by μ, denoted by $\overline{\mu}$ is defined on $\mathcal{P}(\Omega)$ as*

$$\overline{\mu}(E) = \inf\{\mu(V) \;: V \in \mathcal{S} \text{ and } V \supseteq E\}.$$

We also say a measure space, $(\mathcal{S}, \Omega, \mu)$ is σ finite if there exist measurable sets, Ω_i with $\mu(\Omega_i) < \infty$ and $\Omega = \cup_{i=1}^{\infty} \Omega_i$.

The following lemma deals with the outer measure generated by a measure which is σ finite. It says that if the given measure is σ finite and complete then no new measurable sets are gained by going to the induced outer measure and then considering the measurable sets in the sense of Caratheodory.

Lemma 6.6 *Let $(\Omega, \mathcal{S}, \mu)$ be any measure space and let $\overline{\mu} : \mathcal{P}(\Omega) \to [0, \infty]$ be the outer measure induced by μ. Then $\overline{\mu}$ is an outer measure as claimed and if $\overline{\mathcal{S}}$ is the set of $\overline{\mu}$ measurable sets in the sense of Caratheodory, then $\overline{\mathcal{S}} \supseteq \mathcal{S}$ and $\overline{\mu} = \mu$ on $\mathcal{S}$. Furthermore, if μ is σ finite and $(\Omega, \mathcal{S}, \mu)$ is complete, then $\overline{\mathcal{S}} = \mathcal{S}$.*

Proof: It is easy to see that $\overline{\mu}$ is an outer measure. Let $E \in \mathcal{S}$. We need to show $E \in \overline{\mathcal{S}}$ and $\overline{\mu}(E) = \mu(E)$. Let $S \subseteq \Omega$. We need to show

$$\overline{\mu}(S) \geq \overline{\mu}(S \cap E) + \overline{\mu}(S \setminus E). \tag{6.8}$$

If $\overline{\mu}(S) = \infty$, there is nothing to prove, so assume $\overline{\mu}(S) < \infty$. Thus there exists $T \in \mathcal{S}$, $T \supseteq S$, and

$$\begin{aligned} \overline{\mu}(S) &> \mu(T) - \varepsilon = \mu(T \cap E) + \mu(T \setminus E) - \varepsilon \\ &\geq \overline{\mu}(T \cap E) + \overline{\mu}(T \setminus E) - \varepsilon \\ &\geq \overline{\mu}(S \cap E) + \overline{\mu}(S \setminus E) - \varepsilon. \end{aligned}$$

Since ε is arbitrary, this proves (6.8) and verifies $\mathcal{S} \subseteq \overline{\mathcal{S}}$. Now if $E \in \mathcal{S}$ and $V \supseteq E$,

$$\mu(E) \leq \mu(V).$$

Hence, taking inf,

$$\mu(E) \leq \overline{\mu}(E).$$

But also $\mu(E) \geq \overline{\mu}(E)$ since $E \in \mathcal{S}$ and $E \supseteq E$. Hence

$$\overline{\mu}(E) \leq \mu(E) \leq \overline{\mu}(E).$$

Now suppose $(\Omega, \mathcal{S}, \mu)$ is complete. Thus if $E, D \in \mathcal{S}$, and $\mu(E \setminus D) = 0$, then if $D \subseteq F \subseteq E$, it follows

$$F \in \mathcal{S} \tag{6.9}$$

because

$$F \setminus D \subseteq E \setminus D \in \mathcal{S},$$

a set of measure zero. Therefore,

$$F \setminus D \in \mathcal{S}$$

and so $F = D \cup (F \setminus D) \in \mathcal{S}$.

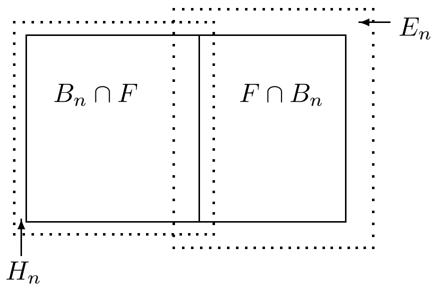
We know already that $\overline{\mathcal{S}} \supseteq \mathcal{S}$ so let $F \in \overline{\mathcal{S}}$. Using the assumption that the measure space is σ finite, let $\{B_n\} \subseteq \mathcal{S}$, $\cup B_n = \Omega$, $B_n \cap B_m = \emptyset$, $\mu(B_n) < \infty$. Let

$$E_n \supseteq F \cap B_n \ , \ \mu(E_n) = \bar{\mu}(F \cap B_n), \tag{6.10}$$

where $E_n \in \mathcal{S}$, and let

$$H_n \supseteq B_n \setminus F = B_n \cap F^C \ , \ \mu(H_n) = \bar{\mu}(B_n \setminus F), \tag{6.11}$$

where $H_n \in \mathcal{S}$. The following picture may be helpful in visualizing this situation.



Thus

$$H_n \supseteq B_n \cap F^C$$

and so

$$H_n^C \subseteq B_n^C \cup F$$

which implies

$$H_n^C \cap B_n \subseteq F \cap B_n.$$

We have

$$H_n^C \cap B_n \subseteq F \cap B_n \subseteq E_n,\ H_n^C \cap B_n, E_n \in \mathcal{S}. \tag{6.12}$$

Claim: If $A, B, D \in \overline{\mathcal{S}}$ and if $A \supseteq B$ with $\overline{\mu}(A \setminus B) = 0$. Then $\overline{\mu}(A \cap D) = \overline{\mu}(B \cap D)$.
Proof of claim: This follows from the observation that $(A \cap D) \setminus (B \cap D) \subseteq A \setminus B$.
Now from (6.10) and (6.11) and this claim,

$$\mu\left(E_n \setminus \left(H_n^C \cap B_n\right)\right) = \overline{\mu}\left((F \cap B_n) \setminus \left(H_n^C \cap B_n\right)\right) = \overline{\mu}\left(F \cap B_n \cap \left(B_n^C \cup H_n\right)\right)$$

$$= \overline{\mu}(F \cap H_n \cap B_n) = \overline{\mu}\left(F \cap \left(B_n \cap F^C\right) \cap B_n\right) = \overline{\mu}(\emptyset) = 0.$$

Therefore, from (6.9) and (6.12) $F \cap B_n \in \mathcal{S}$. Therefore,

$$F = \cup_{n=1}^{\infty} F \cap B_n \in \mathcal{S}.$$

This proves the lemma.

Note that it was not necessary to assume μ was σ finite in order to consider $\overline{\mu}$ and conclude that $\overline{\mu} = \mu$ on $\mathcal{S}$. This is sometimes referred to as the process of completing a measure because $\overline{\mu}$ is a complete measure and $\overline{\mu}$ extends μ.

6.2 Positive linear functionals

One of the most important theorems related to the construction of measures is the Riesz. representation theorem. The situation is that there exists a positive linear functional Λ defined on the space $C_c(\Omega)$ where Ω is a topological space of some sort and Λ is said to be a positive linear functional if it satisfies the following definition.

Definition 6.7 *Let Ω be a topological space. We say $f : \Omega \to \mathbb{C}$ is in $C_c(\Omega)$ if f is continuous and*

$$spt(f) \equiv \overline{\{x \in \Omega : f(x) \neq 0\}}$$

is a compact set. (The symbol, $spt(f)$ is read as "support of f ".) If we write $C_c(V)$ for V an open set, we mean that $spt(f) \subseteq V$ and We say Λ is a positive linear functional defined on $C_c(\Omega)$ if Λ is linear,

$$\Lambda(af + bg) = a\Lambda f + b\Lambda g$$

for all $f, g \in C_c(\Omega)$ and $a, b \in \mathbb{C}$. It is called positive because

$$\Lambda f \geq 0 \text{ whenever } f(x) \geq 0 \text{ for all } x \in \Omega.$$

The most general versions of the theory about to be presented involve locally compact Hausdorff spaces but here we will assume the topological space is a metric space, (Ω, d) which is also σ compact, defined below, and has the property that the closure of any open ball, $\overline{B(x,r)}$ is compact.

Definition 6.8 *We say a topological space, Ω, is σ compact if $\Omega = \cup_{k=1}^{\infty} \Omega_k$ where Ω_k is a compact subset of Ω.*

To begin with we need some technical results and notation. In all that follows, Ω will be a σ compact metric space with the property that the closure of any open ball is compact. An obvious example of such a thing is any closed subset of $\mathbb{R}^n$ or $\mathbb{R}^n$ itself and it is these cases which interest us the most. The terminology of metric spaces is used because it is convenient and contains all the necessary ideas for the proofs which follow while being general enough to include the cases just described.

Definition 6.9 *If K is a compact subset of an open set, V, we say $K \prec \phi \prec V$ if*

$$\phi \in C_c(V),\ \phi(K) = \{1\},\ \phi(\Omega) \subseteq [0,1].$$

Also for $\phi \in C_c(\Omega)$, we say $K \prec \phi$ if

$$\phi(\Omega) \subseteq [0,1] \text{ and } \phi(K) = 1.$$

We say $\phi \prec V$ if

$$\phi(\Omega) \subseteq [0,1] \text{ and } spt(\phi) \subseteq V.$$

The next theorem is a very important result known as the partition of unity theorem. Before we present it, we need a simple lemma which will be used repeatedly.

Lemma 6.10 *Let K be a compact subset of the open set, V. Then there exists an open set, W such that $\overline{W}$ is a compact set and*

$$K \subseteq W \subseteq \overline{W} \subseteq V.$$

Also, if K and V are as just described there exists a continuous function, ψ such that $K \prec \psi \prec V$.

Proof: For each $k \in K$, let $B(k, r_k) \equiv B_k$ be such that $\overline{B_k} \subseteq V$. Since K is compact, finitely many of these balls, $B_{k_1}, \cdots, B_{k_l}$ cover K. Let $W \equiv \cup_{i=1}^{l} B_{k_i}$. Then it follows that $\overline{W} = \cup_{i=1}^{l} \overline{B_{k_i}}$ and satisfies the conclusion of the lemma. Now we define ψ as

$$\psi(x) \equiv \frac{dist\left(x, W^C\right)}{dist\left(x, W^C\right) + dist(x, K)}.$$

Note the denominator is never equal to zero because if $dist(x, K) = 0$, then $x \in W$ and so is at a positive distance from W^C because W is open. This proves the lemma. Also note that $spt(\psi) = \overline{W}$.

Theorem 6.11 *(Partition of unity) Let K be a compact subset of Ω and suppose*

$$K \subseteq V = \cup_{i=1}^{n} V_i, \ V_i \text{ open.}$$

Then there exist $\psi_i \prec V_i$ with

$$\sum_{i=1}^{n} \psi_i(x) = 1$$

for all $x \in K$.

Proof: Let $K_1 = K \setminus \cup_{i=2}^{n} V_i$. Thus K_1 is compact and $K_1 \subseteq V_1$. By the above lemma, we let

$$K_1 \subseteq W_1 \subseteq \overline{W}_1 \subseteq V_1$$

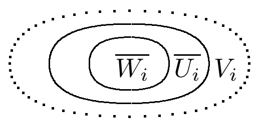
with $\overline{W}_1$compact and f be such that $K_1 \prec f \prec V_1$ with

$$W_1 \equiv \{x : f(x) \neq 0\}.$$

Thus $W_1, V_2, \cdots, V_n$ covers K and $\overline{W}_1 \subseteq V_1$. Let

$$K_2 = K \setminus (\cup_{i=3}^{n} V_i \cup W_1).$$

Then K_2 is compact and $K_2 \subseteq V_2$. Let $K_2 \subseteq W_2 \subseteq \overline{W}_2 \subseteq V_2$, $\overline{W}_2$ compact. Continue this way finally obtaining $W_1, \cdots, W_n$, $K \subseteq W_1 \cup \cdots \cup W_n$, and $\overline{W}_i \subseteq V_i$ $\overline{W}_i$ compact. Now let $\overline{W}_i \subseteq U_i \subseteq \overline{U}_i \subseteq V_i$,$\overline{U}_i$ compact.



By the lemma again, we may define ϕ_i and γ such that

$$\overline{U}_i \prec \phi_i \prec V_i, \ \cup_{i=1}^{n} \overline{W}_i \prec \gamma \prec \cup_{i=1}^{n} U_i.$$

Now define

$$\psi_i(x) = \begin{cases} \gamma(x)\phi_i(x)/\sum_{j=1}^{n} \phi_j(x) \text{ if } \sum_{j=1}^{n} \phi_j(x) \neq 0, \\ 0 \text{ if } \ \sum_{j=1}^{n} \phi_j(x) = 0. \end{cases}$$

If x is such that $\sum_{j=1}^{n} \phi_j(x) = 0$, then $x \notin \cup_{i=1}^{n} \overline{U}_i$. Consequently $\gamma(y) = 0$ for all y near x and so $\psi_i(y) = 0$ for all y near x. Hence ψ_i is continuous at such x. If $\sum_{j=1}^{n} \phi_j(x) \neq 0$, this situation persists near x and so ψ_i is continuous at such points. Therefore ψ_i is continuous. If $x \in K$, then $\gamma(x) = 1$ and so $\sum_{j=1}^{n} \psi_j(x) = 1$. Clearly $0 \leq \psi_i(x) \leq 1$ and $spt(\psi_j) \subseteq V_j$. This proves the theorem.

We don't need the following corollary at this time but it is useful later.

Corollary 6.12 *If H is a compact subset of V_i, we can pick our partition of unity in such a way that $\psi_i(x) = 1$ for all $x \in H$ in addition to the conclusion of Theorem 6.11.*

Proof: Keep V_i the same but replace V_j with $\widetilde{V_j} \equiv V_j \setminus H$. Now in the proof above, applied to this modified collection of open sets, we see that if $j \neq i, \phi_j(x) = 0$ whenever $x \in H$. Therefore, $\psi_i(x) = 1$ on H.

Next we consider a fundamental theorem known as Caratheodory's criterion which gives an easy to check condition which, if satisfied by an outer measure defined on the power set of a metric space, implies that the σ algebra of measurable sets contains the Borel sets.

Definition 6.13 *For two sets, A, B in a metric space, we define*

$$dist\,(A,B) \equiv \inf\{d\,(x,y) : x \in A, y \in B\}.$$

Theorem 6.14 *Let μ be an outer measure on the subsets of (X,d), a metric space. If*

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

whenever $dist(A,B) > 0$, then the σ algebra of measurable sets contains the Borel sets.

Proof: We only need show that closed sets are in $\mathcal{S}$, the σ-algebra of measurable sets, because then the open sets are also in $\mathcal{S}$ and so $\mathcal{S} \supseteq$ Borel sets. Let K be closed and let S be a subset of Ω. We need to show $\mu(S) \geq \mu(S \cap K) + \mu(S \setminus K)$. Therefore, we may assume without loss of generality that $\mu(S) < \infty$. Let

$$K_n = \{x : dist(x,K) \leq \frac{1}{n}\} = \text{closed set}$$

(Recall that $x \to dist\,(x,K)$ is continuous.)

$$\mu(S) \geq \mu((S \cap K) \cup (S \setminus K_n)) = \mu(S \cap K) + \mu(S \setminus K_n) \tag{6.13}$$

by assumption, since $S \cap K$ and $S \setminus K_n$ are a positive distance apart. Now

$$\mu(S \setminus K_n) \leq \mu(S \setminus K) \leq \mu(S \setminus K_n) + \mu((K_n \setminus K) \cap S). \tag{6.14}$$

We look at $\mu((K_n \setminus K) \cap S)$. Note that since K is closed, a point, $x \notin K$ must be at a positive distance from K and so

$$K_n \setminus K = \cup_{k=n}^{\infty} K_k \setminus K_{k+1}.$$

Therefore

$$\mu(S \cap (K_n \setminus K)) \leq \sum_{k=n}^{\infty} \mu(S \cap (K_k \setminus K_{k+1})). \tag{6.15}$$

Now

$$\sum_{k=1}^{\infty} \mu(S \cap (K_k \setminus K_{k+1})) = \sum_{k\ even} \mu(S \cap (K_k \setminus K_{k+1})) +$$

$$+ \sum_{k\ odd} \mu(S \cap (K_k \setminus K_{k+1})). \tag{6.16}$$

Note that if $A = \cup_{i=1}^{\infty} A_i$ and the distance between any pair of sets is positive, then

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i),$$

because

$$\sum_{i=1}^{\infty} \mu(A_i) \geq \mu(A) \geq \mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i).$$

Therefore, from (6.16),

$$\sum_{k=1}^{\infty} \mu(S \cap (K_k \setminus K_{k+1}))$$

$$= \mu(\bigcup_{k\ even} S \cap (K_k \setminus K_{k+1})) + \mu(\bigcup_{k\ odd} S \cap (K_k \setminus K_{k+1}))$$

$$< 2\mu(S) < \infty.$$

Therefore from (6.15)

$$\lim_{n\to\infty} \mu(S \cap (K_n \setminus K)) = 0.$$

From (6.14)

$$0 \leq \mu(S \setminus K) - \mu(S \setminus K_n) \leq \mu(S \cap (K_n \setminus K))$$

and so

$$\lim_{n\to\infty} \mu(S \setminus K_n) = \mu(S \setminus K).$$

From (6.13)

$$\mu(S) \geq \mu(S \cap K) + \mu(S \setminus K).$$

This shows $K \in \mathcal{S}$ and proves the theorem.

The following technical lemma will also prove useful in what follows.

Lemma 6.15 *Suppose ν is a measure defined on a σ algebra, $\mathcal{S}$ of sets of Ω, where (Ω, d) is a metric space having the property that $\Omega = \cup_{k=1}^{\infty} \Omega_k$ where Ω_k is a compact set and for all $k, \Omega_k \subseteq \Omega_{k+1}$. Suppose that $\mathcal{S}$ contains the Borel sets and ν is finite on compact sets. Suppose that ν also has the property that for every $E \in \mathcal{S}$,*

$$\nu(E) = \inf\{\nu(V) : V \supseteq E,\ V \text{ open}\}. \tag{6.17}$$

Then it follows that for all $E \in \mathcal{S}$

$$\nu(E) = \sup\{\nu(K) : K \subseteq E,\ K \text{ compact}\}. \tag{6.18}$$

Proof: Let $E \in \mathcal{S}$ and let $l < \nu(E)$. By Theorem 5.11 we may choose k large enough that

$$l < \nu(E \cap \Omega_k).$$

Now let $F \equiv \Omega_k \setminus E$. Thus $F \cup (E \cap \Omega_k) = \Omega_k$. By assumption, there is an open set, V containing F with

$$\nu(V) - \nu(F) = \nu(V \setminus F) < \nu(E \cap \Omega_k) - l.$$

We define the compact set, $K \equiv V^C \cap \Omega_k$. Then $K \subseteq E \cap \Omega_k$ and

$$E \cap \Omega_k \setminus K = E \cap \Omega_k \cap \left(V \cup \Omega_k^C\right)$$

$$= E \cap \Omega_k \cap V \subseteq \Omega_k \cap F^C \cap V \subseteq V \setminus F.$$

Therefore,

$$\nu(E \cap \Omega_k) - \nu(K) = \nu((E \cap \Omega_k) \setminus K)$$

$$\leq \nu(V \setminus F) < \nu(E \cap \Omega_k) - l$$

which implies

$$l < \nu(K).$$

This proves the lemma because $l < \nu(E)$ was arbitrary.

Definition 6.16 *We say a measure which satisfies (6.17) for all E measurable, is outer regular and a measure which satisfies (6.18) for all E measurable is inner regular. A measure which satisfies both is called regular.*

Thus Lemma 6.15 gives a condition under which outer regular implies inner regular.

With this preparation we are ready to prove the Riesz representation theorem for positive linear functionals.

Theorem 6.17 *Let (Ω, d) be a σ compact metric space with the property that the closures of balls are compact and let Λ be a positive linear functional on $C_c(\Omega)$. Then there exists a unique σ algebra and measure, μ, such that*

$$\mu \text{ is complete, Borel, and regular,} \tag{6.19}$$

$$\mu(K) < \infty \text{ for all } K \text{ compact,} \tag{6.20}$$

$$\Lambda f = \int f d\mu \text{ for all } f \in C_c(\Omega). \tag{6.21}$$

Such measures satisfying (6.19) and (6.20) are called Radon measures.

Proof: First we deal with the question of existence and then we will consider uniqueness. In all that follows V will denote an open set and K will denote a compact set. Define

$$\mu(V) \equiv \sup\{\Lambda(f) : f \prec V\}, \mu(\emptyset) \equiv 0, \tag{6.22}$$

and for an arbitrary set, T,

$$\mu(T) \equiv \inf\{\mu(V) : V \supseteq T\}.$$

We need to show first that this is well defined because there are two ways of defining $\mu(V)$.

Lemma 6.18 *μ is a well defined outer measure on $\mathcal{P}(\Omega)$.*

Proof: First we consider the question of whether μ is well defined. To clarify the argument, denote by μ_1 the first definition for open sets given in (6.22).

$$\mu(V) \equiv \inf\{\mu_1(U) : U \supseteq V\} \leq \mu_1(V).$$

But also, whenever $U \supseteq V$, $\mu_1(U) \geq \mu_1(V)$ and so

$$\mu(V) \geq \mu_1(V).$$

This proves that μ is well defined. Next we verify μ is an outer measure. It is clear that if $A \subseteq B$ then $\mu(A) \leq \mu(B)$. First we verify countable subadditivity for open sets. Thus let $V = \cup_{i=1}^{\infty} V_i$ and let $l < \mu(V)$. Then there exists $f \prec V$ such that $\Lambda f > l$. Now $spt(f)$ is a compact subset of V and so there exists m such that $\{V_i\}_{i=1}^{m}$ covers $spt(f)$. Then, letting ψ_i be a partition of unity from Theorem 6.11 with $spt(\psi_i) \subseteq V_i$, it follows that

$$l < \Lambda(f) = \sum_{i=1}^{n} \Lambda(\psi_i f) \leq \sum_{i=1}^{\infty} \mu(V_i).$$

Since $l < \mu(V)$ is arbitrary, it follows that

$$\mu(V) \leq \sum_{i=1}^{\infty} \mu(V_i).$$

Now we must verify that for any sets, A_i,

$$\mu(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

It suffices to consider the case that $\mu(A_i) < \infty$ for all i. Let $V_i \supseteq A_i$ and $\mu(A_i) + \frac{\varepsilon}{2^i} > \mu(V_i)$. Then from countable subadditivity on open sets,

$$\mu(\cup_{i=1}^{\infty} A_i) \leq \mu(\cup_{i=1}^{\infty} V_i)$$

$$\leq \sum_{i=1}^{\infty} \mu(V_i) \leq \sum_{i=1}^{\infty} \mu(A_i) + \frac{\varepsilon}{2^i} \leq \varepsilon + \sum_{i=1}^{\infty} \mu(A_i).$$

Since ε is arbitrary, this proves the lemma.

We will denote by $\mathcal{S}$ the σ algebra of μ measurable sets.

Lemma 6.19 *The outer measure, μ is finite on all compact sets and in fact, if $K \prec g$, then*

$$\mu(K) \leq \Lambda(g) \tag{6.23}$$

Also $\mathcal{S} \supseteq$ Borel sets so μ is a Borel measure.

Proof: Let $V_\alpha \equiv \{x \in \Omega : g(x) > \alpha\}$ where $\alpha \in (0,1)$ is arbitrary. Now let $h \prec V_\alpha$. Thus $h(x) \leq 1$ and equals zero off V_α while $\alpha^{-1} g(x) \geq 1$ on V_α. Therefore,

$$\Lambda\left(\alpha^{-1} g\right) \geq \Lambda(h).$$

Since $h \prec V_\alpha$ was arbitrary, this shows $\alpha^{-1}\Lambda(g) \geq \mu(V_\alpha) \geq \mu(K)$. Letting $\alpha \to 1$ yields the formula (6.23).

Next we verify that $\mathcal{S} \supseteq$ Borel sets. First suppose that V_1 and V_2 are disjoint open sets with $\mu(V_1 \cup V_2) < \infty$. Let $f_i \prec V_i$ be such that $\Lambda(f_i) + \varepsilon > \mu(V_i)$. Then

$$\mu(V_1 \cup V_2) \geq \Lambda(f_1 + f_2) = \Lambda(f_1) + \Lambda(f_2) \geq \mu(V_1) + \mu(V_2) - 2\varepsilon.$$

Since ε is arbitrary, this shows that $\mu(V_1 \cup V_2) = \mu(V_1) + \mu(V_2)$.

Now suppose that $dist(A, B) = r > 0$ and $\mu(A \cup B) < \infty$. Let

$$\widetilde{V}_1 \equiv \cup\left\{B\left(a, \frac{r}{2}\right) : a \in A\right\}, \widetilde{V}_2 \equiv \cup\left\{B\left(b, \frac{r}{2}\right) : b \in B\right\}.$$

Now let W be an open set containing $A \cup B$ such that $\mu(A \cup B) + \varepsilon > \mu(W)$. Now define $V_i \equiv W \cap \widetilde{V}_i$ and $V \equiv V_1 \cup V_2$. Then

$$\mu(A \cup B) + \varepsilon > \mu(W) \geq \mu(V) = \mu(V_1) + \mu(V_2) \geq \mu(A) + \mu(B).$$

Since ε is arbitrary, the conditions of Caratheodory's criterion are satisfied showing that $\mathcal{S} \supseteq$ Borel sets. This proves the lemma.

It is now easy to verify condition (6.19) and (6.20). Condition (6.20) and that μ is Borel is proved in Lemma 6.19. The measure space just described is complete because it comes from an outer measure using the Caratheodory procedure. The construction of μ shows outer regularity and the inner regularity follows from (6.20), shown in Lemma 6.19, and Lemma 6.15. It only remains to verify Condition (6.21), that the measure reproduces the functional in the desired manner and that the given measure and σ algebra is unique.

Lemma 6.20 *$\int f d\mu = \Lambda f$ for all $f \in C_c(\Omega)$.*

Proof: It suffices to verify this for $f \in C_c(\Omega)$, f real-valued. Suppose f is such a function and $f(\Omega) \subseteq [a,b]$. Choose $t_0 < a$ and let $t_0 < t_1 < \cdots < t_n = b$, $t_i - t_{i-1} < \varepsilon$. Let

$$E_i = f^{-1}((t_{i-1}, t_i]) \cap spt(f). \tag{6.24}$$

Note that $\cup_{i=1}^n E_i$ is a closed set and in fact

$$\cup_{i=1}^n E_i = spt(f) \tag{6.25}$$

since $\Omega = \cup_{i=1}^n f^{-1}((t_{i-1}, t_i])$. From outer regularity and continuity of f, let $V_i \supseteq E_i, V_i$ is open and let V_i satisfy

$$f(x) < t_i + \varepsilon \text{ for all } x \in V_i, \tag{6.26}$$

$$\mu(V_i \setminus E_i) < \varepsilon/n.$$

By Theorem 6.11 there exists $h_i \in C_c(\Omega)$ such that

$$h_i \prec V_i, \ \sum_{i=1}^n h_i(x) = 1 \text{ on } spt(f).$$

Now note that for each i,

$$f(x)h_i(x) \leq h_i(x)(t_i + \varepsilon).$$

(If $x \in V_i$, this follows from (6.26). If $x \notin V_i$ both sides equal 0.) Therefore,

$$\begin{aligned}
\Lambda f &= \Lambda(\sum_{i=1}^n f h_i) \leq \Lambda(\sum_{i=1}^n h_i(t_i + \varepsilon)) \\
&= \sum_{i=1}^n (t_i + \varepsilon)\Lambda(h_i) \\
&= \sum_{i=1}^n (|t_0| + t_i + \varepsilon)\Lambda(h_i) - |t_0|\Lambda\left(\sum_{i=1}^n h_i\right).
\end{aligned}$$

Now note that $|t_0| + t_i + \varepsilon \geq 0$ and so from the definition of μ and Lemma 6.19, this is no larger than

$$\sum_{i=1}^n (|t_0| + t_i + \varepsilon)\mu(V_i) - |t_0|\mu(spt(f))$$

$$\leq \sum_{i=1}^{n}(|t_0|+t_i+\varepsilon)(\mu(E_i)+\varepsilon/n)-|t_0|\mu(spt(f))$$

$$\leq |t_0|\sum_{i=1}^{n}\mu(E_i)+|t_0|\varepsilon+\sum_{i=1}^{n}t_i\mu(E_i)+\varepsilon(|t_0|+|b|)$$

$$+\varepsilon\sum_{i=1}^{n}\mu(E_i)+\varepsilon^2-|t_0|\mu(spt(f)).$$

From (6.25) and (6.24), the first and last terms cancel. Therefore this is no larger than

$$(2|t_0|+|b|+\mu(spt(f))+\varepsilon)\varepsilon+\sum_{i=1}^{n}t_{i-1}\mu(E_i)+\varepsilon\mu(spt(f))$$

$$\leq \int f d\mu+(2|t_0|+|b|+2\mu(spt(f))+\varepsilon)\varepsilon.$$

Since $\varepsilon>0$ is arbitrary,

$$\Lambda f\leq \int f d\mu \tag{6.27}$$

for all $f\in C_c(\Omega)$, f real. Hence equality holds in (6.27) because $\Lambda(-f)\leq -\int f d\mu$ so $\Lambda(f)\geq \int f d\mu$. Thus $\Lambda f=\int f d\mu$ for all $f\in C_c(\Omega)$. Just apply the result for real functions to the real and imaginary parts of f. This proves the Lemma.

Now that we have shown that μ satisfies the conditions of the Riesz representation theorem, we show that μ is the only measure that does so.

Lemma 6.21 *The measure and σ algebra of Theorem 6.17 are unique.*

Proof: If $(\mu_1,\mathcal{S}_1)$ and $(\mu_2,\mathcal{S}_2)$ both work, let

$$K\subseteq V,\ K\prec f\prec V.$$

Then

$$\mu_1(K)\leq \int f d\mu_1=\Lambda f=\int f d\mu_2\leq \mu_2(V).$$

Thus $\mu_1(K)\leq\mu_2(K)$ because of the outer regularity of μ_2. Similarly, $\mu_1(K)\geq\mu_2(K)$ and this shows that $\mu_1=\mu_2$ on all compact sets. It follows from inner regularity that the two measures coincide on all open sets as well. Now let $E\in\mathcal{S}_1$, the σ algebra associated with μ_1, and let $E_n=E\cap\Omega_n$. By the regularity of the measures, there exist sets G and H such that G is a countable intersection of decreasing open sets and H is a countable union of increasing compact sets which satisfy

$$G\supseteq E_n\supseteq H,\ \mu_1(G\setminus H)=0.$$

Since the two measures agree on all open and compact sets, it follows that $\mu_2(G)=\mu_1(G)$ and a similar equation holds for H in place of G. Therefore $\mu_2(G\setminus H)=\mu_1(G\setminus H)=0$. By completeness of $\mu_2, E_n\in\mathcal{S}_2$, the σ algebra associated with μ_2. Thus $E\in\mathcal{S}_2$ since $E=\cup_{n=1}^{\infty}E_n$, showing that $\mathcal{S}_1\subseteq\mathcal{S}_2$. Similarly $\mathcal{S}_2\subseteq\mathcal{S}_1$.

Since the two σ algebras are equal and the two measures are equal on every open set, regularity of these measures shows they coincide on all measurable sets and this proves the theorem.

The following theorem is an interesting application of the Riesz representation theorem for measures defined on subsets of $\mathbb{R}^n$.

Let M be a closed subset of $\mathbb{R}^n$. Then we may consider M as a metric space which has closures of balls compact if we let the topology on M consist of intersections of open sets from the standard topology of $\mathbb{R}^n$ with M or equivalently, use the usual metric on $\mathbb{R}^n$ restricted to M.

Proposition 6.22 *Let τ be the relative topology of M consisting of intersections of open sets of $\mathbb{R}^n$ with M and let $\mathcal{B}$ be the Borel sets of the topological space (M, τ). Then*

$$\mathcal{B} = \mathcal{S} \equiv \{E \cap M : E \text{ is a Borel set of } \mathbb{R}^n\}.$$

Proof: It is clear that $\mathcal{S}$ defined above is a σ algebra containing τ and so $\mathcal{S} \supseteq \mathcal{B}$. Now define

$$\mathcal{M} \equiv \{E \text{ Borel in } \mathbb{R}^n \text{ such that } E \cap M \in \mathcal{B}\}.$$

Then $\mathcal{M}$ is clearly a σ algebra which contains the open sets of $\mathbb{R}^n$. Therefore, $\mathcal{M} \supseteq$ Borel sets of $\mathbb{R}^n$ which shows $\mathcal{S} \subseteq \mathcal{B}$. This proves the proposition.

Theorem 6.23 *Suppose μ is a measure defined on the Borel sets of M where M is a closed subset of $\mathbb{R}^n$. Suppose also that μ is finite on compact sets. Then $\overline{\mu}$, the outer measure determined by μ, is a Radon measure on a σ algebra containing the Borel sets of (M, τ) where τ is the relative topology described above.*

Proof: Since μ is Borel and finite on compact sets, we may define a positive linear functional on $C_c(M)$ as

$$Lf \equiv \int_M f d\mu.$$

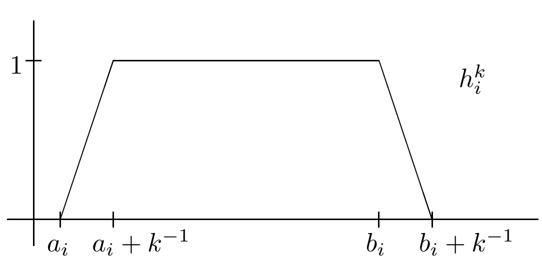
By the Riesz representation theorem, there exists a unique Radon measure and σ algebra, μ_1 and $\mathcal{S}(\mu_1)$ respectively, such that for all $f \in C_c(M)$,

$$\int_M f d\mu = \int_M f d\mu_1.$$

Let $\mathcal{R}$ and $\mathcal{E}$ be as described in Example 5.8 and let $Q_r = (-r, r]^n$. Then if $R \in \mathcal{R}$, it follows that $R \cap Q_r$ has the form,

$$R \cap Q_r = \prod_{i=1}^n (a_i, b_i].$$

Let $f_k(x_1 \cdots, x_n) = \prod_{i=1}^n h_i^k(x_i)$ where h_i^k is given by the following graph.



Then $\operatorname{spt}(f_k) \subseteq [-r, r+1]^n \equiv K$. Thus

$$\int_{K \cap M} f_k d\mu = \int_M f_k d\mu = \int_M f_k d\mu_1 = \int_{K \cap M} f_k d\mu_1.$$

Since $f_k \to \mathcal{X}_{R\cap Q_r}$ pointwise and both measures are finite, on $K \cap M$, it follows from the dominated convergence theorem that

$$\mu(R \cap Q_r \cap M) = \mu_1(R \cap Q_r \cap M)$$

and so it follows that this holds for R replaced with any $A \in \mathcal{E}$. Now define

$$\mathcal{M} \equiv \{E \text{ Borel} : \mu(E \cap Q_r \cap M) = \mu_1(E \cap Q_r \cap M)\}.$$

Then $\mathcal{E} \subseteq \mathcal{M}$ and it is clear that $\mathcal{M}$ is a monotone class. By the theorem on monotone classes, it follows that $\mathcal{M} \supseteq \sigma(\mathcal{E})$, the smallest σ algebra containing $\mathcal{E}$. This σ algebra contains the open sets because

$$\prod_{i=1}^{n}(a_i, b_i) = \cup_{k=1}^{\infty}\prod_{i=1}^{n}(a_i, b_i - k^{-1}] \in \sigma(\mathcal{E})$$

and every open set can be written as a countable union of sets of the form $\prod_{i=1}^{n}(a_i, b_i)$. Therefore,

$$\mathcal{M} \supseteq \sigma(\mathcal{E}) \supseteq \text{Borel sets} \supseteq \mathcal{M}.$$

Thus,

$$\mu(E \cap Q_r \cap M) = \mu_1(E \cap Q_r \cap M)$$

for all E Borel. Letting $r \to \infty$, it follows that $\mu(E \cap M) = \mu_1(E \cap M)$ for all E a Borel set of $\mathbb{R}^n$. By Proposition 6.22 $\mu(F) = \mu_1(F)$ for all F Borel in (M, τ). Consequently,

$$\begin{aligned}
\overline{\mu_1}(S) &\equiv \inf\{\mu_1(E) : E \supseteq S,\ E \in \mathcal{S}(\mu_1)\} \\
&= \inf\{\mu_1(F) : F \supseteq S,\ F \text{ Borel}\} \\
&= \inf\{\mu(F) : F \supseteq S,\ F \text{ Borel}\} \\
&= \overline{\mu}(S).
\end{aligned}$$

Therefore, by Lemma 6.6, the $\overline{\mu}$ measurable sets consist of $\mathcal{S}(\mu_1)$ and $\overline{\mu} = \mu_1$ on $\mathcal{S}(\mu_1)$ and this shows $\overline{\mu}$ is regular as claimed. This proves the theorem.

6.3 Exercises

1. Let $\Omega = \mathbb{N}$, the natural numbers and let $d(p, q) = |p - q|$, the usual distance in $\mathbb{R}$. Show that (Ω, d) is σ compact and the closures of the balls are compact. Now let $\Lambda f \equiv \sum_{k=1}^{\infty} f(k)$ whenever $f \in C_c(\Omega)$. Show this is a well defined positive linear functional on the space $C_c(\Omega)$. Describe the measure of the Riesz representation theorem which results from this positive linear functional. What if $\Lambda(f) = f(1)$? What measure would result from this functional?

2. Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right continuous. Let $\Lambda f \equiv \int f dF$ where the integral is the Riemann Stieltjes integral of f. Show the measure μ from the Riesz representation theorem satisfies

$$\begin{aligned}
\mu([a, b]) &= F(b) - F(a-), \mu((a, b]) = F(b) - F(a), \\
\mu([a, a]) &= F(a) - F(a-).
\end{aligned}$$

3. Let Ω be a σ compact metric space with the closed balls compact and suppose μ is a measure defined on the Borel sets of Ω which is finite on compact sets. Show there exists a unique Radon measure, $\overline{\mu}$ which equals μ on the Borel sets.

4. ↑ Random vectors are measurable functions, $\mathbf{X}$, mapping a probability space, $(\Omega, P, \mathcal{F})$ to $\mathbb{R}^n$. Thus $\mathbf{X}(\omega) \in \mathbb{R}^n$ for each $\omega \in \Omega$ and P is a probability measure defined on the sets of $\mathcal{F}$, a σ algebra of subsets of Ω. For E a Borel set in $\mathbb{R}^n$, define

$$\mu(E) \equiv P\left(\mathbf{X}^{-1}(E)\right) \equiv \text{ probability that } \mathbf{X} \in E.$$

Show this is a well defined measure on the Borel sets of $\mathbb{R}^n$ and use Problem 3 to obtain a Radon measure, $\lambda_{\mathbf{X}}$ defined on a σ algebra of sets of $\mathbb{R}^n$ including the Borel sets such that for E a Borel set, $\lambda_{\mathbf{X}}(E) =$Probability that $(\mathbf{X} \in E)$.

5. Let (X, d_X) and (Y, d_Y) be metric spaces and make $X \times Y$ into a metric space in the following way.

$$d_{X \times Y}((x, y), (x_1, y_1)) \equiv \max(d_X(x, x_1), d_Y(y, y_1)).$$

Show this is a metric space.

6. ↑ Show $(X \times Y, d_{X \times Y})$ is also a σ compact metric space having closed balls compact if both X and Y are σ compact metric spaces having the closed balls compact. Let

$$\mathcal{A} \equiv \{E \times F : E \text{ is a Borel set in } X, F \text{ is a Borel set in } Y\}.$$

Show $\sigma(\mathcal{A})$, the smallest σ algebra containing $\mathcal{A}$ contains the Borel sets. **Hint:** Show every open set in a σ compact metric space can be obtained as a countable union of compact sets. Next show this implies every open set can be obtained as a countable union of open sets of the form $U \times V$ where U is open in X and V is open in Y.

7. ↑ Let μ and ν be Radon measures on X and Y respectively. Define for

$$f \in C_c(X \times Y),$$

the linear functional Λ given by the iterated integral,

$$\Lambda f \equiv \int_X \int_Y f(x, y)\, d\nu d\mu.$$

Show this is well defined and yields a positive linear functional on $C_c(X \times Y)$. Let $\overline{\mu \times \nu}$ be the Radon measure representing Λ. Show for $f \geq 0$ and Borel measurable, that

$$\int_Y \int_X f(x, y)\, d\mu d\nu = \int_X \int_Y f(x, y)\, d\nu d\mu = \int_{X \times Y} f d\left(\overline{\mu \times \nu}\right)$$

Lebesgue Measure

7.1 Lebesgue measure

In this chapter, n dimensional Lebesgue measure and many of its properties are obtained from the Riesz representation theorem. This is done by using a positive linear functional familiar to anyone who has had a course in calculus. The positive linear functional is

$$\Lambda f \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n) dx_1 \cdots dx_n \tag{7.1}$$

for $f \in C_c(\mathbb{R}^n)$. This is the ordinary Riemann iterated integral and we need to observe that it makes sense.

Lemma 7.1 *Let $f \in C_c(\mathbb{R}^n)$ for $n \geq 2$. Then*

$$h(x_n) \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1 \cdots x_{n-1} x_n) dx_1 \cdots dx_{n-1}$$

is well defined and $h \in C_c(\mathbb{R})$.

Proof: Assume this is true for all $2 \leq k \leq n-1$. Then fixing x_n,

$$x_{n-1} \rightarrow \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1 \cdots x_{n-2}, x_{n-1}, x_n) dx_1 \cdots dx_{n-2}$$

is a function in $C_c(\mathbb{R})$. Therefore, it makes sense to write

$$h(x_n) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1 \cdots x_{n-2}, x_{n-1}, x_n) dx_1 \cdots dx_{n-1}.$$

We need to verify $h \in C_c(\mathbb{R})$. Since f vanishes whenever $|\mathbf{x}|$ is large enough, it follows $h(x_n) = 0$ whenever $|x_n|$ is large enough. It only remains to show h is continuous. But f is uniformly continuous, so if $\varepsilon > 0$ is given there exists a δ such that

$$|f(\mathbf{x}_1) - f(\mathbf{x})| < \varepsilon$$

whenever $|\mathbf{x}_1 - \mathbf{x}| < \delta$. Thus, letting $|x_n - \bar{x}_n| < \delta$,

$$|h(x_n) - h(\bar{x}_n)| \leq$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |f(x_1 \cdots x_{n-1}, x_n) - f(x_1 \cdots x_{n-1}, \bar{x}_n)| dx_1 \cdots dx_{n-1}$$

$$\leq \varepsilon(b-a)^{n-1}$$

where $\text{spt}(f) \subseteq [a,b]^n \equiv [a,b] \times \cdots \times [a,b]$. This argument also shows the lemma is true for $n = 2$. This proves the lemma.

From Lemma 7.1 it is clear that (7.1) makes sense and also that Λ is a positive linear functional for $n = 1, 2, \cdots$.

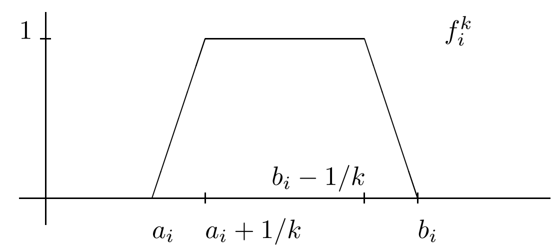
Definition 7.2 *m_n is the unique Radon measure representing Λ. Thus for all $f \in C_c(\mathbb{R}^n)$,*

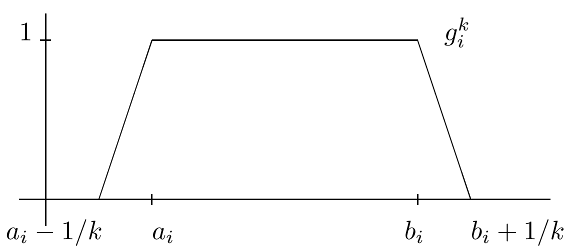
$$\Lambda f = \int f dm_n.$$

Let $R = \prod_{i=1}^n [a_i, b_i]$, $R_0 = \prod_{i=1}^n (a_i, b_i)$. What are $m_n(R)$ and $m_n(R_0)$? We show that both of these equal $\prod_{i=1}^n (b_i - a_i)$. To see this is the case, let k be large enough that

$$a_i + 1/k < b_i - 1/k$$

for $i = 1, \cdots, n$. Consider functions g_i^k and f_i^k having the following graphs.





Let

$$g^k(\mathbf{x}) = \prod_{i=1}^n g_i^k(x_i), \quad f^k(\mathbf{x}) = \prod_{i=1}^n f_i^k(x_i).$$

Then

$$\prod_{i=1}^n (b_i - a_i + 2/k) \geq \Lambda g^k = \int g^k dm_n \geq m_n(R) \geq m_n(R_0)$$

$$\geq \int f^k dm_n = \Lambda f^k \geq \prod_{i=1}^n (b_i - a_i - 2/k).$$

Letting $k \to \infty$, it follows that

$$m_n(R) = m_n(R_0) = \prod_{i=1}^{n}(b_i - a_i)$$

as expected.

We say R is a half open box if

$$R = \prod_{i=1}^{n}[a_i, a_i + r).$$

Lemma 7.3 *Every open set in $\mathbb{R}^n$ is the countable disjoint union of half open boxes of the form*

$$\prod_{i=1}^{n}(a_i, a_i + 2^{-k}]$$

where $a_i = l2^{-k}$ for some integers, l, k.

Proof: Let

$$\mathcal{C}_k = \{\text{All half open boxes } \prod_{i=1}^{n}(a_i, a_i + 2^{-k}] \text{ where}$$

$$a_i = l2^{-k} \text{ for some integer } l.\}$$

Thus $\mathcal{C}_k$ consists of a countable disjoint collection of boxes whose union is $\mathbb{R}^n$. This is sometimes called a tiling of $\mathbb{R}^n$. Note that each box has diameter $2^{-k}\sqrt{n}$. Let U be open and let $\mathcal{B}_1 \equiv$ all sets of $\mathcal{C}_1$ which are contained in U. If $\mathcal{B}_1, \cdots, \mathcal{B}_k$ have been chosen, $\mathcal{B}_{k+1} \equiv$ all sets of $\mathcal{C}_{k+1}$ contained in

$$U \setminus \cup(\cup_{i=1}^{k}\mathcal{B}_i).$$

Let $\mathcal{B}_\infty = \cup_{i=1}^{\infty}\mathcal{B}_i$. We claim $\cup\mathcal{B}_\infty = U$. Clearly $\cup\mathcal{B}_\infty \subseteq U$. If $p \in U$, let k be the smallest integer such that p is contained in a box from $\mathcal{C}_k$ which is also a subset of U. Thus

$$p \in \cup\mathcal{B}_k \subseteq \cup\mathcal{B}_\infty.$$

Hence $\mathcal{B}_\infty$ is the desired countable disjoint collection of half open boxes whose union is U. This proves the lemma.

Lebesgue measure is translation invariant. This means roughly that if you take a Lebesgue measurable set, and slide it around, it remains Lebesgue measurable and the measure does not change.

Theorem 7.4 *Lebesgue measure is translation invariant, i.e.,*

$$m_n(\mathbf{v}{+}E) = m_n(E),$$

for E Lebesgue measurable.

Proof: First note that if E is Borel, then so is $\mathbf{v} + E$. To show this, let

$$\mathcal{S} \ = \{E \in \text{Borel sets such that } \mathbf{v} + E \text{ is Borel}\}.$$

Then from Lemma 7.3, $\mathcal{S}$ contains the open sets and is easily seen to be a σ algebra, so $\mathcal{S} =$ Borel sets. Now let E be a Borel set. Choose V open such that

$$m_n(V) < m_n(E \cap B(0,k)) + \varepsilon, \ \ V \supseteq E \cap B(0,k).$$

Then

$$m_n(\mathbf{v} + E \cap B(0, k)) \le m_n(\mathbf{v} + V) = m_n(V)$$

$$\le m_n(E \cap B(0, k)) + \varepsilon.$$

The equal sign is valid because the conclusion of Theorem 7.4 is clearly true for all open sets thanks to Lemma 7.3 and the simple observation that the theorem is true for boxes. Since ε is arbitrary,

$$m_n(\mathbf{v} + E \cap B(0, k)) \le m_n(E \cap B(0, k)).$$

Letting $k \to \infty$,

$$m_n(\mathbf{v} + E) \le m_n(E).$$

Since $\mathbf{v}$ is arbitrary,

$$m_n(-\mathbf{v} + (\mathbf{v} + E)) \le m_n(E + \mathbf{v}).$$

Hence

$$m_n(\mathbf{v} + E) \le m_n(E) \le m_n(\mathbf{v} + E)$$

proving the theorem in the case where E is Borel. Now suppose that $m_n(S) = 0$. Then there exists $E \supseteq S$, E Borel, and $m_n(E) = 0$.

$$m_n(E{+}\mathbf{v}) = m_n(E) = 0$$

Now $S + \mathbf{v} \subseteq E + \mathbf{v}$ and so by completeness of the measure, $S + \mathbf{v}$ is Lebesgue measurable and has measure zero. Thus,

$$m_n(S) = m_n(S + \mathbf{v}).$$

Now let F be an arbitrary Lebesgue measurable set and let $F_r = F \cap B(0, r)$. Then there exists a Borel set E, $E \supseteq F_r$, and $m_n(E \setminus F_r) = 0$. Then since

$$(E + \mathbf{v}) \setminus (F_r + \mathbf{v}) = (E \setminus F_r) + \mathbf{v},$$

it follows

$$m_n((E + \mathbf{v}) \setminus (F_r + \mathbf{v})) = m_n((E \setminus F_r) + \mathbf{v}) = m_n(E \setminus F_r) = 0.$$

By completeness of m_n, $F_r + \mathbf{v}$ is Lebesgue measurable and $m_n(F_r + \mathbf{v}) = m_n(E + \mathbf{v})$. Hence

$$m_n(F_r) = m_n(E) = m_n(E + \mathbf{v}) = m_n(F_r + \mathbf{v}).$$

Letting $r \to \infty$, we obtain

$$m_n(F) = m_n(F + \mathbf{v})$$

and this proves the theorem.

7.2 Iterated integrals

The positive linear functional used to define Lebesgue measure was an iterated integral. Of course one could take the iterated integral in another order. What would happen to the resulting Radon measure if another order was used? This question will be considered in this section. First, here is a simple lemma.

Lemma 7.5 *If μ and ν are two Radon measures defined on σ algebras, $\mathcal{S}_\mu$ and $\mathcal{S}_\nu$, of subsets of $\mathbb{R}^n$ and if $\mu(V) = \nu(V)$ for all V open, then $\mu = \nu$ and $\mathcal{S}_\mu = \mathcal{S}_\nu$.*

Proof: Let $\overline{\mu}$ and $\overline{\nu}$ be the outer measures determined by μ and ν. Then if E is a Borel set,

$$\begin{aligned} \mu(E) &= \inf\{\mu(V) : V \supseteq E \text{ and } V \text{ is open}\} \\ &= \inf\{\nu(V) : V \supseteq E \text{ and } V \text{ is open}\} = \nu(E). \end{aligned}$$

Now if S is any subset of $\mathbb{R}^n$,

$$\overline{\mu}(S) \equiv \inf\{\mu(E) : E \supseteq S, E \in \mathcal{S}_\mu\}$$

$$= \inf\{\mu(F) : F \supseteq S, F \text{ is Borel}\}$$

$$= \inf\{\nu(F) : F \supseteq S, F \text{ is Borel}\}$$

$$= \inf\{\nu(E) : E \supseteq S, E \in \mathcal{S}_\nu\} \equiv \overline{\nu}(S)$$

where the second and fourth equalities follow from the outer regularity which is assumed to hold for both measures. Therefore, the two outer measures are identical and so the measurable sets determined in the sense of Caratheodory are also identical. By Lemma 6.6 of Chapter 6 this implies both of the given σ algebras are equal and $\mu = \nu$.

Lemma 7.6 *If μ and ν are two Radon measures on $\mathbb{R}^n$ and $\mu = \nu$ on every half open box, then $\mu = \nu$.*

Proof: From Lemma 7.3, $\mu(U) = \nu(U)$ for all U open. Therefore, by Lemma 7.5, the two measures coincide with their respective σ algebras.

Corollary 7.7 *Let $\{1, 2, \cdots, n\} = \{k_1, k_2, \cdots, k_n\}$ and the k_i are distinct. For $f \in C_c(\mathbb{R}^n)$ let*

$$\tilde{\Lambda} f = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n) dx_{k_1} \cdots dx_{k_n},$$

an iterated integral in a different order. Then for all $f \in C_c(\mathbb{R}^n)$,

$$\Lambda f = \tilde{\Lambda} f,$$

and if $\widetilde{m}_n$ is the Radon measure representing $\widetilde{\Lambda}$ and m_n is the Radon measure representing m_n, then $m_n = \widetilde{m}_n$.

Proof: Let $\tilde{m}_n$ be the Radon measure representing $\tilde{\Lambda}$. Then clearly $m_n = \tilde{m}_n$ on every half open box. By Lemma 7.6, $m_n = \tilde{m}_n$. Thus,

$$\Lambda f = \int f dm_n = \int f d\tilde{m}_n = \tilde{\Lambda} f.$$

This Corollary 7.7 is pretty close to Fubini's theorem for Riemann integrable functions. Now we will generalize it considerably by allowing f to only be Borel measurable. To begin with, we consider the question of existence of the iterated integrals for $\mathcal{X}_E$ where E is a Borel set.

Lemma 7.8 *Let E be a Borel set and let $\{k_1, k_2, \cdots, k_n\}$ distinct integers from $\{1, \cdots, n\}$ and if $Q_k \equiv \prod_{i=1}^{n}(-p, p]$. Then for each $r = 2, \cdots, n-1$, the function,*

$$x_{k_{r+1}} \to \overbrace{\int \cdots \int}^{r \text{ integrals}} \mathcal{X}_{E \cap Q_p}(x_1, \cdots, x_n)\, dm(x_{k_1}) \cdots dm(x_{k_r}) \tag{7.2}$$

is Lebesgue measurable. Thus we can add another iterated integral and write

$$\int \overbrace{\int \cdots \int}^{r \text{ integrals}} \mathcal{X}_{E \cap Q_p}(x_1, \cdots, x_n)\, dm(x_{k_1}) \cdots dm(x_{k_r})\, dm\left(x_{k_{r+1}}\right).$$

Here the notation $dm(x_{k_i})$ means we integrate the function of x_{k_i} with respect to one dimensional Lebesgue measure.

Proof: If E is an element of $\mathcal{E}$, the algebra of Example 5.8, we leave the conclusion of this lemma to the reader. If $\mathcal{M}$ is the collection of Borel sets such that (7.2) holds, then the dominated convergence and monotone convergence theorems show that $\mathcal{M}$ is a monotone class. Therefore, $\mathcal{M}$ equals the Borel sets by the monotone class theorem. This proves the lemma.

The following lemma is just a generalization of this one.

Lemma 7.9 *Let f be any nonnegative Borel measurable function. Then for each $r = 2, \cdots, n-1$, the function,*

$$x_{k_{r+1}} \to \overbrace{\int \cdots \int}^{r \text{ integrals}} f(x_1, \cdots, x_n)\, dm(x_{k_1}) \cdots dm(x_{k_r}) \tag{7.3}$$

is one dimensional Lebesgue measurable.

Proof: Letting $p \to \infty$ in the conclusion of Lemma 7.8 we see the conclusion of this lemma holds without the intersection with Q_p. Thus, if s is a nonnegative Borel measurable function (7.3) holds with f replaced with s. Now let s_n be an increasing sequence of nonnegative Borel measurable functions which converge pointwise to f. Then by the monotone convergence theorem applied r times,

$$\begin{aligned} &\lim_{n\to\infty} \overbrace{\int \cdots \int}^{r \text{ integrals}} s_n(x_1, \cdots, x_n)\, dm(x_{k_1}) \cdots dm(x_{k_r}) \\ = &\overbrace{\int \cdots \int}^{r \text{ integrals}} f(x_1, \cdots, x_n)\, dm(x_{k_1}) \cdots dm(x_{k_r}) \end{aligned}$$

and so we may draw the desired conclusion because the given function of $x_{k_{r+1}}$ is a limit of a sequence of measurable functions of this variable.

To summarize this discussion, we have shown that if f is a nonnegative Borel measurable function, we may take the iterated integrals in any order and everything which needs to be measurable in order for the expression to make sense, is. The next lemma shows that different orders of integration in the iterated integrals yield the same answer.

Lemma 7.10 *Let E be any Borel set. Then*

$$\int_{\mathbb{R}^n} \mathcal{X}_E(\mathbf{x})\, dm_n = \int \cdots \int \mathcal{X}_E(x_1, \cdots, x_n)\, dm(x_1) \cdots dm(x_n)$$

$$= \int \cdots \int \mathcal{X}_E (x_1, \cdots, x_n) \, dm (x_{k_1}) \cdots dm (x_{k_n}) \tag{7.4}$$

where $\{k_1, \cdots, k_n\} = \{1, \cdots, n\}$ and everything which needs to be measurable is. Here the notation involving the iterated integrals refers to one-dimensional Lebesgue integrals.

Proof: Let $Q_k = (-k, k]^n$ and let

$$\mathcal{M} \equiv \{\text{Borel sets, } E, \text{ such that (7.4) holds for } E \cap Q_k$$

$$\text{and there are no measurability problems}\}.$$

If $\mathcal{E}$ is the algebra of Example 5.8, we see easily that for all such sets, A of $\mathcal{E}$, (7.4) holds for $A \cap Q_k$ and so $\mathcal{M} \supseteq \mathcal{E}$. Now the theorem on monotone classes implies $\mathcal{M} \supseteq \sigma(\mathcal{E})$ which, by Lemma 7.3, equals the Borel sets. Therefore, $\mathcal{M} =$ Borel sets. Letting $k \to \infty$, and using the monotone convergence theorem, yields the conclusion of the lemma.

The next theorem is referred to as Fubini's theorem. Although a more abstract version will be presented later, the version in the next theorem is particularly useful when dealing with Lebesgue measure.

Theorem 7.11 *Let $f \geq 0$ be Borel measurable. Then everything which needs to be measurable in order to write the following formulae is measurable, and*

$$\int_{\mathbb{R}^n} f(\mathbf{x}) \, dm_n = \int \cdots \int f(x_1, \cdots, x_n) \, dm(x_1) \cdots dm(x_n)$$

$$= \int \cdots \int f(x_1, \cdots, x_n) \, dm(x_{k_1}) \cdots dm(x_{k_n})$$

where $\{k_1, \cdots, k_n\} = \{1, \cdots, n\}$.

Proof: This follows from the previous lemma since the conclusion of this lemma holds for nonnegative simple functions in place of $\mathcal{X}_E$ and we may obtain f as the pointwise limit of an increasing sequence of nonnegative simple functions. The conclusion follows from the monotone convergence theorem.

Corollary 7.12 *Suppose f is complex valued and for some*

$$\{k_1, \cdots, k_n\} = \{1, \cdots, n\},$$

it follows that

$$\int \cdots \int |f(x_1, \cdots, x_n)| \, dm(x_{k_1}) \cdots dm(x_{k_n}) < \infty. \tag{7.5}$$

Then $f \in L^1(\mathbb{R}^n, m_n)$ and if $\{l_1, \cdots, l_n\} = \{1, \cdots, n\}$, then

$$\int_{\mathbb{R}^n} f dm_n = \int \cdots \int f(x_1, \cdots, x_n) \, dm(x_{l_1}) \cdots dm(x_{l_n}). \tag{7.6}$$

Proof: Applying Theorem 7.11 to the positive and negative parts of the real and imaginary parts of f, (7.5) implies all these integrals are finite and all iterated integrals taken in any order are equal for these functions. Therefore, the definition of the integral implies (7.6) holds.

7.3 Change of variables

In this section we show that if $F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, then $m_n(F(E)) = \Delta F m_n(E)$ whenever E is Lebesgue measurable. The constant ΔF will also be shown to be $|\det(F)|$. In order to prove this theorem, we recall Theorem 2.29 which is listed here for convenience.

Theorem 7.13 *Let $F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ where $m \geq n$. Then there exists $R \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and $U \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ such that $U = U^*$, all eigenvalues of U are nonnegative,*

$$U^2 = F^*F, \; R^*R = I, \; F = RU,$$

and $|R\mathbf{x}| = |\mathbf{x}|$.

The following corollary follows as a simple consequence of this theorem.

Corollary 7.14 *Let $F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and suppose $n \geq m$. Then there exists a symmetric nonnegative element of $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$, U, and an element of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, R, such that*

$$F = UR, \; RR^* = I.$$

Proof: We recall that if $M, L \in \mathcal{L}(\mathbb{R}^s, \mathbb{R}^p)$, then $L^{**} = L$ and $(ML)^* = L^*M^*$. Now apply Theorem 2.29 to $F^* \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. Thus,

$$F^* = R^*U$$

where R^* and U satisfy the conditions of that theorem. Then

$$F = UR$$

and $RR^* = R^{**}R^* = I$. This proves the corollary.

The next few lemmas involve the consideration of $F(E)$ where E is a measurable set. They show that $F(E)$ is Lebesgue measurable. We will have occasion to establish similar theorems in other contexts later in the book. In each case, the overall approach will be to show the mapping in question takes sets of measure zero to sets of measure zero and then to exploit the continuity of the mapping and the regularity and completeness of some measure to obtain the final result. The next lemma gives the first part of this procedure here. First we give a simple definition.

Definition 7.15 *Let $F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. We define $||F|| \equiv \max\{|F(x)| : |\mathbf{x}| \leq 1\}$. This number exists because the closed unit ball is compact.*

Now we note that from this definition, if $\mathbf{v}$ is any nonzero vector, then

$$|F(\mathbf{v})| = \left|F\left(\frac{\mathbf{v}}{|\mathbf{v}|}|\mathbf{v}|\right)\right| = |\mathbf{v}| \, F\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right) \leq ||F|| \, |\mathbf{v}| \, .$$

Lemma 7.16 *Let $F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, and suppose E is a measurable set having finite measure. Then*

$$\overline{m}_n(F(E)) \leq \left(2 \, ||F|| \sqrt{n}\right)^n m_n(E)$$

where $\overline{m}_n(\cdot)$ refers to the outer measure,

$$\overline{m}_n(S) \equiv \inf\{m_n(E) : E \supseteq S \text{ and } E \text{ is measurable}\} .$$

Proof: Let $\epsilon > 0$ be given and let $E \subseteq V$, an open set with $m_n(V) < m_n(E) + \epsilon$. Then let $V = \cup_{i=1}^{\infty} Q_i$ where Q_i is a half open box all of whose sides have length 2^{-l} for some $l \in \mathbb{N}$ and $Q_i \cap Q_j = \emptyset$ if $i \neq j$. Then $\operatorname{diam}(Q_i) = \sqrt{n} a_i$ where a_i is the length of the sides of Q_i. Thus, if y_i is the center of Q_i, then

$$B(Fy_i, ||F|| \operatorname{diam}(Q_i)) \supseteq FQ_i.$$

Let Q_i^* denote the cube with sides of length $2\,||F|| \operatorname{diam}(Q_i)$ and center at Fy_i. Then

$$Q_i^* \supseteq B(Fy_i, ||F|| \operatorname{diam}(Q_i)) \supseteq FQ_i$$

and so

$$\overline{m}_n(F(E)) \leq m_n(F(V)) \leq \sum_{i=1}^{\infty} (2\,||F|| \operatorname{diam}(Q_i))^n$$

$$\leq \left(2\,||F|| \sqrt{n}\right)^n \sum_{i=1}^{\infty} (m_n(Q_i)) = \left(2\,||F|| \sqrt{n}\right)^n m_n(V)$$

$$\leq \left(2\,||F|| \sqrt{n}\right)^n (m_n(E) + \epsilon).$$

Since $\epsilon > 0$ is arbitrary, this proves the lemma.

Lemma 7.17 *If E is Lebesgue measurable, then $F(E)$ is also Lebesgue measurable.*

Proof: First note that if K is compact, $F(K)$ is also compact and is therefore a Borel set. Also, if V is an open set, it can be written as a countable union of compact sets, $\{K_i\}$, and so

$$F(V) = \cup_{i=1}^{\infty} F(K_i)$$

which shows that $F(V)$ is a Borel set also. Now take any Lebesgue measurable set, E, which is bounded and use regularity of Lebesgue measure to obtain open sets, V_i, and compact sets, K_i, such that

$$K_i \subseteq E \subseteq V_i,$$

$V_i \supseteq V_{i+1}$, $K_i \subseteq K_{i+1}$, $K_i \subseteq E \subseteq V_i$, and $m_n(V_i \setminus K_i) < 2^{-i}$. Let

$$Q \equiv \cap_{i=1}^{\infty} F(V_i),\ P \equiv \cup_{i=1}^{\infty} F(K_i).$$

Thus both Q and P are Borel sets (hence Lebesgue) measurable. Observe that

$$Q \setminus P \subseteq \cap_{i,j} (F(V_i) \setminus F(K_j))$$

$$\subseteq \cap_{i=1}^{\infty} F(V_i) \setminus F(K_i) \subseteq \cap_{i=1}^{\infty} F(V_i \setminus K_i)$$

which means, by Lemma 7.16,

$$\overline{m}_n(Q \setminus P) \leq \overline{m}_n(F(V_i \setminus K_i)) \leq \left(2\,||F|| \sqrt{n}\right)^n 2^{-i}$$

which implies $Q \setminus P$ is a set of Lebesgue measure zero since i is arbitrary. Also,

$$P \subseteq F(E) \subseteq Q.$$

By completeness of Lebesgue measure, this shows $F(E)$ is Lebesgue measurable.

If E is not bounded but is measurable, consider $E \cap B(0,k)$. Then

$$F(E) = \cup_{k=1}^{\infty} F(E \cap B(0,k))$$

and, thus, $F(E)$ is measurable. This proves the lemma.

Lemma 7.18 *Let $Q_0 \equiv [0,1)^n$ and let $\Delta F \equiv m_n(F(Q_0))$. Then if Q is any half open box whose sides are of length 2^{-k}, $k \in \mathbb{N}$, and F is one to one, it follows*

$$m_n(FQ) = m_n(Q)\,\Delta F$$

and if F is not one to one we can say

$$m_n(FQ) \geq m_n(Q)\,\Delta F.$$

Proof: There are $\left(2^k\right)^n \equiv l$ nonintersecting half open boxes, Q_i, each having measure $\left(2^{-k}\right)^n$ whose union equals Q_0. If F is one to one, translation invariance of Lebesgue measure and the assumption F is linear imply

$$\left(2^k\right)^n m_n(F(Q)) = \sum_{i=1}^{l} m_n(F(Q_i)) =$$

$$m_n\left(\cup_{i=1}^{l} F(Q_i)\right) = m_n(F(Q_0)) = \Delta F. \tag{$*$}$$

Therefore,

$$m_n(F(Q)) = \left(2^{-k}\right)^n \Delta F = m_n(Q)\,\Delta F.$$

If F is not one to one, the sets $F(Q_i)$ are not necessarily disjoint and so the second equality sign in $(*)$ should be $\geq$. This proves the lemma.

Theorem 7.19 *If E is any Lebesgue measurable set, then*

$$m_n(F(E)) = \Delta F m_n(E).$$

If $R^*R = I$ and R preserves distances, then

$$\Delta R = 1.$$

Also, if $F, G \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, then

$$\Delta(FG) = \Delta F \Delta G. \tag{7.7}$$

Proof: Let V be any open set and let $\{Q_i\}$ be half open disjoint boxes of the sort discussed earlier whose union is V and suppose first that F is one to one. Then

$$m_n(F(V)) = \sum_{i=1}^{\infty} m_n(F(Q_i)) = \Delta F \sum_{i=1}^{\infty} m_n(Q_i) = \Delta F m_n(V). \tag{7.8}$$

Now let E be an arbitrary bounded measurable set and let V_i be a decreasing sequence of open sets containing E with

$$i^{-1} \geq m_n(V_i \setminus E).$$

Then let $S \equiv \cap_{i=1}^{\infty} F(V_i)$

$$S \setminus F(E) = \cap_{i=1}^{\infty} (F(V_i) \setminus F(E)) \subseteq \cap_{i=1}^{\infty} F(V_i \setminus E)$$

and so from Lemma 7.16,

$$m_n(S \setminus F(E)) \leq \limsup_{i\to\infty} m_n(F(V_i \setminus E))$$

$$\leq \limsup_{i\to\infty} \left(2\,||F||\,\sqrt{n}\right)^n i^{-1} = 0.$$

Thus

$$m_n\left(F\left(E\right)\right) = m_n\left(S\right) = \lim_{i\to\infty} m_n\left(F\left(V_i\right)\right)$$

$$= \lim_{i\to\infty} \Delta F m_n\left(V_i\right) = \Delta F m_n\left(E\right). \tag{7.9}$$

If E is not bounded, apply the above to $E \cap B(0,k)$ and let $k \to \infty$.

To see the second claim of the theorem,

$$\Delta R m_n\left(B\left(0,1\right)\right) = m_n\left(RB\left(0,1\right)\right) = m_n\left(B\left(0,1\right)\right).$$

Now suppose F is not one to one. Then let $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ such that for some $r < n, \{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ is an orthonormal basis for $F(\mathbb{R}^n)$. Let $R\mathbf{v}_i \equiv \mathbf{e}_i$ where the $\mathbf{e}_i$ are the standard unit basis vectors. Then $RF(\mathbb{R}^n) \subseteq span\left(\mathbf{e}_1, \cdots, \mathbf{e}_r\right)$ and so by what we know about the Lebesgue measure of boxes, whose sides are parallel to the coordinate axes,

$$m_n\left(RF\left(Q\right)\right) = 0$$

whenever Q is a box. Thus,

$$m_n\left(RF\left(Q\right)\right) = \Delta R m_n\left(F\left(Q\right)\right) = 0$$

and this shows that in the case where F is not one to one, $m_n(F(Q)) = 0$ which shows from Lemma 7.18 that $m_n(F(Q)) = \Delta F m_n(Q)$ even if F is not one to one. Therefore, (7.8) continues to hold even if F is not one to one and this implies (7.9). (7.7) follows from this. This proves the theorem.

Lemma 7.20 *Suppose $U = U^*$ and U has all nonnegative eigenvalues, $\{\lambda_i\}$. Then*

$$\Delta U = \prod_{i=1}^n \lambda_i.$$

Proof: Suppose $U_0 \equiv \sum_{i=1}^n \lambda_i \mathbf{e}_i \otimes \mathbf{e}_i$. Note that

$$Q_0 = \left\{ \sum_{i=1}^n t_i \mathbf{e}_i : t_i \in [0,1) \right\}.$$

Thus

$$U_0\left(Q_0\right) = \left\{ \sum_{i=1}^n \lambda_i t_i \mathbf{e}_i : t_i \in [0,1) \right\}$$

and so

$$\Delta U_0 \equiv m_n\left(U_0 Q_0\right) = \prod_{i=1}^n \lambda_i.$$

Now by linear algebra, since $U = U^*$,

$$U = \sum_{i=1}^n \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i$$

where $\{\mathbf{v}_i\}$ is an orthonormal basis of eigenvectors. Define $R \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ such that

$$R\mathbf{v}_i = \mathbf{e}_i.$$

Then R preserves distances and $RUR^* = U_0$ where U_0 is given above. Therefore, if E is any measurable set,

$$\prod_{i=1}^{n} \lambda_i m_n(E) = \Delta U_0 m_n(E) = m_n(U_0(E)) = m_n(RUR^*(E))$$

$$= \Delta R \Delta U \Delta R^* m_n(E) = \Delta U m_n(E).$$

Hence $\prod_{i=1}^{n} \lambda_i = \Delta U$ as claimed. This proves the theorem.

Theorem 7.21 *Let $F \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Then $\Delta F = |\det(F)|$. Thus*

$$m_n(F(E)) = |\det(F)|\, m_n(E)$$

for all E Lebesgue measurable.

Proof: By Theorem 2.29, $F = RU$ where R and U are described in that theorem. Then

$$\Delta F = \Delta R \Delta U = \Delta U = \det(U).$$

Now $F^*F = U^2$ and so $(\det(U))^2 = \det\left(U^2\right) = \det(F^*F) = (\det(F))^2$. Therefore,

$$\det(U) = |\det F|$$

and this proves the theorem.

7.4 Polar coordinates

One of the most useful of all techniques in establishing estimates which involve integrals taken with respect to Lebesgue measure on $\mathbb{R}^n$ is the technique of polar coordinates. This section presents the polar coordinate formula. To begin with we give a general lemma.

Lemma 7.22 *Let X and Y be topological spaces. Then if E is a Borel set in X and F is a Borel set in Y, then $E \times F$ is a Borel set in $X \times Y$.*

Proof: Let E be an open set in X and let

$$\mathcal{S}_E \equiv \{F \text{ Borel in } Y \text{ such that } E \times F \text{ is Borel in } X \times Y\}.$$

Then $\mathcal{S}_E$ contains the open sets and is clearly closed with respect to countable unions. Let $F \in \mathcal{S}_E$. Then

$$E \times F^C \cup E \times F = E \times Y = \text{ a Borel set.}$$

Therefore, since $E \times F$ is Borel, it follows $E \times F^C$ is Borel. Therefore, $\mathcal{S}_E$ is a σ algebra. It follows $\mathcal{S}_E =$ Borel sets, and so, we have shown- open $\times$ Borel = Borel. Now let F be a fixed Borel set in Y and define

$$\mathcal{S}_F \equiv \{E \text{ Borel in } X \text{ such that } E \times F \text{ is Borel in } X \times Y\}.$$

The same argument which was just used shows $\mathcal{S}_F$ is a σ algebra containing the open sets. Therefore, $\mathcal{S}_F =$ the Borel sets, and this proves the lemma since F was an arbitrary Borel set.

Now we define the unit sphere in $\mathbb{R}^n$, S^{n-1}, by

$$S^{n-1} \equiv \{\mathbf{w} \in \mathbb{R}^n : |\mathbf{w}| = 1\}.$$

Then S^{n-1} is a compact metric space using the usual metric on $\mathbb{R}^n$. We define a map

$$\theta : S^{n-1} \times (0, \infty) \to \mathbb{R}^n \setminus \{\mathbf{0}\}$$

by

$$\theta(\mathbf{w}, \rho) \equiv \rho \mathbf{w}.$$

It is clear that θ is one to one and onto with a continuous inverse. Therefore, if $\mathcal{B}_1$ is the set of Borel sets in $S^{n-1} \times (0, \infty)$, and $\mathcal{B}$ are the Borel sets in $\mathbb{R}^n \setminus \{\mathbf{0}\}$, it follows

$$\mathcal{B} = \{\theta(F) : F \in \mathcal{B}_1\}. \tag{7.10}$$

Observe also that the Borel sets of S^{n-1} satisfy the conditions of Lemma 5.6 with Z defined as $S^{n-1}$and the same is true of the sets $(a, b] \cap (0, \infty)$ where $0 \leq a, b \leq \infty$ if Z is defined as $(0, \infty)$. By Corollary 5.7, finite disjoint unions of sets of the form

$$\Big\{E \times I : E \text{ is Borel in } S^{n-1}$$

$$\text{and } I = (a, b] \cap (0, \infty) \text{ where } 0 \leq a, b \leq \infty\Big\}$$

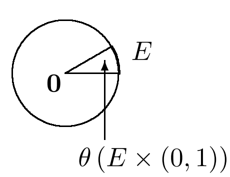
form an algebra of sets, $\mathcal{A}$. It is also clear that $\sigma(\mathcal{A})$ contains the open sets and so $\sigma(\mathcal{A}) = \mathcal{B}_1$ because every set in $\mathcal{A}$ is in $\mathcal{B}_1$ thanks to Lemma 7.22. Let $A_r \equiv S^{n-1} \times (0, r]$ and let

$$\mathcal{M} \equiv \Bigg\{F \in \mathcal{B}_1 : \int_{\mathbb{R}^n} \mathcal{X}_{\theta(F \cap A_r)} dm_n$$

$$= \int_{(0,\infty)} \int_{S^{n-1}} \mathcal{X}_{\theta(F \cap A_r)}(\rho \mathbf{w}) \rho^{n-1} d\sigma dm\Bigg\},$$

where for E a Borel set in S^{n-1},

$$\sigma(E) \equiv n m_n(\theta(E \times (0, 1))). \tag{7.11}$$



Then if $F \in \mathcal{A}$, say $F = E \times (a, b]$, we can show $F \in \mathcal{M}$. This follows easily from the observation that

$$\int_{\mathbb{R}^n} \mathcal{X}_{\theta(F)} dm_n = \int_{\mathbb{R}^n} \mathcal{X}_{\theta(E \times (0,b])}(\mathbf{y}) \, dm_n - \int_{\mathbb{R}^n} \mathcal{X}_{\theta(E \times (0,a])}(\mathbf{y}) \, dm_n$$

$$= m_n(\theta(E \times (0, 1)) \, b^n - m_n(\theta(E \times (0, 1)) \, a^n = \sigma(E) \frac{(b^n - a^n)}{n},$$

a consequence of the change of variables theorem applied to $\mathbf{y} = a\mathbf{x}$, and

$$\int_{(0,\infty)} \int_{S^{n-1}} \mathcal{X}_{\theta(E\times(a,b])}(\rho\mathbf{w})\,\rho^{n-1} d\sigma dm = \int_a^b \int_E \rho^{n-1} d\sigma d\rho$$

$$= \sigma(E)\frac{(b^n - a^n)}{n}.$$

Since it is clear that $\mathcal{M}$ is a monotone class, it follows from the monotone class theorem that $\mathcal{M} = \mathcal{B}_1$. Letting $r \to \infty$, we may conclude that for all $F \in \mathcal{B}_1$,

$$\int_{\mathbb{R}^n} \mathcal{X}_{\theta(F)} dm_n = \int_{(0,\infty)} \int_{S^{n-1}} \mathcal{X}_{\theta(F)}(\rho\mathbf{w})\,\rho^{n-1} d\sigma dm.$$

By (7.10), if A is any Borel set in $\mathbb{R}^n$, then $A \setminus \{\mathbf{0}\} = \theta(F)$ for some $F \in \mathcal{B}_1$. Thus

$$\int_{\mathbb{R}^n} \mathcal{X}_A dm_n = \int_{\mathbb{R}^n} \mathcal{X}_{\theta(F)} dm_n =$$

$$\int_{(0,\infty)} \int_{S^{n-1}} \mathcal{X}_{\theta(F)}(\rho\mathbf{w})\,\rho^{n-1} d\sigma dm = \int_{(0,\infty)} \int_{S^{n-1}} \mathcal{X}_A(\rho\mathbf{w})\,\rho^{n-1} d\sigma dm. \tag{7.12}$$

With this preparation, it is easy to prove the main result which is the following theorem.

Theorem 7.23 *Let $f \geq 0$ and f is Borel measurable on $\mathbb{R}^n$. Then*

$$\int_{\mathbb{R}^n} f(\mathbf{y})\,dm_n = \int_{(0,\infty)} \int_{S^{n-1}} f(\rho\mathbf{w})\,\rho^{n-1} d\sigma dm \tag{7.13}$$

where σ is defined by (7.11) and $\mathbf{y} = \rho\mathbf{w}$, for $\mathbf{w} \in S^{n-1}$.

Proof: From (7.12), (7.13) holds for f replaced with a nonnegative simple function. Now the monotone convergence theorem applied to a sequence of simple functions increasing to f yields the desired conclusion.

7.5 The Lebesgue integral and the Riemann integral

We assume the reader is familiar with the Riemann integral of a function of one variable. It is natural to ask how this is related to the Lebesgue integral where the Lebesgue integral is taken with respect to Lebesgue measure. The following gives the essential result.

Lemma 7.24 *Suppose f is a non negative Riemann integrable function defined on $[a,b]$. Then $\mathcal{X}_{[a,b]}f$ is Lebesgue measurable and $\int_a^b f dx = \int f\mathcal{X}_{[a,b]} dm$ where the first integral denotes the usual Riemann integral and the second integral denotes the Lebesgue integral taken with respect to Lebesgue measure.*

Proof: Since f is Riemann integral, there exist step functions, u_n and l_n of the form

$$\sum_{i=1}^n c_i \mathcal{X}_{[t_{i-1},t_i)}(t)$$

such that $u_n \geq f \geq l_n$ and

$$\begin{aligned} \int_a^b u_n dx &\geq \int_a^b f dx \geq \int_a^b l_n dx, \\ \left| \int_a^b u_n dx - \int_a^b l_n dx \right| &\leq \frac{1}{2^n} \end{aligned}$$

Here $\int_a^b u_n dx$ is an upper sum and $\int_a^b l_n dx$ is a lower sum. We also note that these step functions are Borel measurable simple functions and so $\int_a^b u_n dx = \int u_n dm$ with a similar formula holding for l_n. Replacing l_n with $\max(l_1, \cdots, l_n)$ if necessary, we may assume that the functions, l_n are increasing and similarly, we may assume the u_n are decreasing. Therefore, we may define a Borel measurable function, g, by $g(x) = \lim_{n\to\infty} l_n(x)$ and a Borel measurable function, $h(x)$ by $h(x) = \lim_{n\to\infty} u_n(x)$. We claim that $f(x) = g(x)$ a.e. To see this note that $h(x) \geq f(x) \geq g(x)$ for all x and by the dominated convergence theorem, we can say that

$$\begin{aligned} 0 &= \lim_{n\to\infty} \int_{[a,b]} (u_n - v_n)\, dm \\ &= \int_{[a,b]} (h - g)\, dm \end{aligned}$$

Therefore, $h = g$ a.e. and so off a Borel set of measure zero, $f = g = h$. By completeness of Lebesgue measure, it follows that f is Lebesgue measurable and that both the Lebesgue integral, $\int_{[a,b]} f dm$ and the Riemann integral, $\int_a^b f dx$ are contained in the interval of length 2^{-n},

$$\left[\int_a^b l_n dx, \int_a^b u_n dx\right]$$

showing that these two integrals are equal.

7.6 Exercises

1. If $\mathcal{A}$ is the algebra of sets of Example 5.8, show $\sigma(\mathcal{A})$, the smallest σ algebra containing the algebra, is the Borel sets.

2. Consider the following nested sequence of compact sets, $\{P_n\}$. The set P_n consists of 2^n disjoint closed intervals contained in $[0,1]$. The first interval, P_1, equals $[0,1]$ and P_n is obtained from P_{n-1} by deleting the open interval which is the middle third of each closed interval in P_n. Let $P = \cap_{n=1}^{\infty} P_n$. Show

$$P \neq \emptyset,\ m(P) = 0,\ P \sim [0,1].$$

(There is a 1-1 onto mapping of $[0,1]$ to P.) The set P is called the Cantor set.

3. $\uparrow$ Consider the sequence of functions defined in the following way. We let $f_1(x) = x$ on $[0,1]$. To get from f_n to f_{n+1}, let $f_{n+1} = f_n$ on all intervals where f_n is constant. If f_n is nonconstant on $[a,b]$, let $f_{n+1}(a) = f_n(a)$, $f_{n+1}(b) = f_n(b)$, f_{n+1} is piecewise linear and equal to $\frac{1}{2}(f_n(a) + f_n(b))$ on the middle third of $[a,b]$. Sketch a few of these and you will see the pattern. Show

$$\{f_n\} \text{ converges uniformly on } [0,1]. \tag{a.}$$

If $f(x) = \lim_{n\to\infty} f_n(x)$, show that

$$f(0) = 0,\ f(1) = 1,\ f \text{ is continuous,}$$

and

$$f'(x) = 0 \text{ for all } x \notin P \tag{b.}$$

where P is the Cantor set of Problem 2. This function is called the Cantor function.

4. Suppose X, Y are two locally compact, σ compact, metric spaces. Let $\mathcal{A}$ be the collection of finite disjoint unions of sets of the form $E \times F$ where E and F are Borel sets. Show that $\mathcal{A}$ is an algebra and that the smallest σ algebra containing $\mathcal{A}$, $\sigma(\mathcal{A})$, contains the Borel sets of $X \times Y$. **Hint:** Show $X \times Y$, with the usual product topology, is a σ compact metric space. Next show every open set can be written as a countable union of compact sets. Using this, show every open set can be written as a countable union of open sets of the form $U \times V$ where U is an open set in X and V is an open set in Y.

5. $\uparrow$ Suppose X, Y are two locally compact, σ compact, metric spaces and let μ and ν be Radon measures on X and Y respectively. Define for $f \in C_c(X \times Y)$,
$$\Lambda f \equiv \int_X \int_Y f(x, y)\, d\nu d\mu.$$
Show this is well defined and is a positive linear functional on
$$C_c(X \times Y).$$
Let $(\overline{\mu \times \nu})$be the measure representing Λ. Show that for $f \geq 0$, and f Borel measurable,
$$\int_{X \times Y} f d(\overline{\mu \times \nu}) = \int_X \int_Y f(x, y)\, d v d\mu = \int_Y \int_X f(x, y)\, d\mu d\nu.$$
Hint: First show, using the dominated convergence theorem, that if $E \times F$ is the Cartesian product of two Borel sets each of whom have finite measure, then
$$(\overline{\mu \times \nu})(E \times F) = \mu(E)\,\nu(F) = \int_X \int_Y \mathcal{X}_{E \times F}(x, y)\, d\mu d\nu.$$

6. Let f:$\mathbb{R}^n \rightarrow \mathbb{R}$ be defined by $f(\mathbf{x}) \equiv \left(1 + |\mathbf{x}|^2\right)^k$. Find the values of k for which f is in $L^1(\mathbb{R}^n)$. **Hint:** This is easy and reduces to a one-dimensional problem if you use the formula for integration using polar coordinates.

7. Let B be a Borel set in $\mathbb{R}^n$ and let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^n$. Suppose B has the following property. For each $\mathbf{x} \in \mathbb{R}^n$, $m(\{t : \mathbf{x} + t\mathbf{v} \in B\}) = 0$. Then show $m_n(B) = 0$. Note the condition on B says roughly that B is thin in one direction.

8. Let $f(y) = \mathcal{X}_{(0,1]}(y) \frac{\sin(1/y)}{\sqrt{|y|}}$ and let $g(y) = \mathcal{X}_{(0,1]}(y) \frac{1}{\sqrt{y}}$. For which values of x does it make sense to write the integral $\int_{\mathbb{R}} f(x - y)\, g(y)\, dy$?

9. If $f : \mathbb{R}^n \rightarrow [0, \infty]$ is Lebesgue measurable, show there exists $g : \mathbb{R}^n \rightarrow [0, \infty]$ such that $g = f$ a.e. and g is Borel measurable.

10. $\uparrow$ Let $f \in L^1(\mathbb{R}),\ g \in L^1(\mathbb{R})$. Whereever the integral makes sense, define
$$(f * g)(x) \equiv \int_{\mathbb{R}} f(x - y) g(y) dy.$$
Show the above integral makes sense for a.e. x and that if we define $f * g(x) \equiv 0$ at every point where the above integral does not make sense, it follows that $|(f * g)(x)| < \infty$ a.e. and
$$||f * g||_{L^1} \leq ||f||_{L^1} ||g||_{L^1}. \text{ Here } ||f||_{L^1} \equiv \int |f| dx.$$
Hint: If f is Lebesgue measurable, there exists g Borel measurable with $g(x) = f(x)$ a.e.

11. $\uparrow$ Let $f : [0, \infty) \to \mathbb{R}$ be in $L^1(\mathbb{R}, m)$. The Laplace transform is given by $\widehat{f}(x) = \int_0^\infty e^{-xt} f(t)dt$. Let f, g be in $L^1(\mathbb{R}, m)$, and let $h(x) = \int_0^x f(x - t)g(t)dt$. Show $h \in L^1$, and $\widehat{h} = \widehat{f}\widehat{g}$.

12. Show $\lim_{A\to\infty} \int_0^A \frac{\sin x}{x} dx = \frac{\pi}{2}$. **Hint:** Use $\frac{1}{x} = \int_0^\infty e^{-xt} dt$ and Fubini's theorem. This limit is sometimes called the Cauchy principle value. Note that the function $\sin(x)/x$ is not in L^1 so we are not finding a Lebesgue integral.

13. Let $\mathcal{D}$ consist of functions, $g \in C_c(\mathbb{R}^n)$ which are of the form

$$g(\mathbf{x}) \equiv \prod_{i=1}^n g_i(x_i)$$

where each $g_i \in C_c(\mathbb{R})$. Show that if $f \in C_c(\mathbb{R}^n)$, then there exists a sequence of functions, $\{g_k\}$ in $\mathcal{D}$ which satisfies

$$\lim_{k\to\infty} \sup\{|f(\mathbf{x}) - g_k(\mathbf{x})| : \mathbf{x} \in \mathbb{R}^n\} = 0. \tag{*}$$

Now for $g \in \mathcal{D}$ given as above, let

$$\Lambda_0(g) \equiv \int \cdots \int \prod_{i=1}^n g_i(x_i)\, dm(x_1) \cdots dm(x_n),$$

and define, for arbitrary $f \in C_c(\mathbb{R}^n)$,

$$\Lambda f \equiv \lim_{k\to\infty} \Lambda_0 g_k$$

where * holds. Show this is a well-defined positive linear functional which yields Lebesgue measure. Establish all theorems in this chapter using this as a basis for the definition of Lebesgue measure. Note this approach is arguably less fussy than the presentation in the chapter. **Hint:** You might want to use the Stone Weierstrass theorem.

14. If $f : \mathbb{R}^n \to [0, \infty]$ is Lebesgue measurable, show there exists $g : \mathbb{R}^n \to [0, \infty]$ such that $g = f$ a.e. and g is Borel measurable.

15. Let E be countable subset of $\mathbb{R}$. Show $m(E) = 0$. **Hint:** Let the set be $\{e_i\}_{i=1}^\infty$ and let e_i be the center of an open interval of length $\epsilon/2^i$.

16. $\uparrow$ If S is an uncountable set of irrational numbers, is it necessary that S has a rational number as a limit point? **Hint:** Consider the proof of Problem 15 when applied to the rational numbers. (This problem was shown to me by Lee Earlbach.)

Product Measure

There is a general procedure for constructing a measure space on a σ algebra of subsets of the Cartesian product of two given measure spaces. This leads naturally to a discussion of iterated integrals. In calculus, we learn how to obtain multiple integrals by evaluation of iterated integrals. We are asked to believe that the iterated integrals taken in the different orders give the same answer. The following simple example shows that sometimes when iterated integrals are performed in different orders, the results differ.

Example 8.1 *Let $0 < \delta_1 < \delta_2 < \cdots < \delta_n \cdots < 1, \lim_{n\to\infty} \delta_n = 1$. Let g_n be a real continuous function with $g_n = 0$ outside of (δ_n, δ_{n+1}) and $\int_0^1 g_n(x)dx = 1$ for all n. Define*

$$f(x,y) = \sum_{n=1}^{\infty} (g_n(x) - g_{n+1}(x))g_n(y).$$

Then you can show the following:

a.) f is continuous on $[0,1) \times [0,1)$

b.) $\int_0^1 \int_0^1 f(x,y)dydx = 1\,,\ \int_0^1 \int_0^1 f(x,y)dxdy = 0.$

Nevertheless, it is often the case that the iterated integrals are equal and give the value of an appropriate multiple integral. The best theorems of this sort are to be found in the theory of Lebesgue integration and this is what will be discussed in this chapter.

Definition 8.2 *A measure space $(X, \mathcal{F}, \mu)$ is said to be σ finite if*

$$X = \cup_{n=1}^{\infty} X_n, \ \ X_n \in \mathcal{F}, \ \ \mu(X_n) < \infty.$$

In the rest of this chapter, unless otherwise stated, $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{F}, \lambda)$ will be two σ finite measure spaces. Note that a Radon measure on a σ compact, locally compact space gives an example of a σ finite space. In particular, Lebesgue measure is σ finite.

Definition 8.3 *A measurable rectangle is a set $A \times B \subseteq X \times Y$ where $A \in \mathcal{S}, B \in \mathcal{F}$. An elementary set will be any subset of $X \times Y$ which is a finite union of disjoint measurable rectangles. $\mathcal{S} \times \mathcal{F}$ will denote the smallest σ algebra of sets in $\mathcal{P}(X \times Y)$ containing all elementary sets.*

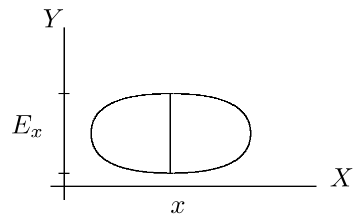
Example 8.4 *It follows from Lemma 5.6 or more easily from Corollary 5.7 that the elementary sets form an algebra.*

Definition 8.5 *Let $E \subseteq X \times Y$,*

$$E_x = \{y \in Y : (x,y) \in E\},$$

$$E^y = \{x \in X : (x,y) \in E\}.$$

These are called the x and y sections.



Theorem 8.6 *If $E \in \mathcal{S} \times \mathcal{F}$, then $E_x \in \mathcal{F}$ and $E^y \in \mathcal{S}$ for all $x \in X$ and $y \in Y$.*

Proof: Let

$$\mathcal{M} = \{E \subseteq \mathcal{S} \times \mathcal{F} \text{ such that } E_x \in \mathcal{F},$$

$$E^y \in \mathcal{S} \text{ for all } x \in X \text{ and } y \in Y.\}$$

Then $\mathcal{M}$ contains all measurable rectangles. If $E_i \in \mathcal{M}$,

$$(\cup_{i=1}^{\infty} E_i)_x = \cup_{i=1}^{\infty} (E_i)_x \in \mathcal{F}.$$

Similarly, $(\cup_{i=1}^{\infty} E_i)^y \in \mathcal{S}$. $\mathcal{M}$ is thus closed under countable unions. If $E \in \mathcal{M}$,

$$\left(E^C\right)_x = (E_x)^C \in \mathcal{F}.$$

Similarly, $\left(E^C\right)^y \in \mathcal{S}$. Thus $\mathcal{M}$ is closed under complementation. Therefore $\mathcal{M}$ is a σ-algebra containing the elementary sets. Hence, $\mathcal{M} \supseteq \mathcal{S} \times \mathcal{F}$. But $\mathcal{M} \subseteq \mathcal{S} \times \mathcal{F}$. Therefore $\mathcal{M} = \mathcal{S} \times \mathcal{F}$ and the theorem is proved.

It follows from Lemma 5.6 that the elementary sets form an algebra because clearly the intersection of two measurable rectangles is a measurable rectangle and

$$(A \times B) \setminus (A_0 \times B_0) = (A \setminus A_0) \times B \cup (A \cap A_0) \times (B \setminus B_0),$$

an elementary set. We use this in the next theorem.

Theorem 8.7 *If $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{F}, \lambda)$ are both finite measure spaces ($\mu(X)$, $\lambda(Y) < \infty$), then for every $E \in \mathcal{S} \times \mathcal{F}$,*
a.) $x \to \lambda(E_x)$ is μ measurable, $y \to \mu(E^y)$ is λ measurable
b.) $\int_X \lambda(E_x) d\mu = \int_Y \mu(E^y) d\lambda$.

Proof: Let $\mathcal{M} = \{E \in \mathcal{S} \times \mathcal{F}$ such that both a.) and b.) hold$\}$. Since μ and λ are both finite, the monotone convergence and dominated convergence theorems imply that $\mathcal{M}$ is a monotone class. Clearly $\mathcal{M}$ contains the algebra of elementary sets. By the monotone class theorem, $\mathcal{M} \supseteq \mathcal{S} \times \mathcal{F}$.

Theorem 8.8 *If $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{F}, \lambda)$ are both σ-finite measure spaces, then for every $E \in \mathcal{S} \times \mathcal{F}$,*
a.) $x \to \lambda(E_x)$ is μ measurable, $y \to \mu(E^y)$ is λ measurable.
b.) $\int_X \lambda(E_x) d\mu = \int_Y \mu(E^y) d\lambda$.

Proof: Let $X = \cup_{n=1}^{\infty} X_n, Y = \cup_{n=1}^{\infty} Y_n$ where,

$$X_n \subseteq X_{n+1}, Y_n \subseteq Y_{n+1}, \mu(X_n) < \infty, \lambda(Y_n) < \infty.$$

Let

$$\mathcal{S}_n = \{A \cap X_n : A \in \mathcal{S}\},\ \mathcal{F}_n = \{B \cap Y_n : B \in \mathcal{F}\}.$$

Thus $(X_n, \mathcal{S}_n, \mu)$ and $(Y_n, \mathcal{F}_n, \lambda)$ are both finite measure spaces.

Claim: If $E \in \mathcal{S} \times \mathcal{F}$, then $E \cap (X_n \times Y_n) \in \mathcal{S}_n \times \mathcal{F}_n$.

Proof: Let $\mathcal{M}_n = \{E \in \mathcal{S} \times \mathcal{F} : E \cap (X_n \times Y_n) \in \mathcal{S}_n \times \mathcal{F}_n\}$. Clearly $\mathcal{M}_n$ contains the algebra of elementary sets. It is also clear that $\mathcal{M}_n$ is a monotone class. Thus $\mathcal{M}_n = \mathcal{S} \times \mathcal{F}$.

Now let $E \in \mathcal{S} \times \mathcal{F}$. By Theorem 8.7,

$$\int_{X_n} \lambda((E \cap (X_n \times Y_n))_x) d\mu = \int_{Y_n} \mu((E \cap (X_n \times Y_n))^y) d\lambda \tag{8.1}$$

where the integrands are measurable. Now

$$(E \cap (X_n \times Y_n))_x = \emptyset$$

if $x \notin X_n$ and a similar observation holds for the second integrand in (8.1). Therefore,

$$\int_X \lambda((E \cap (X_n \times Y_n))_x) d\mu = \int_Y \mu((E \cap (X_n \times Y_n))^y) d\lambda.$$

Then letting $n \to \infty$, we use the monotone convergence theorem to get b.). The measurability assertions of a.) are valid because the limit of a sequence of measurable functions is measurable.

Definition 8.9 *For $E \in \mathcal{S} \times \mathcal{F}$ and $(X, \mathcal{S}, \mu), (Y, \mathcal{F}, \lambda)$ σ-finite, $(\mu \times \lambda)(E) \equiv \int_X \lambda(E_x) d\mu = \int_Y \mu(E^y) d\lambda$.*

This definition is well defined because of Theorem 8.8. We also have the following theorem.

Theorem 8.10 *If $A \in \mathcal{S}, B \in \mathcal{F}$, then $(\mu \times \lambda)(A \times B) = \mu(A)\lambda(B)$, and $\mu \times \lambda$ is a measure on $\mathcal{S} \times \mathcal{F}$ called product measure.*

The proof of Theorem 8.10 is obvious and is left to the reader. Use the Monotone Convergence theorem. The next theorem is one of several theorems due to Fubini and Tonelli. These theorems all have to do with interchanging the order of integration in a multiple integral. The main ideas are illustrated by the next theorem which is often referred to as Fubini's theorem.

Theorem 8.11 *Let $f : X \times Y \to [0, \infty]$ be measurable with respect to $\mathcal{S} \times \mathcal{F}$ and suppose μ and λ are σ-finite. Then*

$$\int_{X \times Y} f d(\mu \times \lambda) = \int_X \int_Y f(x, y) d\lambda d\mu = \int_Y \int_X f(x, y) d\mu d\lambda \tag{8.2}$$

and all integrals make sense.

Proof: For $E \in \mathcal{S} \times \mathcal{F}$, we note

$$\int_Y \mathcal{X}_E(x, y) d\lambda = \lambda(E_x),\ \int_X \mathcal{X}_E(x, y) d\mu = \mu(E^y).$$

Thus from Definition 8.9, (8.2) holds if $f = \mathcal{X}_E$. It follows that (8.2) holds for every nonnegative simple function. By Theorem 5.31, there exists an increasing sequence, $\{f_n\}$, of simple functions converging pointwise to f. Then

$$\int_Y f(x, y) d\lambda = \lim_{n \to \infty} \int_Y f_n(x, y) d\lambda,$$

$$\int_X f(x,y)d\mu = \lim_{n\to\infty} \int_X f_n(x,y)d\mu.$$

This follows from the monotone convergence theorem. Since

$$x \to \int_Y f_n(x,y)d\lambda$$

is measurable with respect to $\mathcal{S}$, it follows that $x \to \int_Y f(x,y)d\lambda$ is also measurable with respect to $\mathcal{S}$. A similar conclusion can be drawn about $y \to \int_X f(x,y)d\mu$. Thus the two iterated integrals make sense. Since (8.2) holds for f_n, another application of the Monotone Convergence theorem shows (8.2) holds for f. This proves the theorem.

Corollary 8.12 *Let $f : X\times Y \to \mathbb{C}$ be $\mathcal{S}\times\mathcal{F}$ measurable. Suppose either $\int_X \int_Y |f|\, d\lambda d\mu$ or $\int_Y \int_X |f|\, d\mu d\lambda < \infty$. Then $f \in L^1(X\times Y, \mu\times\lambda)$ and*

$$\int_{X\times Y} fd(\mu\times\lambda) = \int_X\int_Y fd\lambda d\mu = \int_Y\int_X fd\mu d\lambda \tag{8.3}$$

with all integrals making sense.

Proof: Suppose first that f is real-valued. Apply Theorem 8.11 to $f^+$and f^-. (8.3) follows from observing that $f = f^+ - f^-$; and that all integrals are finite. If f is complex valued, consider real and imaginary parts. This proves the corollary.

How can we tell if f is $\mathcal{S}\times\mathcal{F}$ measurable? The following theorem gives a convenient way for many examples.

Theorem 8.13 *If X and Y are topological spaces having a countable basis of open sets and if $\mathcal{S}$ and $\mathcal{F}$ both contain the open sets, then $\mathcal{S}\times\mathcal{F}$ contains the Borel sets.*

Proof: We need to show $\mathcal{S}\times\mathcal{F}$ contains the open sets in $X\times Y$. If $\mathcal{B}$ is a countable basis for the topology of X and if $\mathcal{C}$ is a countable basis for the topology of Y, then

$$\{B\times C : B\in\mathcal{B}, C\in\mathcal{C}\}$$

is a countable basis for the topology of $X\times Y$. (Remember a basis for the topology of $X\times Y$ is the collection of sets of the form $U\times V$ where U is open in X and V is open in Y.) Thus every open set is a countable union of sets $B\times C$ where $B\in\mathcal{B}$ and $C\in\mathcal{C}$. Since $B\times C$ is a measurable rectangle, it follows that every open set in $X\times Y$ is in $\mathcal{S}\times\mathcal{F}$. This proves the theorem.

The importance of this theorem is that we can use it to assert that a function is product measurable if it is Borel measurable. For an example of how this can sometimes be done, see Problem 5 in this chapter.

Theorem 8.14 *Suppose $\mathcal{S}$ and $\mathcal{F}$ are Borel, μ and λ are regular on $\mathcal{S}$ and $\mathcal{F}$ respectively, and $\mathcal{S}\times\mathcal{F}$ is Borel. Then $\mu\times\lambda$ is regular on $\mathcal{S}\times\mathcal{F}$. (Recall Theorem 8.13 for a sufficient condition for $\mathcal{S}\times\mathcal{F}$ to be Borel.)*

Proof: Let $\mu(X_n) < \infty, \lambda(Y_n) < \infty$, and $X_n \uparrow X, Y_n \uparrow Y$. Let $R_n = X_n\times Y_n$ and define

$$\mathcal{G}_n = \{S\in\mathcal{S}\times\mathcal{F} : \mu\times\lambda \text{ is regular on } S\cap R_n\}.$$

By this we mean that for $S\in\mathcal{G}_n$

$$(\mu\times\lambda)(S\cap R_n) = \inf\{(\mu\times\lambda)(V) : V \text{ is open and } V\supseteq S\cap R_n\}$$

and

$$(\mu \times \lambda)(S \cap R_n) =$$

$$\sup\{(\mu \times \lambda)(K) : K \text{ is compact and } K \subseteq S \cap R_n\}.$$

If $P \times Q$ is a measurable rectangle, then

$$(P \times Q) \cap R_n = (P \cap X_n) \times (Q \cap Y_n).$$

Let $K_x \subseteq (P \cap X_n)$ and $K_y \subseteq (Q \cap Y_n)$ be such that

$$\mu(K_x) + \varepsilon > \mu(P \cap X_n)$$

and

$$\lambda(K_y) + \varepsilon > \lambda(Q \cap Y_n).$$

By Theorem 3.30 $K_x \times K_y$ is compact and from the definition of product measure,

$$(\mu \times \lambda)(K_x \times K_y) = \mu(K_x)\lambda(K_y)$$

$$\geq \mu(P \cap X_n)\lambda(Q \cap Y_n) - \varepsilon(\lambda(Q \cap Y_n) + \mu(P \cap X_n)) + \varepsilon^2.$$

Since ε is arbitrary, this verifies that $(\mu \times \lambda)$ is inner regular on $S \cap R_n$ whenever S is an elementary set. Similarly, $(\mu \times \lambda)$ is outer regular on $S \cap R_n$ whenever S is an elementary set. Thus $\mathcal{G}_n$ contains the elementary sets.

Next we show that $\mathcal{G}_n$ is a monotone class. If $S_k \downarrow S$ and $S_k \in \mathcal{G}_n$, let K_k be a compact subset of $S_k \cap R_n$ with

$$(\mu \times \lambda)(K_k) + \varepsilon 2^{-k} > (\mu \times \lambda)(S_k \cap R_n).$$

Let $K = \cap_{k=1}^{\infty} K_k$. Then

$$S \cap R_n \setminus K \subseteq \cup_{k=1}^{\infty}(S_k \cap R_n \setminus K_k).$$

Therefore

$$\begin{aligned} (\mu \times \lambda)(S \cap R_n \setminus K) &\leq \sum_{k=1}^{\infty}(\mu \times \lambda)(S_k \cap R_n \setminus K_k) \\ &\leq \sum_{k=1}^{\infty} \varepsilon 2^{-k} = \varepsilon. \end{aligned}$$

Now let $V_k \supseteq S_k \cap R_n$, V_k is open and

$$(\mu \times \lambda)(S_k \cap R_n) + \varepsilon > (\mu \times \lambda)(V_k).$$

Let k be large enough that

$$(\mu \times \lambda)(S_k \cap R_n) - \varepsilon < (\mu \times \lambda)(S \cap R_n).$$

Then $(\mu \times \lambda)(S \cap R_n) + 2\varepsilon > (\mu \times \lambda)(V_k)$. This shows $\mathcal{G}_n$ is closed with respect to intersections of decreasing sequences of its elements. The consideration of increasing sequences is similar. By the monotone class theorem, $\mathcal{G}_n = \mathcal{S} \times \mathcal{F}$.

Now let $S \in \mathcal{S} \times \mathcal{F}$ and let $l < (\mu \times \lambda)(S)$. Then $l < (\mu \times \lambda)(S \cap R_n)$ for some n. It follows from the first part of this proof that there exists a compact subset of $S \cap R_n, K$, such that $(\mu \times \lambda)(K) > l$. It follows that $(\mu \times \lambda)$ is inner regular on $\mathcal{S} \times \mathcal{F}$. To verify that the product measure is outer regular on $\mathcal{S} \times \mathcal{F}$, let V_n be an open set such that

$$V_n \supseteq S \cap R_n, \ (\mu \times \lambda)(V_n \setminus (S \cap R_n)) < \varepsilon 2^{-n}.$$

Let $V = \cup_{n=1}^{\infty} V_n$. Then $V \supseteq S$ and

$$V \setminus S \subseteq \cup_{n=1}^{\infty} V_n \setminus (S \cap R_n).$$

Thus,

$$(\mu \times \lambda)(V \setminus S) \leq \sum_{n=1}^{\infty} \varepsilon 2^{-n} = \varepsilon$$

and so

$$(\mu \times \lambda)(V) \leq \varepsilon + (\mu \times \lambda)(S).$$

This proves the theorem.

8.1 Measures on infinite products

It is important in some applications to consider measures on infinite, even uncountably many products. In order to accomplish this, we first give a simple and fundamental theorem of Caratheodory called the Caratheodory extension theorem.

Definition 8.15 *Let $\mathcal{E}$ be an algebra of sets of Ω and let μ_0 be a finite measure on $\mathcal{E}$. By this we mean that μ_0 is finitely additive and if E_i, E are sets of $\mathcal{E}$ with the E_i disjoint and*

$$E = \cup_{i=1}^{\infty} E_i,$$

then

$$\mu_0 (E) = \sum_{i=1}^{\infty} \mu (E_i)$$

while $\mu_0 (\Omega) < \infty$.

In this definition, μ_0 is trying to be a measure and acts like one whenever possible. Under these conditions, we can show that μ_0 can be extended uniquely to a complete measure, μ, defined on a σ algebra of sets containing $\mathcal{E}$ and such that μ agrees with μ_0 on $\mathcal{E}$. We will prove the following important theorem which is the main result in this section.

Theorem 8.16 *Let μ_0 be a measure on an algebra of sets, $\mathcal{E}$, which satisfies $\mu_0 (\Omega) < \infty$. Then there exists a complete measure space $(\Omega, \mathcal{S}, \mu)$ such that*

$$\mu (E) = \mu_0 (E)$$

for all $E \in \mathcal{E}$. Also if ν is any such measure which agrees with μ_0 on $\mathcal{E}$, then $\nu = \mu$ on $\sigma (\mathcal{E})$, the σ algebra generated by $\mathcal{E}$.

Proof: We define an outer measure as follows.

$$\mu(S) \equiv \inf \left\{ \sum_{i=1}^{\infty} \mu_0(E_i) : S \subseteq \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{E} \right\}$$

Claim 1: μ is an outer measure.
Proof of Claim 1: Let $S \subseteq \cup_{i=1}^{\infty} S_i$ and let $S_i \subseteq \cup_{j=1}^{\infty} E_{ij}$, where

$$\mu(S_i) + \frac{\epsilon}{2^i} \geq \sum_{j=1}^{\infty} \mu(E_{ij}).$$

Then

$$\mu(S) \leq \sum_i \sum_j \mu(E_{ij}) = \sum_i \left(\mu(S_i) + \frac{\epsilon}{2^i} \right) = \sum_i \mu(S_i) + \epsilon.$$

Since ϵ is arbitrary, this shows μ is an outer measure as claimed.

By the Caratheodory procedure, there exists a unique σ algebra, $\mathcal{S}$, consisting of the μ measurable sets such that

$$(\Omega, \mathcal{S}, \mu)$$

is a complete measure space. It remains to show μ extends μ_0.
Claim 2: If $\mathcal{S}$ is the σ algebra of μ measurable sets, $\mathcal{S} \supseteq \mathcal{E}$ and $\mu = \mu_0$ on $\mathcal{E}$.
Proof of Claim 2: First we observe that if $A \in \mathcal{E}$, then $\mu(A) \leq \mu_0(A)$ by definition. Letting

$$\mu(A) + \epsilon > \sum_{i=1}^{\infty} \mu_0(E_i), \ \cup_{i=1}^{\infty} E_i \supseteq A,$$

it follows

$$\mu(A) + \epsilon > \sum_{i=1}^{\infty} \mu_0(E_i \cap A) \geq \mu_0(A)$$

since $A = \cup_{i=1}^{\infty} E_i \cap A$. Therefore, $\mu = \mu_0$ on $\mathcal{E}$.

Next we need show $\mathcal{E} \subseteq \mathcal{S}$. Let $A \in \mathcal{E}$ and let $S \subseteq \Omega$ be any set. There exist sets $\{E_i\} \subseteq \mathcal{E}$ such that $\cup_{i=1}^{\infty} E_i \supseteq S$ but

$$\mu(S) + \epsilon > \sum_{i=1}^{\infty} \mu(E_i).$$

Then

$$\mu(S) \leq \mu(S \cap A) + \mu(S \setminus A)$$

$$\leq \mu(\cup_{i=1}^{\infty} E_i \setminus A) + \mu(\cup_{i=1}^{\infty} (E_i \cap A))$$

$$\leq \sum_{i=1}^{\infty} \mu(E_i \setminus A) + \sum_{i=1}^{\infty} \mu(E_i \cap A) = \sum_{i=1}^{\infty} \mu(E_i) < \mu(S) + \epsilon.$$

Since ϵ is arbitrary, this shows $A \in \mathcal{S}$.

With these two claims, we have established the existence part of the theorem. To verify uniqueness, Let

$$\mathcal{M} \equiv \{E \in \sigma(\mathcal{E}) : \mu(E) = \nu(E)\}.$$

Then $\mathcal{M}$ is given to contain $\mathcal{E}$ and is obviously a monotone class. Therefore by the theorem on monotone classes, $\mathcal{M} = \sigma(\mathcal{E})$ and this proves the lemma.

With this result we are ready to consider the Kolmogorov theorem about measures on infinite product spaces. One can consider product measure for infinitely many factors. The Caratheodory extension theorem above implies an important theorem due to Kolmogorov which we will use to give an interesting application of the individual ergodic theorem. The situation involves a probability space, $(\Omega, \mathcal{S}, P)$, an index set, I possibly infinite, even uncountably infinite, and measurable functions, $\{X_t\}_{t\in I}$, $X_t : \Omega \to \mathbb{R}$. These measurable functions are called random variables in this context. It is convenient to consider the topological space

$$[-\infty,\infty]^I \equiv \prod_{t\in I}[-\infty,\infty]$$

with the product topology where a subbasis for the topology of $[-\infty,\infty]$ consists of sets of the form $[-\infty,b)$ and $(a,\infty]$ where $a,b\in\mathbb{R}$. Thus $[-\infty,\infty]$ is a compact set with respect to this topology and by Tychonoff's theorem, so is $[-\infty,\infty]^I$.

Let $J \subseteq I$. Then if $\mathbf{E} \equiv \prod_{t\in I} E_t$, we define

$$\gamma_J \mathbf{E} \equiv \prod_{t\in I} F_t$$

where

$$F_t = \begin{cases} E_t \text{ if } t\in J \\ [-\infty,\infty] \text{ if } t\notin J \end{cases}$$

Thus $\gamma_J\mathbf{E}$ leaves alone E_t for $t\in J$ and changes the other E_t into $[-\infty,\infty]$. Also we define for J a finite subset of I,

$$\pi_J \mathbf{x} \equiv \prod_{t\in J} x_t$$

so π_J is a continuous mapping from $[-\infty,\infty]^I$ to $[-\infty,\infty]^J$.

$$\pi_J\mathbf{E} \equiv \prod_{t\in J} E_t.$$

Note that for J a finite subset of I,

$$\prod_{t\in J}[-\infty,\infty] = [-\infty,\infty]^J$$

is a compact metric space with respect to the metric,

$$d(\mathbf{x},\mathbf{y}) = \max\{d_t(x_t,y_t) : t\in J\},$$

where

$$d_t(x,y) \equiv |\arctan(x) - \arctan(y)|, \text{ for } x,y\in[-\infty,\infty].$$

We leave this assertion as an exercise for the reader. You can show that this is a metric and that the metric just described delivers the usual product topology which will prove the assertion. Now we define for J a finite subset of I,

$$\begin{aligned} \mathcal{R}_J &\equiv \left\{ \mathbf{E} = \prod_{t \in I} E_t \ : \gamma_J \mathbf{E} = \mathbf{E},\ E_t \text{ a Borel set in } [-\infty,\infty]^J \right\} \\ \mathcal{R} &\equiv \cup \{\mathcal{R}_J : J \subseteq I, \text{ and } J \text{ finite}\} \end{aligned}$$

Thus $\mathcal{R}$ consists of those sets of $[-\infty,\infty]^I$ for which every slot is filled with $[-\infty,\infty]$ except for a finite set, $J \subseteq I$ where the slots are filled with a Borel set, E_t. We define $\mathcal{E}$ as finite disjoint unions of sets of $\mathcal{R}$. In fact $\mathcal{E}$ is an algebra of sets.

Lemma 8.17 *The sets, $\mathcal{E}$ defined above form an algebra of sets of*

$$[-\infty,\infty]^I .$$

Proof: Clearly $\emptyset$ and $[-\infty,\infty]^I$ are both in $\mathcal{E}$. Suppose $\mathbf{A}, \mathbf{B} \in \mathcal{R}$. Then for some finite set, J,

$$\gamma_J \mathbf{A} = \mathbf{A},\ \gamma_J \mathbf{B} = \mathbf{B}.$$

Then

$$\gamma_J (\mathbf{A} \setminus \mathbf{B}) = \mathbf{A} \setminus \mathbf{B} \in \mathcal{E}, \gamma_J (\mathbf{A} \cap \mathbf{B}) = \mathbf{A} \cap \mathbf{B} \in \mathcal{R}.$$

By Lemma 5.6 this shows $\mathcal{E}$ is an algebra.

Let $\mathbf{X}_J = (X_{t_1}, \cdots, X_{t_m})$ where $\{t_1, \cdots, t_m\} = J$. We may define a Radon probability measure, $\lambda_{\mathbf{X}_J}$ on a σ algebra of sets of $[-\infty,\infty]^J$ as follows. For E a Borel set in $[-\infty,\infty]^J$,

$$\widehat{\lambda}_{\mathbf{X}_J}(E) \equiv P(\{\omega : \mathbf{X}_J(\omega) \in E\}).$$

(Remember the random variables have values in $\mathbb{R}$ so they do not take the value $\pm\infty$.) Now since $\widehat{\lambda}_{\mathbf{X}_J}$ is a probability measure, it is certainly finite on the compact sets of $[-\infty,\infty]^J$. Also note that if $\overline{\mathcal{B}}$ denotes the Borel sets of $[-\infty,\infty]^J$, then $\mathcal{B} \equiv \left\{E \cap (-\infty,\infty) : E \in \overline{\mathcal{B}}\right\}$ is a σ algebra of sets of $(-\infty,\infty)^J$ which contains the open sets. Therefore, $\mathcal{B}$ contains the Borel sets of $(-\infty,\infty)^J$ and so we can apply Theorem 6.23 to conclude there is a unique Radon measure extending $\widehat{\lambda}_{\mathbf{X}_J}$ which will be denoted as $\lambda_{\mathbf{X}}$.

For $\mathbf{E} \in \mathcal{R}$, with $\gamma_J \mathbf{E} = \mathbf{E}$ we define

$$\lambda_J(\mathbf{E}) \equiv \lambda_{\mathbf{X}_J}(\pi_J \mathbf{E})$$

Theorem 8.18 *(Kolmogorov) There exists a complete probability measure space,*

$$\left([-\infty,\infty]^I, \mathcal{S}, \lambda\right)$$

such that if $\mathbf{E} \in \mathcal{R}$ and $\gamma_J \mathbf{E} = \mathbf{E}$,

$$\lambda(\mathbf{E}) = \lambda_J(\mathbf{E}).$$

The measure is unique on $\sigma(\mathcal{E})$, the smallest σ algebra containing $\mathcal{E}$. If $\mathbf{A} \subseteq [-\infty,\infty]^I$ is a set having measure 1, then

$$\lambda_J(\pi_J \mathbf{A}) = 1$$

for every finite $J \subseteq I$.

Proof: We first describe a measure, λ_0 on $\mathcal{E}$ and then extend this measure using the Caratheodory extension theorem. If $\mathbf{E} \in \mathcal{R}$ is such that $\gamma_J(\mathbf{E}) = \mathbf{E}$, then λ_0 is defined as

$$\lambda_0(\mathbf{E}) \equiv \lambda_J(\mathbf{E}).$$

Note that if $J \subseteq J_1$ and $\gamma_J(\mathbf{E}) = \mathbf{E}$ and $\gamma_{J_1}(\mathbf{E}) = \mathbf{E}$, then $\lambda_J(\mathbf{E}) = \lambda_{J_1}(\mathbf{E})$ because

$$\mathbf{X}_{J_1}^{-1}(\pi_{J_1}\mathbf{E}) = \mathbf{X}_{J_1}^{-1}\left(\pi_J\mathbf{E} \cap \prod_{J_1 \setminus J}[-\infty,\infty]\right) = \mathbf{X}_J^{-1}(\pi_J\mathbf{E}).$$

Therefore λ_0 is finitely additive on $\mathcal{E}$ and is well defined. We need to verify that λ_0 is actually a measure on $\mathcal{E}$. Let $\mathbf{A}_n \downarrow \emptyset$ where $\mathbf{A}_n \in \mathcal{E}$ and $\gamma_{J_n}(\mathbf{A}_n) = \mathbf{A}_n$. We need to show that

$$\lambda_0(\mathbf{A}_n) \downarrow 0.$$

Suppose to the contrary that

$$\lambda_0(\mathbf{A}_n) \downarrow \epsilon > 0.$$

By regularity of the Radon measure, $\lambda_{\mathbf{X}_{J_n}}$, there exists a compact set,

$$\mathbf{K}_n \subseteq \pi_{J_n}\ \mathbf{A}_n$$

such that

$$\lambda_{\mathbf{X}_{J_n}}((\pi_{J_n}\mathbf{A}_n) \setminus \mathbf{K}_n) < \frac{\epsilon}{2^{n+1}}. \tag{8.4}$$

Claim: We can assume $\mathbf{H}_n \equiv \pi_{J_n}^{-1}\mathbf{K}_n$ is in $\mathcal{E}$ in addition to (8.4).

Proof of the claim: Since $\mathbf{A}_n \in \mathcal{E}$, we know $\mathbf{A}_n$ is a finite disjoint union of sets of $\mathcal{R}$. Therefore, it suffices to verify the claim under the assumption that $\mathbf{A}_n \in \mathcal{R}$. Suppose then that

$$\pi_{J_n}(\mathbf{A}_n) = \prod_{t \in J_n} E_t$$

where the E_t are Borel sets. Let K_t be defined by

$$K_t \equiv \left\{x_t : \text{ for some } \mathbf{y} = \prod_{s \neq t, s \in J_n} y_s,\ x_t \times \mathbf{y} \in \mathbf{K}_n\right\}.$$

Then using the compactness of K_n, it is easy to see that K_t is a closed subset of $[-\infty,\infty]$ and is therefore, compact. Now let

$$\mathbf{K}'_n \equiv \prod_{t \in J_n} K_t.$$

It follows $\mathbf{K}'_n \supseteq \mathbf{K}_n$ is a compact subset of $\prod_{t \in J_n} E_t = \pi_{J_n}\mathbf{A}_n$ and $\pi_{J_n}^{-1}\mathbf{K}'_n \in \mathcal{R} \subseteq \mathcal{E}$. It follows we could have picked $\mathbf{K}'_n$. This proves the claim.

Let $\mathbf{H}_n \equiv \pi_{J_n}^{-1}(\mathbf{K}_n)$. By the Tychonoff theorem, $\mathbf{H}_n$ is compact in $[-\infty,\infty]^I$ because it is a closed subset of $[-\infty,\infty]^I$, a compact space. (Recall that π_J is continuous.) Also $\mathbf{H}_n$ is a set of $\mathcal{E}$. Thus,

$$\lambda_0(\mathbf{A}_n \setminus \mathbf{H}_n) = \lambda_{\mathbf{X}_{J_n}}((\pi_{J_n}\mathbf{A}_n) \setminus \mathbf{K}_n) < \frac{\epsilon}{2^{n+1}}.$$

It follows $\{\mathbf{H}_n\}$ has the finite intersection property for if $\cap_{k=1}^m \mathbf{H}_k = \emptyset$, then

$$\epsilon \leq \lambda_0 \left(\mathbf{A}_m \setminus \cap_{k=1}^m \mathbf{H}_k\right) \leq \sum_{k=1}^m \lambda_0 \left(\mathbf{A}_k \setminus \mathbf{H}_k\right) < \sum_{k=1}^\infty \frac{\epsilon}{2^{k+1}} = \frac{\epsilon}{2},$$

a contradiction. Now since these sets have the finite intersection property, it follows

$$\cap_{k=1}^\infty \mathbf{A}_k \supseteq \cap_{k=1}^\infty \mathbf{H}_k \neq \emptyset,$$

a contradiction.

Now we show this implies λ_0 is a measure. If $\{\mathbf{E}_i\}$ are disjoint sets of $\mathcal{E}$ with

$$\mathbf{E} = \cup_{i=1}^\infty \mathbf{E}_i,\ \mathbf{E}, \mathbf{E}_i \in \mathcal{E},$$

then $\mathbf{E} \setminus \cup_{i=1}^n \mathbf{E}_i$ decreases to the empty set and so

$$\lambda_0 \left(\mathbf{E} \setminus \cup_{i=1}^n \mathbf{E}_i\right) = \lambda_0 \left(\mathbf{E}\right) - \sum_{i=1}^n \lambda_0 \left(\mathbf{E}_i\right) \to 0.$$

Thus $\lambda_0 \left(\mathbf{E}\right) = \sum_{k=1}^\infty \lambda_0 \left(\mathbf{E}_k\right)$. Now an application of the Caratheodory extension theorem proves the main part of the theorem. It remains to verify the last assertion.

To verify this assertion, let $\mathbf{A}$ be a set of measure 1. We note that $\pi_J^{-1}\left(\pi_J \mathbf{A}\right) \supseteq \mathbf{A}$, $\gamma_J\left(\pi_J^{-1}\left(\pi_J \mathbf{A}\right)\right) = \pi_J^{-1}\left(\pi_J \mathbf{A}\right)$, and $\pi_J\left(\pi_J^{-1}\left(\pi_J \mathbf{A}\right)\right) = \pi_J \mathbf{A}$. Therefore,

$$1 = \lambda\left(\mathbf{A}\right) \leq \lambda\left(\pi_J^{-1}\left(\pi_J \mathbf{A}\right)\right) = \lambda_J\left(\pi_J\left(\pi_J^{-1}\left(\pi_J \mathbf{A}\right)\right)\right) = \lambda_J\left(\pi_J \mathbf{A}\right) \leq 1.$$

This proves the theorem.

8.2 A strong ergodic theorem

Here we give an application of the individual ergodic theorem, Theorem 5.51. Let $\mathbb{Z}$ denote the integers and let $\{X_n\}_{n\in\mathbb{Z}}$ be a sequence of real valued random variables defined on a probability space, $(\Omega, \mathcal{F}, P)$. Let $T : [-\infty, \infty]^{\mathbb{Z}} \to [-\infty, \infty]^{\mathbb{Z}}$ be defined by

$$T\left(\mathbf{x}\right)_n = x_{n+1},\ \text{where } \mathbf{x} = \prod_{n\in\mathbb{Z}} \{x_n\}$$

Thus T slides the sequence to the left one slot. By the Kolmogorov theorem there exists a unique measure, λ defined on $\sigma\left(\mathcal{E}\right)$ the smallest σ algebra of sets of $[-\infty, \infty]^{\mathbb{Z}}$, which contains $\mathcal{E}$, the algebra of sets described in the proof of Kolmogorov's theorem. We give conditions next which imply that the mapping, T, just defined satisfies the conditions of Theorem 5.51.

Definition 8.19 *We say the sequence of random variables, $\{X_n\}_{n\in\mathbb{Z}}$ is stationary if*

$$\lambda_{(X_n,\cdots,X_{n+p})} = \lambda_{(X_0,\cdots,X_p)}$$

for every $n \in \mathbb{Z}$ and $p \geq 0$.

Theorem 8.20 *Let $\{X_n\}$ be stationary and let T be the shift operator just described. Also let λ be the measure of the Kolmogorov existence theorem defined on $\sigma\left(\mathcal{E}\right)$, the smallest σ algebra of sets of $[-\infty, \infty]^{\mathbb{Z}}$ containing $\mathcal{E}$. Then T is one to one and*

$$T^{-1}\mathbf{A}, T\mathbf{A} \in \sigma\left(\mathcal{E}\right) \ \text{whenever } \mathbf{A} \in \sigma\left(\mathcal{E}\right)$$

$$\lambda\left(T\mathbf{A}\right) = \lambda\left(T^{-1}\mathbf{A}\right) = \lambda\left(\mathbf{A}\right) \ \text{for all } \mathbf{A} \in \sigma\left(\mathcal{E}\right)$$

and T is ergodic.

Proof: It is obvious that T is one to one. We need to show T and T^{-1} map $\sigma(\mathcal{E})$ to $\sigma(\mathcal{E})$. It is clear both these maps take $\mathcal{E}$ to $\mathcal{E}$. Let $\mathcal{F} \equiv \left\{\mathbf{E} \in \sigma(\mathcal{E}) : T^{-1}(\mathbf{E}) \in \sigma(\mathcal{E})\right\}$ then $\mathcal{F}$ is a σ algebra and so it equals $\sigma(\mathcal{E})$. If $\mathbf{E} \in \mathcal{R}$ with $\gamma_J \mathbf{E} = \mathbf{E}$, and $J = \{n, n+1, \cdots, n+p\}$, then

$$\pi_J \mathbf{E} = \prod_{j=n}^{n+p} E_j,\ E_k = [-\infty, \infty],\ k \notin J.$$

Then

$$T\mathbf{E} = \mathbf{F} = \prod_{n \in \mathbb{Z}} F_n$$

where $F_{k-1} = E_k$. Therefore,

$$\lambda(\mathbf{E}) = \lambda_{(X_n, \cdots, X_{n+p})}(E_n \times \cdots \times E_{n+p}),$$

and

$$\lambda(T\mathbf{E}) = \lambda_{(X_{n-1}, \cdots, X_{n+p-1})}(E_n \times \cdots \times E_{n+p}).$$

By the assumption these random variables are stationary, $\lambda(T\mathbf{E}) = \lambda(\mathbf{E})$. A similar assertion holds for T^{-1}. It follows by a monotone class argument that this holds for all $\mathbf{E} \in \sigma(\mathcal{E})$. It is routine to verify T is ergodic. This proves the theorem.

Next we give an interesting lemma which we use in what follows.

Lemma 8.21 *In the context of Theorem 8.20 suppose* $\mathbf{A} \in \sigma(\mathcal{E})$ *and* $\lambda(\mathbf{A}) = 1$. *Then there exists* $\Sigma \in \mathcal{F}$ *such that* $P(\Sigma) = 1$ *and*

$$\Sigma = \left\{\omega \in \Omega : \prod_{k=-\infty}^{\infty} X_k(\omega) \in \mathbf{A}\right\}$$

Proof: Let $J_n \equiv \{-n, \cdots, n\}$ and let $\Sigma_n \equiv \{\omega : \mathbf{X}_{J_n}(\omega) \in \pi_{J_n}(\mathbf{A})\}$. Here

$$\mathbf{X}_{J_n}(\omega) \equiv (X_{-n}(\omega), \cdots, X_n(\omega)).$$

Then from the definition of π_{J_n}, we see that

$$\mathbf{A} = \cap_{n=1}^{\infty} \pi_{J_n}^{-1}(\pi_{J_n}\mathbf{A})$$

and the sets, $\pi_{J_n}^{-1}(\pi_{J_n}\mathbf{A})$ are decreasing in n. Let $\Sigma \equiv \cap_n \Sigma_n$. Then if $\omega \in \Sigma$,

$$\prod_{k=-n}^{n} X_k(\omega) \in \pi_{J_n}\mathbf{A}$$

for all n and so for each n, we have

$$\prod_{k=-\infty}^{\infty} X_k(\omega) \in \pi_{J_n}^{-1}(\pi_{J_n}\mathbf{A})$$

and consequently,

$$\prod_{k=-\infty}^{\infty} X_k(\omega) \in \cap_n \pi_{J_n}^{-1}(\pi_{J_n}\mathbf{A}) = \mathbf{A}$$

showing that

$$\Sigma \subseteq \left\{ \omega \in \Omega : \prod_{k=-\infty}^{\infty} X_k(\omega) \in \mathbf{A} \right\}$$

Now suppose $\prod_{k=-\infty}^{\infty} X_k(\omega) \in \mathbf{A}$. We need to verify that $\omega \in \Sigma$. We know

$$\prod_{k=-n}^{n} X_k(\omega) \in \pi_{J_n} \mathbf{A}$$

for each n and so $\omega \in \Sigma_n$ for each n. Therefore, $\omega \in \cap_{n=1}^{\infty} \Sigma_n$. This proves the lemma.

The following theorem is the main result.

Theorem 8.22 *Let $\{X_i\}_{i\in\mathbb{Z}}$ be stationary and let $J = \{0, \cdots, p\}$. Then if $f \circ \pi_J$ is a function in $L^1(\lambda)$, it follows that*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} f\left(X_{n+i-1}(\omega), \cdots, X_{n+p+i-1}(\omega)\right) = m$$

in $L^1(P)$ where

$$m \equiv \int f(X_0, \cdots, X_p)\, dP$$

and for a.e. $\omega \in \Omega$.

Proof: Let $f \in L^1(\lambda)$ where f is $\sigma(\mathcal{E})$ measurable. Then by Theorem 5.51 it follows that

$$\frac{1}{n} S_n f \equiv \frac{1}{n} \sum_{k=1}^{n} f\left(T^{k-1}(\cdot)\right) \to \int f(\mathbf{x})\, d\lambda \equiv m$$

pointwise λ a.e. and in $L^1(\lambda)$.

Now suppose f is of the form, $f(\mathbf{x}) = f(\pi_J(\mathbf{x}))$ where J is a finite subset of $\mathbb{Z}$,

$$J = \{n, n+1, \cdots, n+p\}.$$

Thus,

$$m = \int f(\pi_J \mathbf{x})\, d\lambda_{(X_n, \cdots, X_{n+p})} = \int f(X_n, \cdots, X_{n+p})\, dP.$$

(To verify the equality between the two integrals, verify for simple functions and then take limits.) Now

$$\begin{aligned} \frac{1}{k} S_k f(\mathbf{x}) &= \frac{1}{k} \sum_{i=1}^{k} f\left(T^{i-1}(\mathbf{x})\right) \\ &= \frac{1}{k} \sum_{i=1}^{k} f\left(\pi_J\left(T^{i-1}(\mathbf{x})\right)\right) \\ &= \frac{1}{k} \sum_{i=1}^{k} f\left(x_{n+i-1}, \cdots, x_{n+p+i-1}\right). \end{aligned}$$

Also, by the assumption that the sequence of random variables is stationary,

$$\int_\Omega \left| \frac{1}{k} S_k f\left(X_n(\omega), \cdots, X_{n+p}(\omega)\right) - m \right| dP =$$

$$\int_\Omega \left| \frac{1}{k} \sum_{i=1}^{k} f\left(X_{n+i-1}(\omega), \cdots, X_{n+p+i-1}(\omega)\right) - m \right| dP =$$

$$\int_{\mathbb{R}^J} \left| \frac{1}{k} \sum_{i=1}^{k} f\left(x_{n+i-1}, \cdots, x_{n+p+i-1}\right) - m \right| d\lambda_{(X_n, \cdots, X_{n+p})} =$$

$$= \int \left| \frac{1}{k} S_k f\left(\pi_J(\cdot)\right) - m \right| d\lambda = \int \left| \frac{1}{k} S_k f(\cdot) - m \right| d\lambda$$

and this last expression converges to 0 as $k \to \infty$ from the above.

By the individual ergodic theorem, we know that

$$\frac{1}{k} S_k f(\mathbf{x}) \to m$$

pointwise a.e. with respect to λ, say for $\mathbf{x} \in \mathbf{A}$ where $\lambda(\mathbf{A}) = 1$. Now by Lemma 8.21, there exists a set of measure 1 in $\mathcal{F}$, Σ, such that for all $\omega \in \Sigma$, $\prod_{k=-\infty}^{\infty} X_k(\omega) \in \mathbf{A}$. Therefore, for such ω,

$$\frac{1}{k} S_k f\left(\pi_J(\mathbf{X}(\omega))\right) = \frac{1}{k} \sum_{i=1}^{k} f\left(X_{n+i-1}(\omega), \cdots, X_{n+p+i-1}(\omega)\right)$$

converges to m. This proves the theorem.

An important example of a situation in which the random variables are stationary is the case when they are identically distributed and independent.

Definition 8.23 *We say the random variables, $\{X_j\}_{j=-\infty}^{\infty}$ are independent if whenever $J = \{i_1, \cdots, i_m\}$ is a finite subset of I, and $\left\{E_{i_j}\right\}_{j=1}^{m}$ are Borel sets in $[-\infty, \infty]$,*

$$\lambda_{\mathbf{X}_J}\left(\prod_{j=1}^{m} E_{i_j}\right) = \prod_{j=1}^{n} \lambda_{X_{i_j}}\left(E_{i_j}\right).$$

We say the random variables, $\{X_j\}_{j=-\infty}^{\infty}$ are identically distributed if whenever $i, j \in \mathbb{Z}$ and $E \subseteq [-\infty, \infty]$ is a Borel set,

$$\lambda_{X_i}(E) = \lambda_{X_j}(E).$$

As a routine lemma we obtain the following.

Lemma 8.24 *Suppose $\{X_j\}_{j=-\infty}^{\infty}$ are independent and identically distributed. Then $\{X_j\}_{j=-\infty}^{\infty}$ are stationary.*

The following corollary is called the strong law of large numbers.

Corollary 8.25 *Suppose* $\{X_j\}_{j=-\infty}^{\infty}$ *are independent and identically distributed random variables which are in* $L^1(\Omega, P)$. *Then*

$$\lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} X_{i-1}(\cdot) = m \equiv \int_{\mathbb{R}} X_0(\omega)\, dP \tag{8.5}$$

a.e. and in $L^1(P)$.

Proof: Let

$$f(\mathbf{x}) = f(\pi_0(\mathbf{x})) = x_0$$

so $f(\mathbf{x}) = \pi_0(\mathbf{x})$.

$$\int |f(\mathbf{x})|\, d\lambda = \int_{\mathbb{R}} |x|\, d\lambda_{X_0} = \int_{\mathbb{R}} |X_0(\omega)|\, dP < \infty.$$

Therefore, from the above strong ergodic theorem, we obtain (8.5).

8.3 Exercises

1. Let $\mathcal{A}$ be an algebra of sets in $\mathcal{P}(Z)$ and suppose μ and ν are two finite measures on $\sigma(\mathcal{A})$, the σ-algebra generated by $\mathcal{A}$. Show that if $\mu = \nu$ on $\mathcal{A}$, then $\mu = \nu$ on $\sigma(\mathcal{A})$.

2. $\uparrow$ Extend Problem 1 to the case where μ, ν are σ finite with

$$Z = \cup_{n=1}^{\infty} Z_n,\ Z_n \in \mathcal{A}$$

and $\mu(Z_n) < \infty$.

3. Show $\lim_{A\to\infty} \int_0^A \frac{\sin x}{x} dx = \frac{\pi}{2}$. **Hint:** Use $\frac{1}{x} = \int_0^\infty e^{-xt} dt$ and Fubini's theorem. This limit is sometimes called the Cauchy principle value. Note that the function $\sin(x)/x$ is not in L^1 so we are not finding a Lebesgue integral.

4. Suppose $g : \mathbb{R}^n \to \mathbb{R}$ has the property that g is continuous in each variable. Can we conclude that g is continuous? **Hint:** Consider

$$g(x, y) \equiv \begin{cases} \frac{xy}{x^2+y^2} \text{ if } (x, y) \neq (0, 0), \\ 0 \text{ if } (x, y) = (0, 0). \end{cases}$$

5. Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is continuous in every variable. Show that g is the pointwise limit of some sequence of continuous functions. Conclude that if g is continuous in each variable, then g is Borel measurable. Give an example of a Borel measurable function on $\mathbb{R}^n$ which is not continuous in each variable. **Hint:** In the case of $n = 2$ let

$$a_i \equiv \frac{i}{n}, i \in \mathbb{Z}$$

and for $(x, y) \in [a_{i-1}, a_i) \times \mathbb{R}$, we let

$$g_n(x, y) \equiv \frac{a_i - x}{a_i - a_{i-1}} g(a_{i-1}, y) + \frac{x - a_{i-1}}{a_i - a_{i-1}} g(a_i, y).$$

Show g_n converges to g and is continuous. Now use induction to verify the general result.

6. Show $(\mathbb{R}^2, m \times m, \mathcal{S} \times \mathcal{S})$ where $\mathcal{S}$ is the set of Lebesgue measurable sets is not a complete measure space. Show there exists $A \in \mathcal{S} \times \mathcal{S}$ and $E \subseteq A$ such that $(m \times m)(A) = 0$, but $E \notin \mathcal{S} \times \mathcal{S}$.

7. Recall that for
$$E \in \mathcal{S} \times \mathcal{F}, (\mu \times \lambda)(E) = \int_X \lambda(E_x) d\mu = \int_Y \mu(E^y) d\lambda.$$
Why is $\mu \times \lambda$ a measure on $\mathcal{S} \times \mathcal{F}$?

8. Suppose $G(x) = G(a) + \int_a^x g(t)\,dt$ where $g \in L^1$ and suppose $F(x) = F(a) + \int_a^x f(t)\,dt$ where $f \in L^1$. Show the usual formula for integration by parts holds,
$$\int_a^b fG dx = FG|_a^b - \int_a^b Fg dx.$$
Hint: You might try replacing $G(x)$ with $G(a) + \int_a^x g(t)\,dt$ in the first integral on the left and then using Fubini's theorem.

9. Let $f : \Omega \to [0, \infty)$ be measurable where $(\Omega, \mathcal{S}, \mu)$ is a σ finite measure space. Let $\phi : [0, \infty) \to [0, \infty)$ satisfy: ϕ is increasing. Show
$$\int_X \phi(f(x)) d\mu = \int_0^\infty \phi'(t) \mu(x : f(x) > t) dt.$$
The function $t \to \mu(x : f(x) > t)$ is called the distribution function. **Hint:**
$$\int_X \phi(f(x)) d\mu = \int_X \int_{\mathbb{R}} \mathcal{X}_{[0, f(x))} \phi'(t) dt dx.$$
Now try to use Fubini's theorem. Be sure to check that everything is appropriately measurable. In doing so, you may want to first consider $f(x)$ a nonnegative simple function. Is it necessary to assume $(\Omega, \mathcal{S}, \mu)$ is σ finite?

10. In the Kolmogorov extension theorem, could X_t be a random vector with values in an arbitrary locally compact Hausdorff space?

11. Can you generalize the strong ergodic theorem to the case where the random variables have values not in $\mathbb{R}$ but $\mathbb{R}^k$ for some k?

Fourier Series

9.1 Definition and basic properties

A Fourier series is an expression of the form

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

where we understand this symbol to mean

$$\lim_{n\to\infty} \sum_{k=-n}^{n} c_k e^{ikx}.$$

Obviously such a sequence of partial sums may or may not converge at a particular value of x.

These series have been important in applied math since the time of Fourier who was an officer in Napolean's army. He was interested in studying the flow of heat in cannons and invented the concept to aid him in his study. Since that time, Fourier series and the mathematical problems related to their convergence have motivated the development of modern methods in analysis. As recently as the mid 1960's a problem related to convergence of Fourier series was solved for the first time and the solution of this problem was a big surprise. In this chapter, we will discuss the classical theory of convergence of Fourier series. We will use standard notation for the integral also. Thus,

$$\int_a^b f(x)\,dx \equiv \int \mathcal{X}_{[a,b]}(x) f(x)\,dm.$$

Definition 9.1 *We say a function, f defined on $\mathbb{R}$ is a periodic function of period T if $f(x+T) = f(x)$ for all x.*

Fourier series are useful for representing periodic functions and no other kind of function. To see this, note that the partial sums are periodic of period 2π. Therefore, it is not reasonable to expect to represent a function which is not periodic of period 2π on $\mathbb{R}$ by a series of the above form. However, if we are interested in periodic functions, there is no loss of generality in studying only functions which are periodic of period 2π. Indeed, if f is a function which has period T, we can study this function in terms of the function, $g(x) \equiv f\left(\frac{Tx}{2\pi}\right)$ and g is periodic of period 2π.

Definition 9.2 *For $f \in L^1(-\pi,\pi)$ and f periodic on $\mathbb{R}$, we define the Fourier series of f as*

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{9.1}$$

where

$$c_k \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iky} dy. \tag{9.2}$$

We also define the nth partial sum of the Fourier series of f by

$$S_n(f)(x) \equiv \sum_{k=-n}^{n} c_k e^{ikx}. \tag{9.3}$$

It may be interesting to see where this formula came from. Suppose then that

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

and we multiply both sides by e^{-imx} and take the integral, $\int_{-\pi}^{\pi}$, so that

$$\int_{-\pi}^{\pi} f(x) e^{-imx} dx = \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} c_k e^{ikx} e^{-imx} dx.$$

Now we switch the sum and the integral on the right side even though we have absolutely no reason to believe this makes any sense. Then we get

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) e^{-imx} dx &= \sum_{k=-\infty}^{\infty} c_k \int_{-\pi}^{\pi} e^{ikx} e^{-imx} dx \\ &= c_m \int_{-\pi}^{\pi} 1 dx = 2\pi c_k \end{aligned}$$

because $\int_{-\pi}^{\pi} e^{ikx} e^{-imx} dx = 0$ if $k \neq m$. It is formal manipulations of the sort just presented which suggest that Definition 9.2 might be interesting.

In case f is real valued, we see that $\overline{c_k} = c_{-k}$ and so

$$S_n f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) dy + \sum_{k=1}^{n} 2 \operatorname{Re}\left(c_k e^{ikx}\right).$$

Letting $c_k \equiv \alpha_k + i\beta_k$

$$S_n f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) dy + \sum_{k=1}^{n} 2 \left[\alpha_k \cos kx - \beta_k \sin kx\right]$$

where

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iky} dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) (\cos ky - i \sin ky) dy$$

which shows that

$$\alpha_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) \cos(ky) dy, \ \beta_k = \frac{-1}{2\pi} \int_{-\pi}^{\pi} f(y) \sin(ky) dy$$

Therefore, letting $a_k = 2\alpha_k$ and $b_k = -2\beta_k$, we see that

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ky) dy, \ b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ky) dy$$

and

$$S_n f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx \tag{9.4}$$

This is often the way Fourier series are presented in elementary courses where it is only real functions which are to be approximated. We will stick with Definition 9.2 because it is easier to use.

The partial sums of a Fourier series can be written in a particularly simple form which we present next.

$$\begin{aligned} S_n f(x) &= \sum_{k=-n}^{n} c_k e^{ikx} \\ &= \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iky} dy \right) e^{ikx} \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{k=-n}^{n} \left(e^{ik(x-y)} \right) f(y)\, dy \\ &\equiv \int_{-\pi}^{\pi} D_n(x-y) f(y)\, dy. \end{aligned} \tag{9.5}$$

The function,

$$D_n(t) \equiv \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikt}$$

is called the Dirichlet Kernel

Theorem 9.3 *The function, D_n satisfies the following:*

1. $\int_{-\pi}^{\pi} D_n(t)\, dt = 1$
2. *D_n is periodic of period 2π*
3. $D_n(t) = (2\pi)^{-1} \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\left(\frac{t}{2}\right)}$.

Proof: Part 1 is obvious because $\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iky} dy = 0$ whenever $k \neq 0$ and it equals 1 if $k = 0$. Part 2 is also obvious because $t \to e^{ikt}$ is periodic of period 2π. It remains to verify Part 3.

$$2\pi D_n(t) = \sum_{k=-n}^{n} e^{ikt},\ 2\pi e^{it} D_n(t) = \sum_{k=-n}^{n} e^{i(k+1)t} = \sum_{k=-n+1}^{n+1} e^{ikt}$$

and so subtracting we get

$$2\pi D_n(t)\left(1 - e^{it}\right) = e^{-int} - e^{i(n+1)t}.$$

Therefore,

$$2\pi D_n(t)\left(e^{-it/2} - e^{it/2}\right) = e^{-i\left(n+\frac{1}{2}\right)t} - e^{i\left(n+\frac{1}{2}\right)t}$$

and so

$$2\pi D_n(t)\left(-2i \sin\frac{t}{2}\right) = -2i \sin\left(n + \frac{1}{2}\right) t$$

which establishes Part 3. This proves the theorem.

It is not reasonable to expect a Fourier series to converge to the function at every point. To see this, change the value of the function at a single point in $(-\pi, \pi)$ and extend to keep the modified function periodic. Then the Fourier series of the modified function is the same as the Fourier series of the original function and so if pointwise convergence did take place, it no longer does. However, it is possible to prove an interesting theorem about pointwise convergence of Fourier series. This is done next.

9.2 Pointwise convergence of Fourier series

The following lemma will make possible a very easy proof of the very important Riemann Lebesgue lemma, the big result which makes possible the study of pointwise convergence of Fourier series.

Lemma 9.4 *Let $f \in L^1(a,b)$ where (a,b) is some interval, possibly an infinite interval like $(-\infty,\infty)$ and let $\varepsilon > 0$. Then there exists an interval of finite length, $[a_1,b_1] \subseteq (a,b)$ and a function, $g \in C_c(a_1,b_1)$ such that also $g' \in C_c(a_1,b_1)$ which has the property that*

$$\int_a^b |f-g|\,dx < \varepsilon \tag{9.6}$$

Proof: Without loss of generality we may assume that $f(x) \geq 0$ for all x since we can always consider the positive and negative parts of the real and imaginary parts of f. Letting $a < a_n < b_n < b$ with $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$, we may use the dominated convergence theorem to conclude

$$\lim_{n\to\infty} \int_a^b \left| f(x) - f(x)\mathcal{X}_{[a_n,b_n]}(x) \right| dx = 0.$$

Therefore, there exist $c > a$ and $d < b$ such that if $h = f\mathcal{X}_{(c,d)}$, then

$$\int_a^b |f-h|\,dx < \frac{\varepsilon}{4}. \tag{9.7}$$

Now from Theorem 5.31 on the pointwise convergence of nonnegative simple functions to nonnegative measurable functions and the monotone convergence theorem, there exists a simple function,

$$s(x) = \sum_{i=1}^{p} c_i \mathcal{X}_{E_i}(x)$$

with the property that

$$\int_a^b |s-h|\,dx < \frac{\varepsilon}{4} \tag{9.8}$$

and we may assume each $E_i \subseteq (c,d)$. Now by regularity of Lebesgue measure, there are compact sets, K_i and open sets, V_i such that $K_i \subseteq E_i \subseteq V_i$, $V_i \subseteq (c,d)$, and $m(V_i \setminus K_i) < \alpha$, where α is a positive number which is arbitrary. Now we let $K_i \prec k_i \prec V_i$. Thus k_i is a continuous function whose support is contained in V_i, which is nonnegative, and equal to one on K_i. Then let

$$g_1(x) \equiv \sum_{i=1}^{p} c_i k_i(x).$$

We see that g_1 is a continuous function whose support is contained in the compact set, $[c,d] \subseteq (a,b)$. Also,

$$\int_a^b |g_1 - s|\,dx \leq \sum_{i=1}^{p} c_i m(V_i \setminus K_i) < \alpha \sum_{i=1}^{p} c_i.$$

Choosing α small enough, we may assume

$$\int_a^b |g_1 - s|\, dx < \frac{\varepsilon}{4}. \tag{9.9}$$

Now choose r small enough that $[c-r, d+r] \subseteq (a,b)$ and for $0 < h < r$, let

$$g_h(x) \equiv \frac{1}{2h} \int_{x-h}^{x+h} g_1(t)\, dt.$$

Then g_h is a continuous function whose derivative is also continuous and for which both g_h and g_h' have support in $[c-r, d+r] \subseteq (a,b)$. Now let $[a_1, b_1] \equiv [c-r, d+r]$.

$$\begin{aligned} \int_a^b |g_1 - g_h|\, dx &\leq \int_{a_1}^{b_1} \frac{1}{2h} \int_{x-h}^{x+h} |g_1(x) - g_1(t)|\, dt dx \\ &< \frac{\varepsilon}{4(b_1 - a_1)} (b_1 - a_1) = \frac{\varepsilon}{4} \end{aligned} \tag{9.10}$$

whenever h is small enough due to the uniform continuity of g_1. Let $g = g_h$ for such h. Using (9.7) - (9.10), we obtain

$$\begin{aligned} \int_a^b |f - g|\, dx &\leq \int_a^b |f - h|\, dx + \int_a^b |h - s|\, dx + \int_a^b |s - g_1|\, dx \\ + \int_a^b |g_1 - g|\, dx &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

This proves the lemma.

With this lemma, we are ready to prove the Riemann Lebesgue lemma.

Lemma 9.5 *(Riemann Lebesgue) Let $f \in L^1(a,b)$ where (a,b) is some interval, possibly an infinite interval like $(-\infty, \infty)$. Then*

$$\lim_{\alpha \to \infty} \int_a^b f(t) \sin(\alpha t + \beta)\, dt = 0. \tag{9.11}$$

Proof: Let $\varepsilon > 0$ be given and use Lemma 9.4 to obtain g such that g and g' are both continuous, the support of both g and $g'$are contained in $[a_1, b_1]$, and

$$\int_a^b |g - f|\, dx < \frac{\varepsilon}{2}. \tag{9.12}$$

Then

$$\begin{aligned} \left| \int_a^b f(t) \sin(\alpha t + \beta)\, dt \right| &\leq \left| \int_a^b f(t) \sin(\alpha t + \beta)\, dt - \int_a^b g(t) \sin(\alpha t + \beta)\, dt \right| \\ &\quad + \left| \int_a^b g(t) \sin(\alpha t + \beta)\, dt \right| \\ &\leq \int_a^b |f - g|\, dx + \left| \int_a^b g(t) \sin(\alpha t + \beta)\, dt \right| \\ &< \frac{\varepsilon}{2} + \left| \int_{a_1}^{b_1} g(t) \sin(\alpha t + \beta)\, dt \right|. \end{aligned}$$

We integrate the last term by parts obtaining

$$\int_{a_1}^{b_1} g(t)\sin(\alpha t+\beta)\,dt = \frac{-\cos(\alpha t+\beta)}{\alpha} g(t)\,|_{a_1}^{b_1} + \int_{a_1}^{b_1} \frac{\cos(\alpha t+\beta)}{\alpha} g'(t)\,dt,$$

an expression which converges to zero since g' is bounded. Therefore, taking α large enough, we see

$$\left|\int_a^b f(t)\sin(\alpha t+\beta)\,dt\right| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

and this proves the lemma.

9.2.1 Dini's criterion

Now we are ready to prove a theorem about the convergence of the Fourier series to the midpoint of the jump of the function. The condition given for convergence in the following theorem is due to Dini. [3].

Theorem 9.6 *Let f be a periodic function of period 2π which is in $L^1(-\pi,\pi)$. Suppose at some x, we have $f(x+)$ and $f(x-)$ both exist and that the function*

$$y \to \frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{y} \tag{9.13}$$

is in $L^1(0,\delta)$ for some $\delta>0$. Then

$$\lim_{n\to\infty} S_n f(x) = \frac{f(x+)+f(x-)}{2}. \tag{9.14}$$

Proof:

$$S_n f(x) = \int_{-\pi}^{\pi} D_n(x-y) f(y)\,dy$$

We change variables $x-y\to y$ and use the periodicity of f and D_n to write this as

$$\begin{aligned}
S_n f(x) &= \int_{-\pi}^{\pi} D_n(y) f(x-y) \\
&= \int_0^{\pi} D_n(y) f(x-y)\,dy + \int_{-\pi}^{0} D_n(y) f(x-y)\,dy \\
&= \int_0^{\pi} D_n(y)\left[f(x-y)+f(x+y)\right]dy \\
&= \int_0^{\pi} \frac{1}{\pi}\frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{\sin\left(\frac{y}{2}\right)}\left[\frac{f(x-y)+f(x+y)}{2}\right]dy. \qquad (9.15)
\end{aligned}$$

Also

$$\begin{aligned}
f(x+)+f(x-) &= \int_{-\pi}^{\pi} D_n(y)\left[f(x+)+f(x-)\right]dy \\
&= 2\int_0^{\pi} D_n(y)\left[f(x+)+f(x-)\right]dy \\
&= \int_0^{\pi} \frac{1}{\pi}\frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{\sin\left(\frac{y}{2}\right)}\left[f(x+)+f(x-)\right]dy
\end{aligned}$$

and so

$$\left|S_n f(x) - \frac{f(x+) + f(x-)}{2}\right| =$$

$$\left|\int_0^{\pi} \frac{1}{\pi} \frac{\sin\left(\left(n + \frac{1}{2}\right) y\right)}{\sin\left(\frac{y}{2}\right)} \left[\frac{f(x-y) - f(x-) + f(x+y) - f(x+)}{2}\right] dy\right|. \tag{9.16}$$

Now the function

$$y \to \frac{f(x-y) - f(x-) + f(x+y) - f(x+)}{2\sin\left(\frac{y}{2}\right)} \tag{9.17}$$

is in $L^1(0,\pi)$. To see this, note the numerator is in L^1 because f is. Therefore, this function is in $L^1(\delta,\pi)$ where δ is given in the Hypotheses of this theorem because $\sin\left(\frac{y}{2}\right)$ is bounded below by $\sin\left(\frac{\delta}{2}\right)$ for such y. The function is in $L^1(0,\delta)$ because

$$\frac{f(x-y) - f(x-) + f(x+y) - f(x+)}{2\sin\left(\frac{y}{2}\right)} = \frac{f(x-y) - f(x-) + f(x+y) - f(x+)}{y} \frac{y}{2\sin\left(\frac{y}{2}\right)},$$

and $y/2\sin\left(\frac{y}{2}\right)$ is bounded on $[0,\delta]$. Thus the function in (9.17) is in $L^1(0,\pi)$ as claimed. It follows from the Riemann Lebesgue lemma, that (9.16) converges to zero as $n \to \infty$. This proves the theorem.

The following corollary is obtained immediately from the above proof with minor modifications.

Corollary 9.7 *Let f be a periodic function of period 2π which is in $L^1(-\pi,\pi)$. Suppose at some x, we have the function*

$$y \to \frac{f(x-y) + f(x+y) - 2s}{y} \tag{9.18}$$

is in $L^1(0,\delta)$ for some $\delta > 0$. Then

$$\lim_{n\to\infty} S_n f(x) = s. \tag{9.19}$$

As pointed out by Apostol, [3], this is a very remarkable result because even though the Fourier coeficients depend on the values of the function on all of $[-\pi,\pi]$, the convergence properties depend in this theorem on very local behavior of the function. Dini's condition is rather like a very weak smoothness condition on f at the point, x. Indeed, in elementary treatments of Fourier series, the assumption given is that the one sided derivatives of the function exist. The following corollary gives an easy to check condition for the Fourier series to converge to the mid point of the jump.

Corollary 9.8 *Let f be a periodic function of period 2π which is in $L^1(-\pi,\pi)$. Suppose at some x, we have $f(x+)$ and $f(x-)$ both exist and there exist positive constants, K and δ such that whenever $0 < y < \delta$*

$$|f(x-y) - f(x-)| \le Ky^{\theta},\ |f(x+y) - f(x+)| < Ky^{\theta} \tag{9.20}$$

where $\theta \in (0,1]$. Then

$$\lim_{n\to\infty} S_n f(x) = \frac{f(x+) + f(x-)}{2}. \tag{9.21}$$

Proof: The condition (9.20) clearly implies Dini's condition, (9.13).

9.2.2 Jordan's criterion

We give a different condition under which the Fourier series converges to the mid point of the jump. In order to prove the theorem, we need to give some introductory lemmas which are interesting for their own sake.

Lemma 9.9 *Let G be an increasing function defined on $[a,b]$. Thus $G(x) \leq G(y)$ whenever $x < y$. Then $G(x-) = G(x+) = G(x)$ for every x except for a countable set of exceptions.*

Proof: We let $S \equiv \{x \in [a,b] : G(x+) > G(x-)\}$. Then there is a rational number in each interval, $(G(x-), G(x+))$ and also, since G is increasing, these intervals are disjoint. It follows that there are only contably many such intervals. Therefore, S is countable and if $x \notin S$, we have $G(x+) = G(x-)$ showing that G is continuous on S^C and the claimed equality holds.

The next lemma is called the second mean value theorem for integrals.

Lemma 9.10 *Let G be an increasing function defined on $[a,b]$ and let f be a continuous function defined on $[a,b]$. Then there exists $t_0 \in [a,b]$ such that*

$$\int_a^b G(s) f(s)\, ds = G(a) \left(\int_a^{t_0} f(s)\, ds \right) + G(b-) \left(\int_{t_0}^b f(s)\, ds \right). \tag{9.22}$$

Proof: Letting $h > 0$ we define

$$G_h(t) \equiv \frac{1}{h^2} \int_{t-h}^t \int_{s-h}^s G(r)\, dr ds$$

where we understand $G(x) = G(a)$ for all $x < a$. Thus $G_h(a) = G(a)$ for all h. Also, from the fundamental theorem of calculus, we see that $G_h'(t) \geq 0$ and is a continuous function of t. Also it is clear that $\lim_{h\to\infty} G_h(t) = G(t-)$ for all $t \in [a,b]$. Letting $F(t) \equiv \int_a^t f(s)\, ds$,

$$\int_a^b G_h(s) f(s)\, ds = F(t) G_h(t)|_a^b - \int_a^b F(t) G_h'(t)\, dt. \tag{9.23}$$

Now letting $m = \min\{F(t) : t \in [a,b]\}$ and $M = \max\{F(t) : t \in [a,b]\}$, since $G_h'(t) \geq 0$, we have

$$m \int_a^b G_h'(t)\, dt \leq \int_a^b F(t) G_h'(t)\, dt \leq M \int_a^b G_h'(t)\, dt.$$

Therefore, if $\int_a^b G_h'(t)\, dt \neq 0$,

$$m \leq \frac{\int_a^b F(t) G_h'(t)\, dt}{\int_a^b G_h'(t)\, dt} \leq M$$

and so by the intermediate value theorem from calculus,

$$F(t_h) = \frac{\int_a^b F(t) G_h'(t)\, dt}{\int_a^b G_h'(t)\, dt}$$

for some $t_h \in [a,b]$. Therefore, substituting for $\int_a^b F(t) G_h'(t)\, dt$ in (9.23) we have

$$\begin{aligned}
\int_a^b G_h(s) f(s)\, ds &= F(t) G_h(t)|_a^b - \left[F(t_h) \int_a^b G_h'(t)\, dt \right] \\
&= F(b) G_h(b) - F(t_h) G_h(b) + F(t_h) G_h(a) \\
&= \left(\int_{t_h}^b f(s)\, ds \right) G_h(b) + \left(\int_a^{t_h} f(s)\, ds \right) G(a).
\end{aligned}$$

Now selecting a subsequence, still denoted by h which converges to zero, we can assume $t_h \to t_0 \in [a,b]$. Therefore, using the dominated convergence theorem, we may obtain the following from the above lemma.

$$\begin{aligned} \int_a^b G(s) f(s)\, ds &= \int_a^b G(s-) f(s)\, ds \\ &= \left(\int_{t_0}^b f(s)\, ds\right) G(b-) + \left(\int_a^{t_0} f(s)\, ds\right) G(a). \end{aligned}$$

and this proves the lemma.

Definition 9.11 *Let $f : [a,b] \to \mathbb{C}$ be a function. We say f is of bounded variation if*

$$\sup\left\{\sum_{i=1}^n |f(t_i) - f(t_{i-1})| : a = t_0 < \cdots < t_n = b\right\} \equiv V(f,[a,b]) < \infty$$

where the sums are taken over all possible lists, $\{a = t_0 < \cdots < t_n = b\}$. The symbol, $V(f,[a,b])$ is known as the total variation on $[a,b]$.

Lemma 9.12 *A real valued function, f, defined on an interval, $[a,b]$ is of bounded variation if and only if there are increasing functions, H and G defined on $[a,b]$ such that $f(t) = H(t) - G(t)$. A complex valued function is of bounded variation if and only if the real and imaginary parts are of bounded variation.*

Proof: For f a real valued function of bounded variation, we may define an increasing function, $H(t) \equiv V(f,[a,t])$ and then note that

$$f(t) = H(t) - \overbrace{[H(t) - f(t)]}^{G(t)}.$$

It is routine to verify that $G(t)$ is increasing. Conversely, if $f(t) = H(t) - G(t)$ where H and G are increasing, we can easily see the total variation for H is just $H(b) - H(a)$ and the total variation for G is $G(b) - G(a)$. Therefore, the total variation for f is bounded by the sum of these.

The last claim follows from the observation that

$$|f(t_i) - f(t_{i-1})| \geq \max\left(|\operatorname{Re} f(t_i) - \operatorname{Re} f(t_{i-1})|, |\operatorname{Im} f(t_i) - \operatorname{Im} f(t_{i-1})|\right)$$

and

$$|\operatorname{Re} f(t_i) - \operatorname{Re} f(t_{i-1})| + |\operatorname{Im} f(t_i) - \operatorname{Im} f(t_{i-1})| \geq |f(t_i) - f(t_{i-1})|.$$

With this lemma, we can now prove the Jordan criterion for pointwise convergence of the Fourier series.

Theorem 9.13 *Suppose f is 2π periodic and is in $L^1(-\pi,\pi)$. Suppose also that for some $\delta > 0$, f is of bounded variation on $[x-\delta, x+\delta]$. Then*

$$\lim_{n\to\infty} S_n f(x) = \frac{f(x+) + f(x-)}{2}. \tag{9.24}$$

Proof: First note that from Definition 9.11, $\lim_{y\to x-} \operatorname{Re} f(y)$ exists because $\operatorname{Re} f$ is the difference of two increasing functions. Similarly this limit will exist for $\operatorname{Im} f$ by the same reasoning and limits of the form $\lim_{y\to x+}$ will also exist. Therefore, the expression on the right in (9.24) exists. If we can verify (9.24) for real functions which are of bounded variation on $[x-\delta, x+\delta]$, we can apply this to the real and imaginary parts of f and obtain the desired result for f. Therefore, we assume without loss of generality that f is real valued and of bounded variation on $[x-\delta, x+\delta]$.

$$S_n f(x) - \frac{f(x+) + f(x-)}{2} = \int_{-\pi}^{\pi} D_n(y) \left(f(x-y) - \frac{f(x+) + f(x-)}{2} \right) dy$$

$$= \int_0^{\pi} D_n(y) \left[(f(x+y) - f(x+)) + (f(x-y) - f(x-)) \right] dy.$$

Now the Dirichlet kernel, $D_n(y)$ is a constant multiple of $\sin((n+1/2)y)/\sin(y/2)$ and so we can apply the Riemann Lebesgue lemma to conclude that

$$\lim_{n\to\infty} \int_{\delta}^{\pi} D_n(y) \left[(f(x+y) - f(x+)) + (f(x-y) - f(x-)) \right] dy = 0$$

and so it suffices to show that

$$\lim_{n\to\infty} \int_0^{\delta} D_n(y) \left[(f(x+y) - f(x+)) + (f(x-y) - f(x-)) \right] dy = 0. \tag{9.25}$$

Now $y \to (f(x+y) - f(x+)) + (f(x-y) - f(x-)) = h(y)$ is of bounded variation for $y \in [0, \delta]$ and $\lim_{y\to 0+} h(y) = 0$. Therefore, we can write $h(y) = H(y) - G(y)$ where H and G are increasing and for $F = G, H$, $\lim_{y\to 0+} F(y) = F(0) = 0$. It suffices to show (9.25) holds with f replaced with either of G or H.

Letting $\varepsilon > 0$ be given, we choose $\delta_1 < \delta$ such that $H(\delta_1), G(\delta_1) < \varepsilon$. Now

$$\int_0^{\delta} D_n(y) G(y) \, dy = \int_{\delta_1}^{\delta} D_n(y) G(y) \, dy + \int_0^{\delta_1} D_n(y) G(y) \, dy$$

and we see from the Riemann Lebesgue lemma that the first integral on the right converges to 0 for any choice of $\delta_1 \in (0, \delta)$. Therefore, we estimate the second integral on the right. Using the second mean value theorem, Lemma 9.10, we see there exists $\delta_n \in [0, \delta_1]$ such that

$$\begin{aligned} \left| \int_0^{\delta_1} D_n(y) G(y) \, dy \right| &= \left| G(\delta_1 -) \int_{\delta_n}^{\delta_1} D_n(y) \, dy \right| \\ &\leq \varepsilon \left| \int_{\delta_n}^{\delta_1} D_n(y) \, dy \right|. \end{aligned}$$

Now

$$\left| \int_{\delta_n}^{\delta_1} D_n(y) \right| = C \left| \int_{\delta_n}^{\delta_1} \frac{y}{\sin(y/2)} \frac{\sin\left(n + \frac{1}{2}\right) y}{y} dt \right|$$

and for small $\delta_1, y/\sin(y/2)$ is approximately equal to 2. Therefore, the expression on the right will be bounded if we can show that

$$\left| \int_{\delta_n}^{\delta_1} \frac{\sin\left(n + \frac{1}{2}\right) y}{y} dt \right|$$

is bounded independent of choice of $\delta_n \in [0, \delta_1]$. Changing variables, we see this is equivalent to showing that

$$\left| \int_a^b \frac{\sin y}{y} dy \right|$$

is bounded independent of the choice of a, b. But this follows from the convergence of the Cauchy principle value integral given by

$$\lim_{A\to\infty}\int_0^A \frac{\sin y}{y}dy$$

which was considered in Problem 3 of Chapter 8 or Problem 12 of Chapter 7. Using the above argument for H as well as G, this shows that there exists a constant, C independent of ε such that

$$\limsup_{n\to\infty}\left|\int_0^\delta D_n(y)\left[(f(x+y)-f(x+))+(f(x-y)-f(x-))\right]dy\right|\le C\varepsilon.$$

Since ε was arbitrary, this proves the theorem.

It is known that neither the Jordan criterion nor the Dini criterion implies the other.

9.2.3 The Fourier cosine series

Suppose now that f is a **real valued** function which is defined on the interval $[0,\pi]$. Then we can define f on the interval, $[-\pi,\pi]$ according to the rule, $f(-x)=f(x)$. Thus the resulting function, still denoted by f is an even function. We can now extend this even function to the whole real line by requiring $f(x+2\pi)=f(x)$ obtaining a 2π periodic function. Note that if f is continuous, then this periodic function defined on the whole line is also continuous. What is the Fourier series of the extended function f? Since f is an even function, the n^{th} coefficient is of the form

$$\begin{aligned} c_n &\equiv \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx=\frac{1}{\pi}\int_0^\pi f(x)\cos(nx)\,dx \text{ if } n\neq 0 \\ c_0 &= \frac{1}{\pi}\int_0^\pi f(x)\,dx \text{ if } n=0. \end{aligned}$$

Thus $c_{-n}=c_n$ and we see the Fourier series of f is of the form

$$\frac{1}{\pi}\left(\int_0^\pi f(x)\,dx\right)+\sum_{k=1}^\infty\left(\frac{2}{\pi}\int_0^\pi f(y)\cos ky\right)\cos kx \tag{9.26}$$

$$=c_0+\sum_{k=1}^\infty 2c_k\cos kx. \tag{9.27}$$

Definition 9.14 *If f is a function defined on $[0,\pi]$ then (9.26) is called the Fourier cosine series of f.*

Observe that Fourier series of even 2π periodic functions yield Fourier cosine series.

We have the following corollary to Theorem 9.6 and Theorem 9.13.

Corollary 9.15 *Let f be an even function defined on $\mathbb{R}$ which has period 2π and is in $L^1(0,\pi)$. Then at every point, x, where $f(x+)$ and $f(x-)$ both exist and the function*

$$y\to\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{y} \tag{9.28}$$

is in $L^1(0,\delta)$ for some $\delta>0$, or for which f is of bounded variation near x we have

$$\lim_{n\to\infty} a_0+\sum_{k=1}^n a_k\cos kx=\frac{f(x+)+f(x-)}{2} \tag{9.29}$$

Here

$$a_0=\frac{1}{\pi}\int_0^\pi f(x)\,dx,\ a_n=\frac{2}{\pi}\int_0^\pi f(x)\cos(nx)\,dx. \tag{9.30}$$

There is another way of approximating periodic piecewise continuous functions as linear combinations of the functions e^{iky} which is clearly superior in terms of pointwise and uniform convergence. This other way does not depend on any hint of smoothness of f near the point in question.

9.3 The Cesaro means

In this section we define the notion of the Cesaro mean and show these converge to the midpoint of the jump under very general conditions.

Definition 9.16 *We define the nth Cesaro mean of a periodic function which is in $L^1(-\pi,\pi)$, $\sigma_n f(x)$ by the formula*

$$\sigma_n f(x) \equiv \frac{1}{n+1}\sum_{k=0}^{n} S_k f(x).$$

Thus the nth Cesaro mean is just the average of the first $n+1$ partial sums of the Fourier series.

Just as in the case of the $S_n f$, we can write the Cesaro means in terms of convolution of the function with a suitable kernel, known as the Fejer kernel. We want to find a formula for the Fejer kernel and obtain some of its properties. First we give a simple formula which follows from elementary trigonometry.

$$\begin{aligned}\sum_{k=0}^{n}\sin\left(k+\frac{1}{2}\right)y &= \frac{1}{2\sin\frac{y}{2}}\sum_{k=0}^{n}(\cos ky-\cos(k+1)y)\\ &= \frac{1-\cos((n+1)y)}{2\sin\frac{y}{2}}.\end{aligned}\tag{9.31}$$

Lemma 9.17 *There exists a unique function, $F_n(y)$ with the following properties.*

1. $\sigma_n f(x) = \int_{-\pi}^{\pi} F_n(x-y) f(y)\,dy$,
2. F_n *is periodic of period* 2π,
3. $F_n(y)\geq 0$ *and if* $\pi > |y| \geq r > 0$, *then* $\lim_{n\to\infty} F_n(y) = 0$,
4. $\int_{-\pi}^{\pi} F_n(y)\,dy = 1$,
5. $F_n(y) = \frac{1-\cos((n+1)y)}{4\pi(n+1)\sin^2\left(\frac{y}{2}\right)}$

Proof: From the definition of σ_n, it follows that

$$\sigma_n f(x) = \int_{-\pi}^{\pi}\left[\frac{1}{n+1}\sum_{k=0}^{n} D_k(x-y)\right] f(y)\,dy.$$

Therefore,

$$F_n(y) = \frac{1}{n+1}\sum_{k=0}^{n} D_k(y).\tag{9.32}$$

That F_n is periodic of period 2π follows from this formula and the fact, established earlier that D_k is periodic of period 2π. Thus we have established parts 1 and 2. Part 4 also follows immediately from the fact that

$\int_{-\pi}^{\pi} D_k(y)\,dy = 1$. We now establish Part 5 and Part 3. From (9.32) and (9.31),

$$
\begin{aligned}
F_n(y) &= \frac{1}{2\pi(n+1)}\frac{1}{\sin\left(\frac{y}{2}\right)}\sum_{k=0}^{n}\sin\left(\left(k+\frac{1}{2}\right)y\right) \\
&= \frac{1}{2\pi(n+1)\sin\left(\frac{y}{2}\right)}\left(\frac{1-\cos((n+1)y)}{2\sin\frac{y}{2}}\right) \\
&= \frac{1-\cos((n+1)y)}{4\pi(n+1)\sin^2\left(\frac{y}{2}\right)}. \qquad (9.33)
\end{aligned}
$$

This verifies Part 5 and also shows that $F_n(y) \geq 0$, the first part of Part 3. If $|y| > r$,

$$
|F_n(y)| \leq \frac{2}{4\pi(n+1)\sin^2\left(\frac{r}{2}\right)} \qquad (9.34)
$$

and so the second part of Part 3 holds. This proves the lemma.

The following theorem is called Fejer's theorem

Theorem 9.18 *Let f be a periodic function with period 2π which is in $L^1(-\pi,\pi)$. Then if $f(x+)$ and $f(x-)$ both exist,*

$$
\lim_{n\to\infty}\sigma_n f(x) = \frac{f(x+)+f(x-)}{2}. \qquad (9.35)
$$

If f is everywhere continuous, then $\sigma_n f$ converges uniformly to f on all of $\mathbb{R}$.

Proof: As before, we may use the periodicity of f and F_n to write

$$
\begin{aligned}
\sigma_n f(x) &= \int_0^{\pi} F_n(y)\left[f(x-y)+f(x+y)\right]dy \\
&= \int_0^{\pi} 2F_n(y)\left[\frac{f(x-y)+f(x+y)}{2}\right]dy.
\end{aligned}
$$

From the formula for F_n, we see that F_n is even and so $\int_0^{\pi} 2F_n(y)\,dy = 1$. Also

$$
\frac{f(x-)+f(x+)}{2} = \int_0^{\pi} 2F_n(y)\left[\frac{f(x-)+f(x+)}{2}\right]dy.
$$

Therefore,

$$
\left|\sigma_n f(x) - \frac{f(x-)+f(x+)}{2}\right| =
$$

$$
\left|\int_0^{\pi} 2F_n(y)\left[\frac{f(x-)+f(x+)}{2} - \frac{f(x-y)+f(x+y)}{2}\right]dy\right| \leq
$$

$$
\int_0^{r} 2F_n(y)\,\varepsilon dy + \int_r^{\pi} 2F_n(y)\,C dy
$$

where r is chosen small enough that

$$
\left|\frac{f(x-)+f(x+)}{2} - \frac{f(x-y)+f(x+y)}{2}\right| < \varepsilon \qquad (9.36)
$$

for all $0 < y \leq r$. Now using the estimate (9.34) we obtain

$$\begin{aligned}\left|\sigma_n f(x) - \frac{f(x-)+f(x+)}{2}\right| &\leq \varepsilon + C\int_r^\pi \frac{1}{\pi(n+1)\sin^2\left(\frac{r}{2}\right)}dy \\ &\leq \varepsilon + \frac{\widetilde{C}}{n+1}.\end{aligned}$$

and so, letting $n \to \infty$, we obtain the desired result.

In case that f is everywhere continuous, then since it is periodic, it must also be uniformly continuous. It follows that $f(x\pm) = f(x)$ and by the uniform continuity of f, we may choose r small enough that (9.36) holds for all x whenever $y < r$. Therefore, we obtain uniform convergence as claimed.

9.4 Gibb's phenomenon

The Fourier series converges to the mid point of the jump in the function under various conditions including those given above. However, in doing so the convergence cannot be uniform due to the discontinuity of the function to which it converges. In this section we show there is a small bump in the partial sums of the Fourier series on either side of the jump which does not disappear as the number of terms in the Fourier series increases. The small bump just gets narrower. To illustrate this phenomenon, known as Gibb's phenomenon, we consider a function, f, which equals -1 for $x < 0$ and 1 for $x > 0$. Thus the nth partial sum of the Fourier series is

$$S_n f(x) = \frac{4}{\pi}\sum_{k=1}^n \frac{\sin((2k-1)x)}{2k-1}.$$

We consider the value of this at the point $\frac{\pi}{2n}$. This equals

$$\frac{4}{\pi}\sum_{k=1}^n \frac{\sin\left((2k-1)\frac{\pi}{2n}\right)}{2k-1} = \frac{2}{\pi}\sum_{k=1}^n \frac{\sin\left((2k-1)\frac{\pi}{2n}\right)}{(2k-1)\left(\frac{\pi}{2n}\right)}\frac{\pi}{n}$$

which is seen to be a Riemann sum for the integral $\frac{2}{\pi}\int_0^\pi \frac{\sin y}{y}dy$. This integral is a positive constant approximately equal to 1.179. Therefore, although the value of the function equals 1 for all $x > 0$, we see that for large n, the value of the nth partial sum of the Fourier series at points near $x = 0$ equals approximately 1.179. To illustrate this phenomenon we graph the Fourier series of this function for large n, say $n = 10$. The following is the graph of the function, $S_{10}f(x) = \frac{4}{\pi}\sum_{k=1}^{10}\frac{\sin((2k-1)x)}{2k-1}$

You see the little blip near the jump which does not disappear. So you will see this happening for even larger n, we graph this for $n = 20$. The following is the graph of $S_{20}f(x) = \frac{4}{\pi}\sum_{k=1}^{20}\frac{\sin((2k-1)x)}{2k-1}$

As you can observe, it looks the same except the wriggles are a little closer together. Nevertheless, it still has a bump near the discontinuity.

9.5 The mean square convergence of Fourier series

We showed that in terms of pointwise convergence, Fourier series are inferior to the Cesaro means. However, there is a type of convergence that Fourier series do better than any other sequence of linear combinations of the functions, e^{ikx}. This convergence is often called mean square convergence. We describe this next.

Definition 9.19 *We say $f \in L^2(-\pi,\pi)$ if f is Lebesgue measurable and*

$$\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty.$$

We say a sequence of functions, $\{f_n\}$ *converges to a function,* f *in the mean square sense if*

$$\lim_{n\to\infty}\int_{-\pi}^{\pi}|f_n-f|^2\,dx=0.$$

Lemma 9.20 *If* $f\in L^2(-\pi,\pi)$, *then* $f\in L^1(-\pi,\pi)$.

Proof: We use the inequality $ab\leq\frac{a^2}{2}+\frac{b^2}{2}$ whenever $a,b\geq 0$, which follows from the inequality $(a-b)^2\geq 0$.

$$\int_{-\pi}^{\pi}|f(x)|\,dx\leq\int_{-\pi}^{\pi}\frac{|f(x)|^2}{2}dx+\int_{-\pi}^{\pi}\frac{1}{2}dx<\infty.$$

This proves the lemma.

From this lemma, we see we can at least discuss the Fourier series of a function in $L^2(-\pi,\pi)$. The following theorem is the main result which shows the superiority of the Fourier series in terms of mean square convergence.

Theorem 9.21 *For* c_k *complex numbers, the choice of* c_k *which minimizes the expression*

$$\int_{-\pi}^{\pi}\left|f(x)-\sum_{k=-n}^{n}c_ke^{ikx}\right|^2dx \tag{9.37}$$

is for c_k *to equal the Fourier coefficient,* α_k *where*

$$\alpha_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\,e^{-ikx}dx. \tag{9.38}$$

Also we have Bessel's inequality,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|f|^2\,dx\geq\sum_{k=-n}^{n}|\alpha_k|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}|S_nf|^2\,dx \tag{9.39}$$

where α_k *denotes the kth Fourier coefficient,*

$$\alpha_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\,e^{-ikx}dx. \tag{9.40}$$

Proof: It is routine to obtain that the expression in (9.37) equals

$$\begin{aligned}&\int_{-\pi}^{\pi}|f|^2\,dx-\sum_{k=-n}^{n}c_k\int_{-\pi}^{\pi}\overline{f}(x)\,e^{ikx}dx-\sum_{k=-n}^{n}\overline{c_k}\int_{-\pi}^{\pi}f(x)\,e^{-ikx}dx+2\pi\sum_{k=-n}^{n}|c_k|^2\\ =\;&\int_{-\pi}^{\pi}|f|^2\,dx-2\pi\sum_{k=-n}^{n}c_k\overline{\alpha_k}-2\pi\sum_{k=-n}^{n}\overline{c_k}\alpha_k+2\pi\sum_{k=-n}^{n}|c_k|^2\end{aligned}$$

where $\alpha_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\,e^{-ikx}dx$, the kth Fourier coefficient. Now

$$-c_k\overline{\alpha_k}-\overline{c_k}\alpha_k+|c_k|^2=|c_k-\alpha_k|^2-|\alpha_k|^2$$

Therefore,

$$\int_{-\pi}^{\pi}\left|f(x)-\sum_{k=-n}^{n}c_ke^{ikx}\right|^2dx=2\pi\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}|f|^2\,dx-\sum_{k=-n}^{n}|\alpha_k|^2+\sum_{k=-n}^{n}|\alpha_k-c_k|^2\right]\geq 0.$$

It is clear from this formula that the minimum occurs when $\alpha_k = c_k$ and that Bessel's inequality holds. It only remains to verify the equal sign in (9.39).

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}|S_n f|^2\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{k=-n}^{n}\alpha_k e^{ikx}\right)\left(\sum_{l=-n}^{n}\overline{\alpha_l}e^{-ilx}\right)dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=-n}^{n}|\alpha_k|^2\,dx = \sum_{k=-n}^{n}|\alpha_k|^2\,.
\end{aligned}$$

This proves the theorem.

This theorem has shown that if we measure the distance between two functions in the mean square sense,

$$d(f,g) = \left(\int_{-\pi}^{\pi}|f-g|^2\,dx\right)^{1/2},$$

then the partial sums of the Fourier series do a better job approximating the given function than any other linear combination of the functions e^{ikx} for $-n \leq k \leq n$. We show now that

$$\lim_{n\to\infty}\int_{-\pi}^{\pi}|f(x)-S_n f(x)|^2\,dx = 0$$

whenever $f \in L^2(-\pi,\pi)$. To begin with we need the following lemma.

Lemma 9.22 *Let $\varepsilon > 0$ and let $f \in L^2(-\pi,\pi)$. Then there exists $g \in C_c(-\pi,\pi)$ such that*

$$\int_{-\pi}^{\pi}|f-g|^2\,dx < \varepsilon.$$

Proof: We can use the dominated convergence theorem to conclude

$$\lim_{r\to 0}\int_{-\pi}^{\pi}\left|f - f\mathcal{X}_{(-\pi+r,\pi-r)}\right|^2 dx = 0.$$

Therefore, picking r small enough, we may define $k \equiv f\mathcal{X}_{(-\pi+r,\pi-r)}$ and have

$$\int_{-\pi}^{\pi}|f-k|^2\,dx < \frac{\varepsilon}{9}. \tag{9.41}$$

Now let $k = h^+ - h^- + i(l^+ - l^-)$ where the functions, h and l are nonnegative. We may then use Theorem 5.31 on the pointwise convergence of nonnegative simple functions to nonnegative measurable functions and the dominated convergence theorem to obtain a nonnegative simple function, $s^+ \leq h^+$ such that

$$\int_{-\pi}^{\pi}\left|h^+ - s^+\right|^2 dx < \frac{\varepsilon}{144}.$$

Similarly, we may obtain simple functions, s^-, t^+, and t^- such that

$$\int_{-\pi}^{\pi}\left|h^- - s^-\right|^2 dx < \frac{\varepsilon}{144}, \int_{-\pi}^{\pi}\left|l^+ - t^+\right|^2 dx < \frac{\varepsilon}{144},$$

$$\int_{-\pi}^{\pi}\left|l^- - t^-\right|^2 dx < \frac{\varepsilon}{144}.$$

Letting $s \equiv s^+ - s^- + i(t^+ - t^-)$, and using the inequality,

$$(a+b+c+d)^2 \leq 4\left(a^2+b^2+c^2+d^2\right),$$

we see that

$$\int_{-\pi}^{\pi} |k-s|^2 \, dx \leq$$

$$4\int_{-\pi}^{\pi} \left(\left|h^+ - s^+\right|^2 + \left|h^- - s^-\right|^2 + \left|l^+ - t^+\right|^2 + \left|l^- - t^-\right|^2\right) dx$$

$$\leq 4\left(\frac{\varepsilon}{144} + \frac{\varepsilon}{144} + \frac{\varepsilon}{144} + \frac{\varepsilon}{144}\right) = \frac{\varepsilon}{9}. \tag{9.42}$$

Let $s(x) = \sum_{i=1}^{n} c_i \mathcal{X}_{E_i}(x)$ where the E_i are disjoint and $E_i \subseteq (-\pi + r, \pi - r)$. Let $\alpha > 0$ be given. Using the regularity of Lebesgue measure, we can get compact sets, K_i and open sets, V_i such that

$$K_i \subseteq E_i \subseteq V_i \subseteq (-\pi + r, \pi - r)$$

and $m(V_i \setminus K_i) < \alpha$. Then letting $K_i \prec r_i \prec V_i$ and $g(x) \equiv \sum_{i=1}^{n} c_i r_i(x)$

$$\int_{-\pi}^{\pi} |s-g|^2 \, dx \leq \sum_{i=1}^{n} |c_i|^2 \, m(V_i \setminus K_i) \leq \alpha \sum_{i=1}^{n} |c_i|^2 < \frac{\varepsilon}{9} \tag{9.43}$$

provided we choose α small enough. Thus $g \in C_c(-\pi, \pi)$ and from (9.41) - (9.42),

$$\int_{-\pi}^{\pi} |f-g|^2 \, dx \leq 3\int_{-\pi}^{\pi} \left(|f-k|^2 + |k-s|^2 + |s-g|^2\right) dx < \varepsilon.$$

This proves the lemma.

With this lemma, we are ready to prove a theorem about the mean square convergence of Fourier series.

Theorem 9.23 *Let $f \in L^2(-\pi, \pi)$. Then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |f - S_n f|^2 \, dx = 0. \tag{9.44}$$

Proof: From Lemma 9.22 there exists $g \in C_c(-\pi, \pi)$ such that

$$\int_{-\pi}^{\pi} |f-g|^2 \, dx < \varepsilon.$$

Extend g to make the extended function 2π periodic. Then from Theorem 9.18, $\sigma_n g$ converges uniformly to g on all of $\mathbb{R}$. In particular, this uniform convergence occurs on $(-\pi, \pi)$. Therefore,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |\sigma_n g - g|^2 \, dx = 0.$$

Also note that $\sigma_n f$ is a linear combination of the functions e^{ikx} for $|k| \leq n$. Therefore,

$$\int_{-\pi}^{\pi} |\sigma_n g - g|^2 \, dx \geq \int_{-\pi}^{\pi} |S_n g - g|^2 \, dx$$

which implies

$$\lim_{n\to\infty}\int_{-\pi}^{\pi}|S_n g - g|^2\,dx = 0.$$

Also from (9.39),

$$\int_{-\pi}^{\pi}|S_n(g-f)|^2\,dx \le \int_{-\pi}^{\pi}|g-f|^2\,dx.$$

Therefore, if n is large enough, this shows

$$\int_{-\pi}^{\pi}|f - S_n f|^2\,dx \le 3\left(\int_{-\pi}^{\pi}|f-g|^2\,dx + \int_{-\pi}^{\pi}|g - S_n g|^2\,dx\right.$$

$$\left. + \int_{-\pi}^{\pi}|S_n(g-f)|^2\,dx\right)$$

$$\le 3(\varepsilon+\varepsilon+\varepsilon) = 9\varepsilon.$$

This proves the theorem.

9.6 Exercises

1. Let f be a continuous function defined on $[-\pi,\pi]$. Show there exists a polynomial, p such that $||p-f|| < \varepsilon$ where $||g|| \equiv \sup\{|g(x)| : x\in[-\pi,\pi]\}$. Extend this result to an arbitrary interval. This is called the Weierstrass approximation theorem. **Hint:** First find a linear function, $ax+b=y$ such that $f-y$ has the property that it has the same value at both ends of $[-\pi,\pi]$. Therefore, you may consider this as the restriction to $[-\pi,\pi]$ of a continuous periodic function, F. Now find a trig polynomial, $\sigma(x)\equiv a_0+\sum_{k=1}^{n}a_k\cos kx + b_k\sin kx$ such that $||\sigma - F|| < \frac{\varepsilon}{3}$. Recall (9.4). Now consider the power series of the trig functions.

2. Show that neither the Jordan nor the Dini criterion for pointwise convergence implies the other criterion. That is, find an example of a function for which Jordan's condition implies pointwise convergence but not Dini's and then find a function for which Dini works but Jordan does not. **Hint:** You might try considering something like $y=[\ln(1/x)]^{-1}$ for $x>0$ to get something for which Jordan works but Dini does not. For the other part, try something like $x\sin(1/x)$.

3. If $f\in L^2(-\pi,\pi)$ show using Bessel's inequality that $\lim_{n\to\infty}\int_{-\pi}^{\pi}f(x)e^{inx}dx = 0$. Can this be used to give a proof of the Riemann Lebesgue lemma for the case where $f\in L^2$?

4. Let $f(x)=x$ for $x\in(-\pi,\pi)$ and extend to make the resulting function defined on $\mathbb{R}$ and periodic of period 2π. Find the Fourier series of f. Verify the Fourier series converges to the midpoint of the jump and use this series to find a nice formula for $\frac{\pi}{4}$. **Hint:** For the last part consider $x=\frac{\pi}{2}$.

5. Let $f(x)=x^2$ on $(-\pi,\pi)$ and extend to form a 2π periodic function defined on $\mathbb{R}$. Find the Fourier series of f. Now obtain a famous formula for $\frac{\pi^2}{6}$ by letting $x=\pi$.

6. Let $f(x)=\cos x$ for $x\in(0,\pi)$ and define $f(x)\equiv-\cos x$ for $x\in(-\pi,0)$. Now extend this function to make it 2π periodic. Find the Fourier series of f.

7. Show that for $f \in L^2(-\pi,\pi)$,

$$\int_{-\pi}^{\pi} f(x)\overline{S_n f(x)}dx = 2\pi \sum_{k=-n}^{n} |\alpha_k|^2$$

where the α_k are the Fourier coefficients of f. Use this and the theorem about mean square convergence, Theorem 9.23, to show that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{k=-\infty}^{\infty} |\alpha_k|^2 \equiv \lim_{n\to\infty} \sum_{k=-n}^{n} |\alpha_k|^2$$

8. Suppose $f, g \in L^2(-\pi,\pi)$. Show using Problem 7

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f\overline{g}dx = \sum_{k=-\infty}^{\infty} \alpha_k \overline{\beta_k},$$

where α_k are the Fourier coefficients of f and β_k are the Fourier coefficients of g.

9. Find a formula for $\sum_{k=1}^{n} \sin kx$. **Hint:** Let $S_n = \sum_{k=1}^{n} \sin kx$. The $\sin\left(\frac{x}{2}\right) S_n = \sum_{k=1}^{n} \sin kx \sin\left(\frac{x}{2}\right)$. Now use a Trig. identity to write the terms of this series as a difference of cosines.

10. Prove the Dirichlet formula which says that $\sum_{k=p}^{q} a_k b_k = A_q b_q - A_{p-1} b_p + \sum_{k=p}^{q-1} A_k (b_k - b_{k+1})$. Here $A_q \equiv \sum_{k=1}^{q} a_k$.

11. Let $\{a_n\}$ be a sequence of positive numbers having the property that $\lim_{n\to\infty} na_n = 0$ and $na_n \geq (n+1)a_{n+1}$. Show that if this is so, it follows that the series, $\sum_{k=1}^{\infty} a_n \sin nx$ converges uniformly on $\mathbb{R}$. This is a variation of a very interesting problem found in Apostol's book, [3]. **Hint:** Use the Dirichlet formula of Problem 10 and consider the Fourier series for the 2π periodic extension of the function $f(x) = \pi - x$ on $(0, 2\pi)$. Show the partial sums for this Fourier series are uniformly bounded for $x \in \mathbb{R}$. To do this it might be of use to maximize the series $\sum_{k=1}^{n} \frac{\sin kx}{k}$ using methods of elementary calculus. Thus you would find the maximum of this function among the points where $\sum_{k=1}^{n} \cos(kx) = 0$. This sum can be expressed in a simple closed form using techniques similar to those in Problem 10. Then, having found the value of x at which the maximum is achieved, plug it in to $\sum_{k=1}^{n} \frac{\sin kx}{k}$ and observe you have a Riemann sum for a certain finite integral.

12. The problem in Apostol's book mentioned in Problem 11 is as follows. Let $\{a_k\}_{k=1}^{\infty}$ be a decreasing sequence of nonnegative numbers which satisfies $\lim_{n\to\infty} na_n = 0$. Then

$$\sum_{k=1}^{\infty} a_k \sin(kx)$$

converges uniformly on $\mathbb{R}$. **Hint: (**Following Jones [18]**)** First show that for $p < q$, and $x \in (0,\pi)$,

$$\left|\sum_{k=p}^{q} a_k \sin(kx)\right| \leq a_p \csc\left(\frac{x}{2}\right).$$

To do this, use summation by parts and establish the formula

$$\sum_{k=p}^{q} \sin(kx) = \frac{\cos\left(\left(p-\frac{1}{2}\right)x\right) - \cos\left(\left(q+\frac{1}{2}\right)x\right)}{2\sin\left(\frac{x}{2}\right)}.$$

Next show that if $a_k \leq \frac{C}{k}$ and $\{a_k\}$ is decreasing, then

$$\left|\sum_{k=1}^{n} a_k \sin(kx)\right| \leq 5C.$$

To do this, establish that on $(0,\pi)$ $\sin x \geq \frac{x}{\pi}$ and for any integer, k, $|\sin(kx)| \leq |kx|$ and then write

$$\begin{aligned}
\left|\sum_{k=1}^{n} a_k \sin(kx)\right| &\leq \sum_{k=1}^{m} a_k |\sin(kx)| + \overbrace{\left|\sum_{k=m+1}^{n} a_k \sin(kx)\right|}^{\text{This equals 0 if } m=n} \\
&\leq \sum_{k=1}^{m} \frac{C}{k}|kx| + a_{m+1}\csc\left(\frac{x}{2}\right) \\
&\leq Cmx + \frac{C}{m+1}\frac{\pi}{x}. \qquad (9.45)
\end{aligned}$$

Now consider two cases, $x \leq 1/n$ and $x > 1/n$. In the first case, let $m = n$ and in the second, choose m such that

$$n > \frac{1}{x} \geq m > \frac{1}{x} - 1.$$

Finally, establish the desired result by modifying a_k making it equal to a_p for all $k \leq p$ and then writing

$$\left|\sum_{k=p}^{q} a_k \sin(kx)\right| \leq$$

$$\left|\sum_{k=1}^{p} a_k \sin(kx)\right| + \left|\sum_{k=1}^{q} a_k \sin(kx)\right| \leq 10e(p)$$

where $e(p) \equiv \sup\{na_n : n \geq p\}$. This will verify uniform convergence on $(0,\pi)$. Now explain why this yields uniform convergence on all of $\mathbb{R}$.

13. Suppose $f(x) = \sum_{k=1}^{\infty} a_k \sin kx$ and that the convergence is uniform. Is it reasonable to suppose that $f'(x) = \sum_{k=1}^{\infty} a_k k \cos kx$? Explain.

14. Suppose $|u_k(x)| \leq K_k$ for all $x \in D$ where

$$\sum_{k=-\infty}^{\infty} K_k = \lim_{n\to\infty} \sum_{k=-n}^{n} K_k < \infty.$$

Show that $\sum_{k=-\infty}^{\infty} u_k(x)$ converges converges uniformly on D in the sense that for all $\varepsilon > 0$, there exists N such that whenever $n > N$,

$$\left|\sum_{k=-\infty}^{\infty} u_k(x) - \sum_{k=-n}^{n} u_k(x)\right| < \varepsilon$$

for all $x \in D$. This is called the Weierstrass M test.

15. Suppose f is a differentiable function of period 2π and suppose that both f and f' are in $L^2(-\pi,\pi)$ such that for all $x \in (-\pi,\pi)$ and y sufficiently small,

$$f(x+y) - f(x) = \int_x^{x+y} f'(t)\, dt.$$

Show that the Fourier series of f converges uniformly to f. **Hint:** First show using the Dini criterion that $S_n f(x) \to f(x)$ for all x. Next let $\sum_{k=-\infty}^{\infty} c_k e^{ikx}$ be the Fourier series for f. Then from the definition of c_k, show that for $k \neq 0, c_k = \frac{1}{ik} c_k'$ where c_k' is the Fourier coefficient of f'. Now use the Bessel's inequality to argue that $\sum_{k=-\infty}^{\infty} |c_k'|^2 < \infty$ and use the Cauchy Schwarz inequality to obtain $\sum |c_k| < \infty$. Then using the version of the Weierstrass M test given in Problem 14 obtain uniform convergence of the Fourier series to f.

16. Suppose $f \in L^2(-\pi,\pi)$ and that E is a measurable subset of $(-\pi,\pi)$. Show that

$$\lim_{n\to\infty} \int_E S_n f(x)\, dx = \int_E f(x)\, dx.$$

Can you conclude that

$$\int_E f(x)\, dx = \sum_{k=-\infty}^{\infty} c_k \int_E e^{ikx} dx?$$

The Frechet derivative

10.1 Norms for finite dimensional vector spaces

This chapter is on the derivative of a function defined on a finite dimensional normed vector space. In this chapter, X and Y are finite dimensional vector spaces which have a norm. We will say a set, $U \subseteq X$ is open if for every $p \in U$, there exists $\delta > 0$ such that

$$B(p, \delta) \equiv \{x : ||x - p|| < \delta\} \subseteq U.$$

Thus, a set is open if every point of the set is an interior point. To begin with we give an important inequality known as the Cauchy Schwartz inequality.

Theorem 10.1 *The following inequality holds for a_i and $b_i \in \mathbb{C}$.*

$$\left|\sum_{i=1}^{n} a_i \overline{b}_i\right| \leq \left(\sum_{i=1}^{n} |a_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |b_i|^2\right)^{1/2}. \tag{10.1}$$

Proof: Let $t \in \mathbb{R}$ and define

$$h(t) \equiv \sum_{i=1}^{n} (a_i + tb_i) \overline{(a_i + tb_i)} = \sum_{i=1}^{n} |a_i|^2 + 2t \operatorname{Re} \sum_{i=1}^{n} a_i \overline{b}_i + t^2 \sum_{i=1}^{n} |b_i|^2.$$

Now $h(t) \geq 0$ for all $t \in \mathbb{R}$. If all b_i equal 0, then the inequality (10.1) clearly holds so assume this does not happen. Then the graph of $y = h(t)$ is a parabola which opens up and intersects the t axis in at most one point. Thus there is either one real zero or none. Therefore, from the quadratic formula,

$$4\left(\operatorname{Re} \sum_{i=1}^{n} a_i \overline{b}_i\right)^2 \leq 4\left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |b_i|^2\right)$$

which shows

$$\left|\operatorname{Re} \sum_{i=1}^{n} a_i \overline{b}_i\right| \leq \left(\sum_{i=1}^{n} |a_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |b_i|^2\right)^{1/2} \tag{10.2}$$

To get the desired result, let $\omega \in \mathbb{C}$ be such that $|\omega| = 1$ and

$$\sum_{i=1}^{n} \omega a_i \overline{b}_i = \omega \sum_{i=1}^{n} a_i \overline{b}_i = \left|\sum_{i=1}^{n} a_i \overline{b}_i\right|.$$

Then apply (10.2) replacing a_i with ωa_i. Then

$$\begin{aligned}\left|\sum_{i=1}^n a_i\overline{b}_i\right| &= \operatorname{Re}\sum_{i=1}^n \omega a_i\overline{b}_i \leq \left(\sum_{i=1}^n |\omega a_i|^2\right)^{1/2}\left(\sum_{i=1}^n |b_i|^2\right)^{1/2} \\ &= \left(\sum_{i=1}^n |a_i|^2\right)^{1/2}\left(\sum_{i=1}^n |b_i|^2\right)^{1/2}.\end{aligned}$$

This proves the theorem.

Recall that a linear space X is a normed linear space if there is a norm defined on X, $||\cdot||$ satisfying

$$||\mathbf{x}|| \geq 0,\ \ ||\mathbf{x}|| = 0 \text{ if and only if } \mathbf{x} = 0,$$

$$||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||\,,$$

$$||c\mathbf{x}|| = |c|\,||\mathbf{x}||$$

whenever c is a scalar.

Definition 10.2 *We say a normed linear space,* $(X, ||\cdot||)$ *is a Banach space if it is complete. Thus, whenever,* $\{\mathbf{x}_n\}$ *is a Cauchy sequence, there exists* $\mathbf{x} \in X$ *such that* $\lim_{n\to\infty} ||\mathbf{x}-\mathbf{x}_n|| = 0$.

Let X be a finite dimensional normed linear space with norm $||\cdot||$where the field of scalars is denoted by $\mathbb{F}$ and is understood to be either $\mathbb{R}$ or $\mathbb{C}$. Let $\{\mathbf{v}_1, \cdot\cdot\cdot, \mathbf{v}_n\}$ be a basis for X. If $\mathbf{x} \in X$, we will denote by x_i the *ith* component of $\mathbf{x}$ with respect to this basis. Thus

$$\mathbf{x} = \sum_{i=1}^n x_i\mathbf{v}_i.$$

Definition 10.3 *For* $\mathbf{x} \in X$ *and* $\{\mathbf{v}_1, \cdot\cdot\cdot, \mathbf{v}_n\}$ *a basis, we define a new norm by*

$$|\mathbf{x}| \equiv \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}.$$

Similarly, for $\mathbf{y} \in Y$ *with basis* $\{\mathbf{w}_1, \cdot\cdot\cdot, \mathbf{w}_m\}$, *and* y_i *its components with respect to this basis,*

$$|\mathbf{y}| \equiv \left(\sum_{i=1}^m |y_i|^2\right)^{1/2}$$

For $A \in \mathcal{L}(X,Y)$, *the space of linear mappings from* X *to* Y,

$$||A|| \equiv \sup\{|A\mathbf{x}| : |\mathbf{x}| \leq 1\}. \tag{10.3}$$

We also say that a set U is an open set if for all $\mathbf{x} \in U$, there exists $r > 0$ such that

$$B(\mathbf{x},r) \subseteq U$$

where

$$B(\mathbf{x},r) \equiv \{\mathbf{y} : |\mathbf{y}-\mathbf{x}| < r\}.$$

Another way to say this is that every point of U is an interior point. The first thing we will show is that these two norms, $||\cdot||$ and $|\cdot|$, are equivalent. This means the conclusion of the following theorem holds.

Theorem 10.4 *Let $(X, ||\cdot||)$ be a finite dimensional normed linear space and let $|\cdot|$ be described above relative to a given basis, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$. Then $|\cdot|$ is a norm and there exist constants $\delta, \Delta > 0$ independent of $\mathbf{x}$ such that*

$$\delta\,||\mathbf{x}|| \leq |\mathbf{x}| \leq \Delta\,||\mathbf{x}||\,. \tag{10.4}$$

Proof: All of the above properties of a norm are obvious except the second, the triangle inequality. To establish this inequality, we use the Cauchy Schwartz inequality to write

$$\begin{aligned}
|\mathbf{x}+\mathbf{y}|^2 &\equiv \sum_{i=1}^n |x_i + y_i|^2 \leq \sum_{i=1}^n |x_i|^2 + \sum_{i=1}^n |y_i|^2 + 2\,\mathrm{Re}\sum_{i=1}^n x_i \overline{y}_i \\
&\leq |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\left(\sum_{i=1}^n |x_i|^2\right)^{1/2}\left(\sum_{i=1}^n |y_i|^2\right)^{1/2} \\
&= |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\,|\mathbf{x}|\,|\mathbf{y}| = (|\mathbf{x}| + |\mathbf{y}|)^2
\end{aligned}$$

and this proves the second property above.

It remains to show the equivalence of the two norms. By the Cauchy Schwartz inequality again,

$$\begin{aligned}
||\mathbf{x}|| &\equiv \left|\left|\sum_{i=1}^n x_i \mathbf{v}_i\right|\right| \leq \sum_{i=1}^n |x_i|\,||\mathbf{v}_i|| \leq |\mathbf{x}|\left(\sum_{i=1}^n ||\mathbf{v}_i||^2\right)^{1/2} \\
&\equiv \delta^{-1}\,|\mathbf{x}|\,.
\end{aligned}$$

This proves the first half of the inequality.

Suppose the second half of the inequality is not valid. Then there exists a sequence $\mathbf{x}^k \in X$ such that

$$\left|\mathbf{x}^k\right| > k\,\left|\left|\mathbf{x}^k\right|\right|,\ k = 1, 2, \cdots.$$

Then define

$$\mathbf{y}^k \equiv \frac{\mathbf{x}^k}{|\mathbf{x}^k|}.$$

It follows

$$\left|\mathbf{y}^k\right| = 1, \quad \left|\mathbf{y}^k\right| > k\,\left|\left|\mathbf{y}^k\right|\right|. \tag{10.5}$$

Letting y_i^k be the components of $\mathbf{y}^k$ with respect to the given basis, it follows the vector

$$\left(y_1^k, \cdots, y_n^k\right)$$

is a unit vector in $\mathbb{F}^n$. By the Heine Borel theorem, there exists a subsequence, still denoted by k such that

$$\left(y_1^k, \cdots, y_n^k\right) \to (y_1, \cdots, y_n)\,.$$

It follows from (10.5) and this that for

$$\mathbf{y} = \sum_{i=1}^n y_i \mathbf{v}_i,$$

$$0 = \lim_{k\to\infty} \left|\left|\mathbf{y}^k\right|\right| = \lim_{k\to\infty}\left|\left|\sum_{i=1}^n y_i^k \mathbf{v}_i\right|\right| = \left|\left|\sum_{i=1}^n y_i \mathbf{v}_i\right|\right|$$

but not all the y_i equal zero. This contradicts the assumption that $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis and this proves the second half of the inequality.

Corollary 10.5 *If $(X, ||\cdot||)$ is a finite dimensional normed linear space with the field of scalars $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, then X is complete.*

Proof: Let $\{\mathbf{x}^k\}$ be a Cauchy sequence. Then letting the components of $\mathbf{x}^k$ with respect to the given basis be

$$x_1^k, \cdot \cdot \cdot, x_n^k,$$

it follows from Theorem 10.4, that

$$\left(x_1^k, \cdot \cdot \cdot, x_n^k\right)$$

is a Cauchy sequence in $\mathbb{F}^n$ and so

$$\left(x_1^k, \cdot \cdot \cdot, x_n^k\right) \to (x_1, \cdot \cdot \cdot, x_n) \in \mathbb{F}^n.$$

Thus,

$$\mathbf{x}^k = \sum_{i=1}^n x_i^k \mathbf{v}_i \to \sum_{i=1}^n x_i \mathbf{v}_i \in X.$$

This proves the corollary.

Corollary 10.6 *Suppose X is a finite dimensional linear space with the field of scalars either $\mathbb{C}$ or $\mathbb{R}$ and $||\cdot||$ and $|||\cdot|||$ are two norms on X. Then there exist positive constants, δ and Δ, independent of $\mathbf{x} \in X$ such that*

$$\delta\, |||\mathbf{x}||| \leq ||\mathbf{x}|| \leq \Delta\, |||\mathbf{x}||| \,.$$

Thus any two norms are equivalent.

Proof: Let $\{\mathbf{v}_1, \cdot \cdot \cdot, \mathbf{v}_n\}$ be a basis for X and let $|\cdot|$ be the norm taken with respect to this basis which was described earlier. Then by Theorem 10.4, there are positive constants $\delta_1, \Delta_1, \delta_2, \Delta_2$, all independent of $\mathbf{x} \in X$ such that

$$\delta_2\, |||\mathbf{x}||| \leq |\mathbf{x}| \leq \Delta_2\, |||\mathbf{x}||| \,,$$

$$\delta_1\, ||\mathbf{x}|| \leq |\mathbf{x}| \leq \Delta_1\, ||\mathbf{x}|| \,.$$

Then

$$\delta_2\, |||\mathbf{x}||| \leq |\mathbf{x}| \leq \Delta_1\, ||\mathbf{x}|| \leq \frac{\Delta_1}{\delta_1}\, |\mathbf{x}| \leq \frac{\Delta_1 \Delta_2}{\delta_1}\, |||\mathbf{x}|||$$

and so

$$\frac{\delta_2}{\Delta_1}\, |||\mathbf{x}||| \leq ||\mathbf{x}|| \leq \frac{\Delta_2}{\delta_1}\, |||\mathbf{x}|||$$

which proves the corollary.

Definition 10.7 *Let X and Y be normed linear spaces with norms $||\cdot||_X$ and $||\cdot||_Y$ respectively. Then $\mathcal{L}(X, Y)$ denotes the space of linear transformations, called bounded linear transformations, mapping X to Y which have the property that*

$$||A|| \equiv \sup \left\{||Ax||_Y : ||x||_X \leq 1\right\} < \infty.$$

Then $||A||$ is referred to as the operator norm of the bounded linear transformation, A.

We leave it as an easy exercise to verify that $||\cdot||$ is a norm on $\mathcal{L}(X,Y)$ and it is always the case that

$$||Ax||_Y \leq ||A||\,||x||_X\,.$$

Theorem 10.8 *Let X and Y be finite dimensional normed linear spaces of dimension n and m respectively and denote by $||\cdot||$ the norm on either X or Y. Then if A is any linear function mapping X to Y, then $A \in \mathcal{L}(X,Y)$ and $(\mathcal{L}(X,Y), ||\cdot||)$ is a complete normed linear space of dimension nm with*

$$||A\mathbf{x}|| \leq ||A||\,||\mathbf{x}||\,.$$

Proof: We need to show the norm defined on linear transformations really is a norm. Again the first and third properties listed above for norms are obvious. We need to show the second and verify $||A|| < \infty$. Letting $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ be a basis and $|\cdot|$ defined with respect to this basis as above, there exist constants $\delta, \Delta > 0$ such that

$$\delta\,||\mathbf{x}|| \leq |\mathbf{x}| \leq \Delta\,||\mathbf{x}||\,.$$

Then,

$$\begin{aligned} ||A+B|| &\equiv \sup\{||(A+B)(\mathbf{x})|| : ||\mathbf{x}|| \leq 1\} \\ &\leq \sup\{||A\mathbf{x}|| : ||\mathbf{x}|| \leq 1\} + \sup\{||B\mathbf{x}|| : ||\mathbf{x}|| \leq 1\} \\ &\equiv ||A|| + ||B||\,. \end{aligned}$$

Next we verify that $||A|| < \infty$. This follows from

$$||A(\mathbf{x})|| = \left|\left|A\left(\sum_{i=1}^n x_i \mathbf{v}_i\right)\right|\right| \leq \sum_{i=1}^n |x_i|\,||A(\mathbf{v}_i)||$$

$$\leq |\mathbf{x}| \left(\sum_{i=1}^n ||A(\mathbf{v}_i)||^2\right)^{1/2} \leq \Delta\,||\mathbf{x}|| \left(\sum_{i=1}^n ||A(\mathbf{v}_i)||^2\right)^{1/2} < \infty.$$

Thus $||A|| \leq \Delta \left(\sum_{i=1}^n ||A(\mathbf{v}_i)||^2\right)^{1/2}$.

Next we verify the assertion about the dimension of $\mathcal{L}(X,Y)$. Let the two sets of bases be

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \text{ and } \{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$$

for X and Y respectively. Let $\mathbf{w}_i \otimes \mathbf{v}_k \in \mathcal{L}(X,Y)$ be defined by

$$\mathbf{w}_i \otimes \mathbf{v}_k \mathbf{v}_l \equiv \begin{cases} \mathbf{0} \text{ if } l \neq k \\ \mathbf{w}_i \text{ if } l = k \end{cases}$$

and let $L \in \mathcal{L}(X,Y)$. Then

$$L\mathbf{v}_r = \sum_{j=1}^m d_{jr}\mathbf{w}_j$$

for some d_{jk}. Also

$$\sum_{j=1}^m \sum_{k=1}^n d_{jk}\mathbf{w}_j \otimes \mathbf{v}_k(\mathbf{v}_r) = \sum_{j=1}^m d_{jr}\mathbf{w}_j.$$

It follows that

$$L = \sum_{j=1}^{m} \sum_{k=1}^{n} d_{jk} \mathbf{w}_j \otimes \mathbf{v}_k$$

because the two linear transformations agree on a basis. Since L is arbitrary this shows

$$\{\mathbf{w}_i \otimes \mathbf{v}_k : i = 1, \cdot \cdot \cdot, m, \ k = 1, \cdot \cdot \cdot, n\}$$

spans $\mathcal{L}(X, Y)$. If

$$\sum_{i,k} d_{ik} \mathbf{w}_i \otimes \mathbf{v}_k = \mathbf{0},$$

then

$$\mathbf{0} = \sum_{i,k} d_{ik} \mathbf{w}_i \otimes \mathbf{v}_k (\mathbf{v}_l) = \sum_{i=1}^{m} d_{il} \mathbf{w}_i$$

and so, since $\{\mathbf{w}_1, \cdot \cdot \cdot, \mathbf{w}_m\}$ is a basis, $d_{il} = 0$ for each $i = 1, \cdot \cdot \cdot, m$. Since l is arbitrary, this shows $d_{il} = 0$ for all i and l. Thus these linear transformations form a basis and this shows the dimension of $\mathcal{L}(X, Y)$ is mn as claimed. By Corollary 10.5 $(\mathcal{L}(X, Y), ||\cdot||)$ is complete. If $\mathbf{x} \neq \mathbf{0}$,

$$||A\mathbf{x}|| \frac{1}{||\mathbf{x}||} = \left|\left| A \frac{\mathbf{x}}{||\mathbf{x}||} \right|\right| \leq ||A||$$

This proves the theorem.

An interesting application of the notion of equivalent norms on $\mathbb{R}^n$ is the process of giving a norm on a finite Cartesian product of normed linear spaces.

Definition 10.9 *Let X_i, $i = 1, \cdot \cdot \cdot, n$ be normed linear spaces with norms, $||\cdot||_i$. For*

$$\mathbf{x} \equiv (x_1, \cdot \cdot \cdot, x_n) \in \prod_{i=1}^{n} X_i$$

define $\theta : \prod_{i=1}^{n} X_i \to \mathbb{R}^n$ by

$$\theta(\mathbf{x}) \equiv (||x_1||_1, \cdot \cdot \cdot, ||x_n||_n)$$

Then if $||\cdot||$ is any norm on $\mathbb{R}^n$, we define a norm on $\prod_{i=1}^{n} X_i$, also denoted by $||\cdot||$ by

$$||\mathbf{x}|| \equiv ||\theta \mathbf{x}||.$$

The following theorem follows immediately from Corollary 10.6.

Theorem 10.10 *Let X_i and $||\cdot||_i$ be given in the above definition and consider the norms on $\prod_{i=1}^{n} X_i$ described there in terms of norms on $\mathbb{R}^n$. Then any two of these norms on $\prod_{i=1}^{n} X_i$ obtained in this way are equivalent.*

For example, we may define

$$||\mathbf{x}||_1 \equiv \sum_{i=1}^{n} |x_i|,$$

$$||\mathbf{x}||_\infty \equiv \max\{||x_i||_i, i = 1, \cdot \cdot \cdot, n\},$$

or

$$||\mathbf{x}||_2 = \left(\sum_{i=1}^{n} ||x_i||_i^2 \right)^{1/2}$$

and all three are equivalent norms on $\prod_{i=1}^{n} X_i$.

10.2 The Derivative

Let U be an open set in X, a normed linear space and let $\mathbf{f} : U \rightarrow Y$ be a function.

Definition 10.11 *We say a function* $\mathbf{g}$ *is* $o(\mathbf{v})$ *if*

$$\lim_{||\mathbf{v}|| \rightarrow 0} \frac{\mathbf{g}(\mathbf{v})}{||\mathbf{v}||} = \mathbf{0} \tag{10.6}$$

We say a function $\mathbf{f} : U \rightarrow Y$ *is differentiable at* $\mathbf{x} \in U$ *if there exists a linear transformation* $L \in \mathcal{L}(X, Y)$ *such that*

$$\mathbf{f}(\mathbf{x} + \mathbf{v}) = \mathbf{f}(\mathbf{x}) + L\mathbf{v} + o(\mathbf{v})$$

This linear transformation L *is the definition of* $D\mathbf{f}(\mathbf{x})$, *the derivative sometimes called the Frechet derivative.*

Note that in finite dimensional normed linear spaces, it does not matter which norm we use in this definition because of Theorem 10.4 and Corollary 10.6. The definition means that the error,

$$\mathbf{f}(\mathbf{x} + \mathbf{v}) - \mathbf{f}(\mathbf{x}) - L\mathbf{v}$$

converges to $\mathbf{0}$ faster than $||\mathbf{v}||$. The term $o(\mathbf{v})$ is notation that is descriptive of the behavior in (10.6) and it is only this behavior that concerns us. Thus,

$$o(\mathbf{v}) = o(\mathbf{v}) + o(\mathbf{v}),\ o(t\mathbf{v}) = o(\mathbf{v}),\ ko(\mathbf{v}) = o(\mathbf{v})$$

and other similar observations hold. This notation is both sloppy and useful because it neglects details which are not important.

Theorem 10.12 *The derivative is well defined.*

Proof: Suppose both L_1 and L_2 work in the above definition. Then let $\mathbf{v}$ be any vector and let t be a real scalar which is chosen small enough that $t\mathbf{v} + \mathbf{x} \in U$. Then

$$\mathbf{f}(\mathbf{x} + t\mathbf{v}) = \mathbf{f}(\mathbf{x}) + L_1 t\mathbf{v} + o(t\mathbf{v}),\ \mathbf{f}(\mathbf{x} + t\mathbf{v}) = \mathbf{f}(\mathbf{x}) + L_2 t\mathbf{v} + o(t\mathbf{v}).$$

Therefore, subtracting these two yields

$$(L_2 - L_1)(t\mathbf{v}) = o(t).$$

Note that $o(t\mathbf{v}) = o(t)$ for fixed $\mathbf{v}$. Therefore, dividing by t yields

$$(L_2 - L_1)(\mathbf{v}) = \frac{o(t)}{t}.$$

Now let $t \rightarrow 0$ to conclude that $(L_2 - L_1)(\mathbf{v}) = 0$. This proves the theorem.

Lemma 10.13 *Let* $\mathbf{f}$ *be differentiable at* $\mathbf{x}$. *Then* $\mathbf{f}$ *is continuous at* $\mathbf{x}$ *and in fact, there exists* $K > 0$ *such that whenever* $||\mathbf{v}||$ *is small enough,*

$$||\mathbf{f}(\mathbf{x} + \mathbf{v}) - \mathbf{f}(\mathbf{x})|| \leq K ||\mathbf{v}||$$

Proof:

$$\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x})=D\mathbf{f}(\mathbf{x})\mathbf{v}+o(\mathbf{v}).$$

Let $||\mathbf{v}||$ be small enough that

$$||o(\mathbf{v})||\leq||\mathbf{v}||$$

Then

$$\begin{aligned}||\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x})|| &\leq ||D\mathbf{f}(\mathbf{x})\mathbf{v}||+||\mathbf{v}|| \\ &\leq (||D\mathbf{f}(\mathbf{x})||+1)||\mathbf{v}||\end{aligned}$$

This proves the lemma with $K=||D\mathbf{f}(\mathbf{x})||+1$.

Theorem 10.14 *(The chain rule) Let X,Y, and Z be normed linear spaces, and let $U\subseteq X$ be an open set and let $V\subseteq Y$ also be an open set. Suppose $\mathbf{f}:U\rightarrow V$ is differentiable at $\mathbf{x}$ and suppose $\mathbf{g}:V\rightarrow Z$ is differentiable at $\mathbf{f}(\mathbf{x})$. Then $\mathbf{g}\circ\mathbf{f}$ is differentiable at $\mathbf{x}$ and*

$$D(\mathbf{g}\circ\mathbf{f})(\mathbf{x})=D(\mathbf{g}(\mathbf{f}(\mathbf{x})))D(\mathbf{f}(\mathbf{x})).$$

Proof: This follows from a computation. Let $B(\mathbf{x},r)\subseteq U$ and let r also be small enough that for

$$||\mathbf{v}||\leq r,$$

$\mathbf{f}(\mathbf{x}+\mathbf{v})\in V$. For such $\mathbf{v}$, using the definition of differentiability of $\mathbf{g}$ and $\mathbf{f}$,

$$\mathbf{g}(\mathbf{f}(\mathbf{x}+\mathbf{v}))-\mathbf{g}(\mathbf{f}(\mathbf{x}))=$$

$$\begin{aligned}& D\mathbf{g}(\mathbf{f}(\mathbf{x}))(\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x}))+o(\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x})) \\ =\ & D\mathbf{g}(\mathbf{f}(\mathbf{x}))[D\mathbf{f}(\mathbf{x})\mathbf{v}+o(\mathbf{v})]+o(\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x})) \\ =\ & D(\mathbf{g}(\mathbf{f}(\mathbf{x})))D(\mathbf{f}(\mathbf{x}))\mathbf{v}+o(\mathbf{v})+o(\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x})).\end{aligned} \tag{10.7}$$

Now by Lemma 10.13, letting $\epsilon>0$, it follows that for $||\mathbf{v}||$ small enough,

$$||o(\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x}))||\leq\epsilon||\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x})||\leq\epsilon K||\mathbf{v}||.$$

Since $\epsilon>0$ is arbitrary, this shows $o(\mathbf{f}(\mathbf{x}+\mathbf{v})-\mathbf{f}(\mathbf{x}))=o(\mathbf{v})$. By (10.7), this shows

$$\mathbf{g}(\mathbf{f}(\mathbf{x}+\mathbf{v}))-\mathbf{g}(\mathbf{f}(\mathbf{x}))=D(\mathbf{g}(\mathbf{f}(\mathbf{x})))D(\mathbf{f}(\mathbf{x}))\mathbf{v}+o(\mathbf{v})$$

which proves the theorem.

We have defined the derivative as a linear transformation. This means that we can consider the matrix of the linear transformation with respect to various bases on X and Y. In the case where $X=\mathbb{R}^n$ and $Y=\mathbb{R}^m$,we shall denote the matrix taken with respect to the standard basis vectors $\mathbf{e}_i$, the vector with a 1 in the ith slot and zeros elsewhere, by $J\mathbf{f}(\mathbf{x})$. Thus, if the components of $\mathbf{v}$ with respect to the standard basis vectors are v_i,

$$\sum_j J\mathbf{f}(\mathbf{x})_{ij}v_j=\pi_i(D\mathbf{f}(\mathbf{x})\mathbf{v}) \tag{10.8}$$

where π_i is the projection onto the ith component of a vector in $Y=\mathbb{R}^m$. What are the entries of $J\mathbf{f}(x)$? Letting

$$\mathbf{f}(\mathbf{x})=\sum_{i=1}^m f_i(\mathbf{x})\mathbf{e}_i,$$

$$f_i(\mathbf{x}+\mathbf{v}) - f_i(\mathbf{x}) = \pi_i(D\mathbf{f}(\mathbf{x})\mathbf{v}) + o(\mathbf{v}).$$

Thus, letting t be a small scalar,

$$f_i(\mathbf{x}+t\mathbf{e}_j) - f_i(\mathbf{x}) = t\pi_i(D\mathbf{f}(\mathbf{x})\mathbf{e}_j) + o(t).$$

Dividing by t, and letting $t \to 0$,

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j} = \pi_i(D\mathbf{f}(\mathbf{x})\mathbf{e}_j).$$

Thus, from (10.8),

$$J\mathbf{f}(\mathbf{x})_{ij} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}. \tag{10.9}$$

This proves the following theorem

Theorem 10.15 *In the case where $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, if $\mathbf{f}$ is differentiable at $\mathbf{x}$ then all the partial derivatives*

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j}$$

exist and if $J\mathbf{f}(\mathbf{x})$ is the matrix of the linear transformation with respect to the standard basis vectors, then the ijth entry is given by (10.9).

What if all the partial derivatives of $\mathbf{f}$ exist? Does it follow that $\mathbf{f}$ is differentiable? Consider the following function. $f : \mathbb{R}^2 \to \mathbb{R}$.

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} \text{ if } (x,y) \neq (0,0) \\ 0 \text{ if } (x,y) = (0,0) \end{cases}.$$

Then from the definition of partial derivatives, this function has both partial derivatives at $(0,0)$. However f is not even continuous at $(0,0)$ which may be seen by considering the behavior of the function along the line $y = x$ and along the line $x = 0$. By Lemma 10.13 this implies f is not differentiable.

Lemma 10.16 *Suppose $X = \mathbb{R}^n$, $f : U \to \mathbb{R}$ and all the partial derivatives of f exist and are continuous in U. Then f is differentiable in U.*

Proof: Let $B(\mathbf{x}, r) \subseteq U$ and let $||\mathbf{v}|| < r$. Then,

$$f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) = \sum_{i=1}^{n}\left(f\left(\mathbf{x}+\sum_{j=1}^{i} v_j\mathbf{e}_j\right) - f\left(\mathbf{x}+\sum_{j=1}^{i-1} v_j\mathbf{e}_j\right)\right)$$

where

$$\sum_{i=1}^{0} v_j\mathbf{e}_j \equiv \mathbf{0}.$$

By the one variable mean value theorem,

$$f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f\left(\mathbf{x}+\sum_{j=1}^{i-1} v_j\mathbf{e}_j + \theta_i v_i\mathbf{e}_i\right)}{\partial x_i} v_i$$

where $\theta_j \in [0,1]$. Therefore,

$$f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i} v_i +$$

$$\sum_{i=1}^{n} \left(\frac{\partial f\left(\mathbf{x} + \sum_{j=1}^{i-1} v_j \mathbf{e}_j + \theta_i v_i \mathbf{e}_i\right)}{\partial x_i} - \frac{\partial f(\mathbf{x})}{\partial x_i} \right) v_i.$$

Consider the last term.

$$\left| \sum_{i=1}^{n} \left(\frac{\partial f\left(\mathbf{x} + \sum_{j=1}^{i-1} v_j \mathbf{e}_j + \theta_j v_j \mathbf{e}_j\right)}{\partial x_i} - \frac{\partial f(\mathbf{x})}{\partial x_i} \right) v_i \right| \leq \left(\sum_{i=1}^{n} |v_i|^2 \right)^{1/2} \cdot$$

$$\left(\sum_{i=1}^{n} \left| \left(\frac{\partial f\left(\mathbf{x} + \sum_{j=1}^{i-1} v_j \mathbf{e}_j + \theta_j v_j \mathbf{e}_j\right)}{\partial x_i} - \frac{\partial f(\mathbf{x})}{\partial x_i} \right) \right|^2 \right)^{1/2}$$

and so it follows from continuity of the partial derivatives that this last term is $o(\mathbf{v})$. Therefore, we define

$$L\mathbf{v} \equiv \sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i} v_i$$

where

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i.$$

Then L is a linear transformation which satisfies the conditions needed for it to equal $Df(\mathbf{x})$ and this proves the lemma.

Theorem 10.17 *Suppose $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ and $\mathbf{f} : U \to Y$ and suppose the partial derivatives,*

$$\frac{\partial f_i}{\partial x_j}$$

all exist and are continuous in U. Then $\mathbf{f}$ is differentiable in U.

Proof: From Lemma 10.16,

$$f_i(\mathbf{x}+\mathbf{v}) - f_i(\mathbf{x}) = Df_i(\mathbf{x})\,\mathbf{v} + o(\mathbf{v}).$$

Letting

$$(D\mathbf{f}(\mathbf{x})\,\mathbf{v})_i \equiv Df_i(\mathbf{x})\,\mathbf{v},$$

we see that

$$\mathbf{f}(\mathbf{x}+\mathbf{v}) - \mathbf{f}(\mathbf{x}) = D\mathbf{f}(\mathbf{x})\,\mathbf{v} + o(\mathbf{v})$$

and this proves the theorem.

When all the partial derivatives exist and are continuous we say the function is a C^1 function. More generally, we give the following definition.

Definition 10.18 *In the case where X and Y are normed linear spaces, and $U \subseteq X$ is an open set, we say $\mathbf{f} : U \to Y$ is $C^1(U)$ if $\mathbf{f}$ is differentiable and the mapping*

$$\mathbf{x} \to D\mathbf{f}(\mathbf{x}),$$

is continuous as a function from U to $\mathcal{L}(X, Y)$.

The following is an important abstract generalization of the concept of partial derivative defined above.

Definition 10.19 *Let X and Y be normed linear spaces. Then we can make $X \times Y$ into a normed linear space by defining a norm,*

$$||(\mathbf{x}, \mathbf{y})|| \equiv \max\left(||x||_X, ||\mathbf{y}||_Y\right).$$

Now let $\mathbf{g} : U \subseteq X \times Y \to Z$, where U is an open set and X, Y, and Z are normed linear spaces, and denote an element of $X \times Y$ by $(\mathbf{x}, \mathbf{y})$ where $\mathbf{x} \in X$ and $\mathbf{y} \in Y$. Then the map $\mathbf{x} \to \mathbf{g}(\mathbf{x}, \mathbf{y})$ is a function from the open set in X,

$$\{\mathbf{x} : (\mathbf{x}, \mathbf{y}) \in U\}$$

to Z. When this map is differentiable, we denote its derivative by

$$D_1\mathbf{g}(\mathbf{x}, \mathbf{y}), \text{ or sometimes by } D_{\mathbf{x}}\mathbf{g}(\mathbf{x}, \mathbf{y}).$$

Thus,

$$\mathbf{g}(\mathbf{x} + \mathbf{v}, \mathbf{y}) - \mathbf{g}(\mathbf{x}, \mathbf{y}) = D_1\mathbf{g}(\mathbf{x}, \mathbf{y})\mathbf{v} + o(\mathbf{v}).$$

A similar definition holds for the symbol $D_{\mathbf{y}}\mathbf{g}$ or $D_2\mathbf{g}$.

The following theorem will be very useful in much of what follows. It is a version of the mean value theorem.

Theorem 10.20 *Suppose X and Y are Banach spaces, U is an open subset of X and $\mathbf{f} : U \to Y$ has the property that $D\mathbf{f}(\mathbf{x})$ exists for all $\mathbf{x}$ in U and that, $\mathbf{x}+t(\mathbf{y} - \mathbf{x}) \in U$ for all $t \in [0, 1]$. (The line segment joining the two points lies in U.) Suppose also that for all points on this line segment,*

$$||D\mathbf{f}(\mathbf{x}+t(\mathbf{y} - \mathbf{x}))|| \leq M.$$

Then

$$||\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})|| \leq M ||\mathbf{y} - \mathbf{x}||.$$

Proof: Let

$$S \equiv \{t \in [0, 1] : \text{ for all } s \in [0, t],$$

$$||\mathbf{f}(\mathbf{x} + s(\mathbf{y} - \mathbf{x})) - \mathbf{f}(\mathbf{x})|| \leq (M + \epsilon) s ||\mathbf{y} - \mathbf{x}||\}.$$

Then $0 \in S$ and by continuity of $\mathbf{f}$, it follows that if $t \equiv \sup S$, then $t \in S$ and if $t < 1$,

$$||\mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - \mathbf{f}(\mathbf{x})|| = (M + \epsilon) t ||\mathbf{y} - \mathbf{x}||. \tag{10.10}$$

If $t < 1$, then there exists a sequence of positive numbers, $\{h_k\}_{k=1}^{\infty}$ converging to 0 such that

$$||\mathbf{f}(\mathbf{x} + (t + h_k)(\mathbf{y} - \mathbf{x})) - \mathbf{f}(\mathbf{x})|| > (M + \epsilon)(t + h_k) ||\mathbf{y} - \mathbf{x}||$$

which implies that

$$||\mathbf{f}(\mathbf{x}+(t+h_k)(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))||$$

$$+||\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})||>(M+\epsilon)(t+h_k)||\mathbf{y}-\mathbf{x}||.$$

By (10.10), this inequality implies

$$||\mathbf{f}(\mathbf{x}+(t+h_k)(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))||>(M+\epsilon)h_k||\mathbf{y}-\mathbf{x}||$$

which yields upon dividing by h_k and taking the limit as $h_k \to 0$,

$$||D\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))(\mathbf{y}-\mathbf{x})||>(M+\epsilon)||\mathbf{y}-\mathbf{x}||.$$

Now by the definition of the norm of a linear operator,

$$M||\mathbf{y}-\mathbf{x}||\geq||D\mathbf{f}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))||\,||\mathbf{y}-\mathbf{x}||>(M+\epsilon)||\mathbf{y}-\mathbf{x}||,$$

a contradiction. Therefore, $t=1$ and so

$$||\mathbf{f}(\mathbf{x}+(\mathbf{y}-\mathbf{x}))-\mathbf{f}(\mathbf{x})||\leq(M+\epsilon)||\mathbf{y}-\mathbf{x}||.$$

Since $\epsilon>0$ is arbitrary, this proves the theorem.

The next theorem is a version of the theorem presented earlier about continuity of the partial derivatives implying differentiability, presented in a more general setting. In the proof of this theorem, we will take

$$||(\mathbf{u},\mathbf{v})||\equiv\max\left(||\mathbf{u}||_X,||\mathbf{v}||_Y\right)$$

and always we will use the operator norm for linear maps.

Theorem 10.21 *Let $\mathbf{g}, U, X, Y$, and Z be given as in Definition 10.19. Then $\mathbf{g}$ is $C^1(U)$ if and only if $D_1\mathbf{g}$ and $D_2\mathbf{g}$ both exist and are continuous on U. In this case we have the formula,*

$$D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{v})=D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}.$$

Proof: Suppose first that $\mathbf{g}\in C^1(U)$. Then if $(\mathbf{x},\mathbf{y})\in U$,

$$\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})-\mathbf{g}(\mathbf{x},\mathbf{y})=D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{0})+o(\mathbf{u}).$$

Therefore, $D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}=D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{0})$. Then

$$||(D_1\mathbf{g}(\mathbf{x},\mathbf{y})-D_1\mathbf{g}(\mathbf{x}',\mathbf{y}'))(\mathbf{u})||=$$

$$||(D\mathbf{g}(\mathbf{x},\mathbf{y})-D\mathbf{g}(\mathbf{x}',\mathbf{y}'))(\mathbf{u},\mathbf{0})||\leq$$

$$||D\mathbf{g}(\mathbf{x},\mathbf{y})-D\mathbf{g}(\mathbf{x}',\mathbf{y}')||\,||(\mathbf{u},\mathbf{0})||.$$

Therefore,

$$||D_1\mathbf{g}(\mathbf{x},\mathbf{y})-D_1\mathbf{g}(\mathbf{x}',\mathbf{y}')||\leq||D\mathbf{g}(\mathbf{x},\mathbf{y})-D\mathbf{g}(\mathbf{x}',\mathbf{y}')||.$$

A similar argument applies for $D_2\mathbf{g}$ and this proves the continuity of the function, $(\mathbf{x},\mathbf{y})\to D_i\mathbf{g}(\mathbf{x},\mathbf{y})$ for $i=1,2$. The formula follows from

$$\begin{aligned}D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{v}) &= D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{0})+D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{0},\mathbf{v})\\ &\equiv D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}.\end{aligned}$$

Now suppose $D_1\mathbf{g}(\mathbf{x},\mathbf{y})$ and $D_2\mathbf{g}(\mathbf{x},\mathbf{y})$ exist and are continuous.

$$\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y})=\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})$$

$$+\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y})$$

$$=\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})-\mathbf{g}(\mathbf{x},\mathbf{y})+\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y})+$$

$$[\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})-(\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y}))]$$

$$=D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}+o(\mathbf{v})+o(\mathbf{u})+$$

$$[\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})-(\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y}))]. \tag{10.11}$$

Let $\mathbf{h}(\mathbf{x},\mathbf{u})\equiv\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})$. Then the expression in [ ] is of the form,

$$\mathbf{h}(\mathbf{x},\mathbf{u})-\mathbf{h}(\mathbf{x},\mathbf{0}).$$

Also

$$D_{\mathbf{u}}\mathbf{h}(\mathbf{x},\mathbf{u})=D_1\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-D_1\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})$$

and so, by continuity of $(\mathbf{x},\mathbf{y})\to D_1\mathbf{g}(\mathbf{x},\mathbf{y})$,

$$||D_{\mathbf{u}}\mathbf{h}(\mathbf{x},\mathbf{u})||<\varepsilon$$

whenever $||(\mathbf{u},\mathbf{v})||$ is small enough. By Theorem 10.20, there exists $\delta>0$ such that if $||(\mathbf{u},\mathbf{v})||<\delta$, the norm of the last term in (10.11) satisfies the inequality,

$$||\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y})-(\mathbf{g}(\mathbf{x},\mathbf{y}+\mathbf{v})-\mathbf{g}(\mathbf{x},\mathbf{y}))||<\varepsilon\,||\mathbf{u}||. \tag{10.12}$$

Therefore, this term is $o((\mathbf{u},\mathbf{v}))$. It follows from (10.12) and (10.11) that

$$\mathbf{g}(\mathbf{x}+\mathbf{u},\mathbf{y}+\mathbf{v})=$$

$$\mathbf{g}(\mathbf{x},\mathbf{y})+D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}+o(\mathbf{u})+o(\mathbf{v})+o((\mathbf{u},\mathbf{v}))$$

$$=\mathbf{g}(\mathbf{x},\mathbf{y})+D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}+o((\mathbf{u},\mathbf{v}))$$

Showing that $D\mathbf{g}(\mathbf{x},\mathbf{y})$ exists and is given by

$$D\mathbf{g}(\mathbf{x},\mathbf{y})(\mathbf{u},\mathbf{v})=D_1\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{u}+D_2\mathbf{g}(\mathbf{x},\mathbf{y})\mathbf{v}.$$

The continuity of $(\mathbf{x},\mathbf{y})\to D\mathbf{g}(\mathbf{x},\mathbf{y})$ follows from the continuity of $(\mathbf{x},\mathbf{y})\to D_i\mathbf{g}(\mathbf{x},\mathbf{y})$. This proves the theorem.

10.3 Higher order derivatives

If $f : U \to Y$, then

$$\mathbf{x} \to D\mathbf{f}(\mathbf{x})$$

is a mapping from U to $\mathcal{L}(X, Y)$, a normed linear space.

Definition 10.22 *The following is the definition of the second derivative.*

$$D^2\mathbf{f}(\mathbf{x}) \equiv D(D\mathbf{f}(\mathbf{x})).$$

Thus,

$$D\mathbf{f}(\mathbf{x}+\mathbf{v}) - D\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})\mathbf{v} + o(\mathbf{v}).$$

This implies

$$D^2\mathbf{f}(\mathbf{x}) \in \mathcal{L}(X, \mathcal{L}(X, Y)),\ D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v}) \in Y,$$

and the map

$$(\mathbf{u}, \mathbf{v}) \to D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})$$

is a bilinear map having values in Y. The same pattern applies to taking higher order derivatives. Thus,

$$D^3\mathbf{f}(\mathbf{x}) \equiv D\left(D^2\mathbf{f}(\mathbf{x})\right)$$

and we can consider $D^3\mathbf{f}(\mathbf{x})$ as a tri linear map. Also, instead of writing

$$D^2 f(\mathbf{x})(\mathbf{u})(\mathbf{v}),$$

we sometimes write

$$D^2 f(\mathbf{x})(\mathbf{u}, \mathbf{v}).$$

We say $\mathbf{f}$ is $C^k(U)$ if $\mathbf{f}$ and its first k derivatives are all continuous. For example, for $\mathbf{f}$ to be $C^2(U)$,

$$\mathbf{x} \to D^2\mathbf{f}(\mathbf{x})$$

would have to be continuous as a map from U to $\mathcal{L}(X, \mathcal{L}(X, Y))$. The following theorem deals with the question of symmetry of the map $D^2\mathbf{f}$.

This next lemma is a finite dimensional result but a more general result can be proved using the Hahn Banach theorem which will also be valid for an infinite dimensional setting. We leave this to the interested reader who has had some exposure to functional analysis. We are primarily interested in finite dimensional situations here, although most of the theorems and proofs given so far carry over to the infinite dimensional case with no change.

Lemma 10.23 *If* $\mathbf{z} \in Y$, *there exists* $L \in \mathcal{L}(Y, \mathbb{F})$ *such that*

$$L\mathbf{z} = |\mathbf{z}|^2,\ |L| \leq |\mathbf{z}|.$$

Here $|\mathbf{z}|^2 \equiv \sum_{i=1}^m |z_i|^2$ *where* $\mathbf{z} = \sum_{i=1}^m z_i\mathbf{w}_i$, *for* $\{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ *a basis for* Y *and*

$$|L| \equiv \sup\{|L\mathbf{x}| : |\mathbf{x}| \leq 1\},$$

the operator norm for L *with respect to this norm.*

Proof of the lemma: Let

$$L\mathbf{x} \equiv \sum_{i=1}^{m} x_i \overline{z}_i$$

where $\sum_{i=1}^{m} z_i \ \mathbf{w}_i = \mathbf{z}$. Then

$$L(\mathbf{z}) \equiv \sum_{i=1}^{m} z_i \overline{z}_i = \sum_{i=1}^{m} |z_i|^2 = |\mathbf{z}|^2 .$$

Also

$$|L\mathbf{x}| = \left| \sum_{i=1}^{m} x_i \overline{z}_i \right| \leq |\mathbf{x}| \, |\mathbf{z}|$$

and so $|L| \leq |\mathbf{z}|$. This proves the lemma.

Actually, the following lemma is valid but its proof involves the Hahn Banach theorem. Infinite dimensional versions of the following theorem will need this version of the lemma.

Lemma 10.24 *If $\mathbf{z} \in (Y, ||\ ||)$ a normed linear space, there exists $L \in \mathcal{L}(Y, \mathbb{F})$ such that*

$$L\mathbf{z} = ||\mathbf{z}||^2 , \ ||L|| \leq ||\mathbf{z}|| .$$

Theorem 10.25 *Suppose $f : U \subseteq X \to Y$ where X and Y are normed linear spaces, $D^2\mathbf{f}(\mathbf{x})$ exists for all $\mathbf{x} \in U$ and $D^2\mathbf{f}$ is continuous at $\mathbf{x} \in U$. Then*

$$D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v}) = D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u}) .$$

Proof: Let $B(\mathbf{x}, r) \subseteq U$ and let $t, s \in (0, r/2]$. Now let $L \in \mathcal{L}(Y, \mathbb{F})$ and define

$$\Delta(s,t) \equiv \frac{\operatorname{Re} L}{st} \{ \mathbf{f}(\mathbf{x}+t\mathbf{u}+s\mathbf{v}) - \mathbf{f}(\mathbf{x}+t\mathbf{u}) - (\mathbf{f}(\mathbf{x}+s\mathbf{v}) - \mathbf{f}(\mathbf{x})) \}. \tag{10.13}$$

Let $h(t) = \operatorname{Re} L(\mathbf{f}(\mathbf{x}+s\mathbf{v}+t\mathbf{u}) - \mathbf{f}(\mathbf{x}+t\mathbf{u}))$. Then by the mean value theorem,

$$\begin{aligned} \Delta(s,t) &= \frac{1}{st}(h(t) - h(0)) = \frac{1}{st} h'(\alpha t) t \\ &= \frac{1}{s} (\operatorname{Re} LD\mathbf{f}(\mathbf{x}+s\mathbf{v}+\alpha t\mathbf{u}) \mathbf{u} - \operatorname{Re} LD\mathbf{f}(\mathbf{x}+\alpha t\mathbf{u}) \mathbf{u}) . \end{aligned}$$

Applying the mean value theorem again,

$$\Delta(s,t) = \operatorname{Re} LD^2\mathbf{f}(\mathbf{x}+\beta s\mathbf{v}+\alpha t\mathbf{u})(\mathbf{v})(\mathbf{u})$$

where $\alpha, \beta \in (0,1)$. If the terms $\mathbf{f}(\mathbf{x}+t\mathbf{u})$ and $\mathbf{f}(\mathbf{x}+s\mathbf{v})$ are interchanged in (10.13), $\Delta(s,t)$ is also unchanged and the above argument shows there exist $\gamma, \delta \in (0,1)$ such that

$$\Delta(s,t) = \operatorname{Re} LD^2\mathbf{f}(\mathbf{x}+\gamma s\mathbf{v}+\delta t\mathbf{u})(\mathbf{u})(\mathbf{v}) .$$

Letting $(s,t) \to (0,0)$ and using the continuity of $D^2\mathbf{f}$ at $\mathbf{x}$,

$$\lim_{(s,t)\to(0,0)} \Delta(s,t) = \operatorname{Re} LD^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v}) = \operatorname{Re} LD^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u}) .$$

By Lemma 10.23, there exists $L \in \mathcal{L}(Y, \mathbb{F})$ such that for some norm on $Y, |\cdot|$,

$$L\left(D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v}) - D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u})\right) =$$

$$\left|D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})-D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u})\right|^2$$

and

$$|L|\leq\left|D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})-D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u})\right|.$$

For this L,

$$0=\operatorname{Re}L\left(D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})-D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u})\right)$$

$$=L\left(D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})-D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u})\right)$$

$$=\left|D^2\mathbf{f}(\mathbf{x})(\mathbf{u})(\mathbf{v})-D^2\mathbf{f}(\mathbf{x})(\mathbf{v})(\mathbf{u})\right|^2$$

and this proves the theorem in the case where the vector spaces are finite dimensional. We leave the general case to the reader. Use the second version of the above lemma, the one which depends on the Hahn Banach theorem in the last step of the proof where an auspicious choice is made for L.

Consider the important special case when $X=\mathbb{R}^n$ and $Y=\mathbb{R}$. If $\mathbf{e}_i$ are the standard basis vectors, what is

$$D^2f(\mathbf{x})(\mathbf{e}_i)(\mathbf{e}_j)?$$

To see what this is, use the definition to write

$$D^2f(\mathbf{x})(\mathbf{e}_i)(\mathbf{e}_j)=t^{-1}s^{-1}D^2f(\mathbf{x})(t\mathbf{e}_i)(s\mathbf{e}_j)$$

$$=t^{-1}s^{-1}\left(Df(\mathbf{x}+t\mathbf{e}_i)-Df(\mathbf{x})+o(t)\right)(s\mathbf{e}_j)$$

$$=t^{-1}s^{-1}\left(f(\mathbf{x}+t\mathbf{e}_i+s\mathbf{e}_j)-f(\mathbf{x}+t\mathbf{e}_i)\right.$$

$$\left.+o(s)-\left(f(\mathbf{x}+s\mathbf{e}_j)-f(\mathbf{x})+o(s)\right)+o(t)s\right).$$

First let $s\to0$ to get

$$t^{-1}\left(\frac{\partial f}{\partial x_j}(\mathbf{x}+t\mathbf{e}_i)-\frac{\partial f}{\partial x_j}(\mathbf{x})+o(t)\right)$$

and then let $t\to0$ to obtain

$$D^2f(\mathbf{x})(\mathbf{e}_i)(\mathbf{e}_j)=\frac{\partial^2 f}{\partial x_i\partial x_j}(\mathbf{x}) \tag{10.14}$$

Thus the theorem asserts that in this special case the mixed partial derivatives are equal at $\mathbf{x}$ if they are defined near $\mathbf{x}$ and continuous at $\mathbf{x}$.

10.4 Implicit function theorem

The following lemma is very useful.

Lemma 10.26 *Let $A \in \mathcal{L}(X, X)$ where X is a Banach space, (complete normed linear space), and suppose $||A|| \leq r < 1$. Then*

$$(I - A)^{-1} \text{ exists} \tag{10.15}$$

and

$$\left|\left|(I - A)^{-1}\right|\right| \leq (1 - r)^{-1}. \tag{10.16}$$

Furthermore, if

$$\mathcal{I} \equiv \left\{A \in \mathcal{L}(X, X) : A^{-1} \text{ exists}\right\}$$

the map $A \to A^{-1}$ is continuous on $\mathcal{I}$ and $\mathcal{I}$ is an open subset of $\mathcal{L}(X, X)$.

Proof: Consider

$$B_k \equiv \sum_{i=0}^{k} A^i.$$

Then if $N < l < k$,

$$||B_k - B_l|| \leq \sum_{i=N}^{k} \left|\left|A^i\right|\right| \leq \sum_{i=N}^{k} ||A||^i \leq \frac{r^N}{1 - r}.$$

It follows B_k is a Cauchy sequence and so it converges to $B \in \mathcal{L}(X, X)$. Also,

$$(I - A) B_k = I - A^{k+1} = B_k (I - A)$$

and so

$$I = \lim_{k \to \infty} (I - A) B_k = (I - A) B,\ I = \lim_{k \to \infty} B_k (I - A) = B (I - A).$$

Thus

$$(I - A)^{-1} = B = \sum_{i=0}^{\infty} A^i.$$

It follows

$$\left|\left|(I - A)^{-1}\right|\right| \leq \sum_{i=1}^{\infty} \left|\left|A^i\right|\right| \leq \sum_{i=0}^{\infty} ||A||^i = \frac{1}{1 - r}.$$

To verify the continuity of the inverse map, let $A \in \mathcal{I}$. Then

$$B = A\left(I - A^{-1}(A - B)\right)$$

and so if $\left|\left|A^{-1}(A - B)\right|\right| < 1$ it follows $B^{-1} = \left(I - A^{-1}(A - B)\right)^{-1} A^{-1}$ which shows $\mathcal{I}$ is open. Now for such B this close to A,

$$\left|\left|B^{-1} - A^{-1}\right|\right| = \left|\left|\left(I - A^{-1}(A - B)\right)^{-1} A^{-1} - A^{-1}\right|\right|$$

$$= \left|\left|\left(\left(I - A^{-1}(A - B)\right)^{-1} - I\right) A^{-1}\right|\right|$$

$$= \left|\left|\sum_{k=1}^{\infty} \left(A^{-1}(A-B)\right)^k A^{-1}\right|\right| \leq \sum_{k=1}^{\infty} \left|\left|A^{-1}(A-B)\right|\right|^k \left|\left|A^{-1}\right|\right|$$

$$= \frac{\left|\left|A^{-1}(A-B)\right|\right|}{1-\left|\left|A^{-1}(A-B)\right|\right|} \left|\left|A^{-1}\right|\right|$$

which shows that if $||A-B||$ is small, so is $\left|\left|B^{-1}-A^{-1}\right|\right|$. This proves the lemma.

The next theorem is a very useful result in many areas. It will be used in this section to give a short proof of the implicit function theorem but it is also useful in studying differential equations and integral equations. It is sometimes called the uniform contraction principle.

Theorem 10.27 *Let (Y,ρ) and (X,d) be complete metric spaces and suppose for each $(x,y) \in X \times Y$, $T(x,y) \in X$ and satisfies*

$$d(T(x,y), T(x',y)) \leq rd(x,x') \tag{10.17}$$

where $0 < r < 1$ and also

$$d(T(x,y), T(x,y')) \leq M\rho(y,y'). \tag{10.18}$$

Then for each $y \in Y$ there exists a unique "fixed point" for $T(\cdot,y)$, $x \in X$, satisfying

$$T(x,y) = x \tag{10.19}$$

and also if $x(y)$ is this fixed point,

$$d(x(y), x(y')) \leq \frac{M}{1-r}\rho(y,y'). \tag{10.20}$$

Proof: First we show there exists a fixed point for the mapping, $T(\cdot,y)$. For a fixed y, let $g(x) \equiv T(x,y)$. Now pick any $x_0 \in X$ and consider the sequence,

$$x_1 = g(x_0),\ x_{k+1} = g(x_k).$$

Then by (10.17),

$$d(x_{k+1}, x_k) = d(g(x_k), g(x_{k-1})) \leq rd(x_k, x_{k-1}) \leq$$

$$r^2 d(x_{k-1}, x_{k-2}) \leq \cdots \leq r^k d(g(x_0), x_0).$$

Now by the triangle inequality,

$$d(x_{k+p}, x_k) \leq \sum_{i=1}^{p} d(x_{k+i}, x_{k+i-1})$$

$$\leq \sum_{i=1}^{p} r^{k+i-1} d(x_0, g(x_0)) \leq \frac{r^k d(x_0, g(x_0))}{1-r}.$$

Since $0 < r < 1$, this shows that $\{x_k\}_{k=1}^{\infty}$ is a Cauchy sequence. Therefore, it converges to a point in X, x. To see x is a fixed point,

$$x = \lim_{k\to\infty} x_k = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) = g(x).$$

This proves (10.19). To verify (10.20),

$$d(x(y), x(y')) = d(T(x(y), y), T(x(y'), y')) \leq$$

$$d(T(x(y), y), T(x(y), y')) + d(T(x(y), y'), T(x(y'), y'))$$

$$\leq M\rho(y, y') + rd(x(y), x(y')).$$

Thus $(1-r) d(x(y), x(y')) \leq M\rho(y, y')$. This also shows the fixed point for a given y is unique. This proves the theorem.

The implicit function theorem is one of the most important results in Analysis. It provides the theoretical justification for such procedures as implicit differentiation taught in Calculus courses and has surprising consequences in many other areas. It deals with the question of solving, $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ for $\mathbf{x}$ in terms of $\mathbf{y}$ and how smooth the solution is. We give a proof of this theorem next. The proof we give will apply with no change to the case where the linear spaces are infinite dimensional once the necessary changes are made in the definition of the derivative. In a more general setting one assumes the derivative is what is called a bounded linear transformation rather than just a linear transformation as in the finite dimensional case. Basically, this means we assume the operator norm is defined. In the case of finite dimensional spaces, this boundedness of a linear transformation can be proved. We will use the norm for $X \times Y$ given by,

$$||(\mathbf{x}, \mathbf{y})|| \equiv \max\{||\mathbf{x}||, ||\mathbf{y}||\}.$$

Theorem 10.28 *(implicit function theorem) Let X, Y, Z be complete normed linear spaces and suppose U is an open set in $X \times Y$. Let $\mathbf{f} : U \to Z$ be in $C^1(U)$ and suppose*

$$\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0},\ D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} \in \mathcal{L}(Z, X). \tag{10.21}$$

Then there exist positive constants, δ, η, such that for every $\mathbf{y} \in B(\mathbf{y}_0, \eta)$ there exists a unique $\mathbf{x}(\mathbf{y}) \in B(\mathbf{x}_0, \delta)$ such that

$$\mathbf{f}(\mathbf{x}(\mathbf{y}), \mathbf{y}) = \mathbf{0}. \tag{10.22}$$

Futhermore, the mapping, $\mathbf{y} \to \mathbf{x}(\mathbf{y})$ is in $C^1(B(\mathbf{y}_0, \eta))$.

Proof: Let $\mathbf{T}(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x} - D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} \mathbf{f}(\mathbf{x}, \mathbf{y})$. Therefore,

$$D_1\mathbf{T}(\mathbf{x}, \mathbf{y}) = I - D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1} D_1\mathbf{f}(\mathbf{x}, \mathbf{y}). \tag{10.23}$$

by continuity of the derivative and Theorem 10.21, it follows that there exists $\delta > 0$ such that if $||(\mathbf{x} - \mathbf{x}_0, \mathbf{y} - \mathbf{y}_0)|| < \delta$, then

$$||D_1\mathbf{T}(\mathbf{x}, \mathbf{y})|| < \frac{1}{2},$$

$$\left|\left|D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1}\right|\right| ||D_2\mathbf{f}(\mathbf{x}, \mathbf{y})|| < M \tag{10.24}$$

where $M > \left|\left|D_1\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)^{-1}\right|\right| ||D_2\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)||$. By Theorem 10.20, whenever $\mathbf{x}, \mathbf{x}' \in B(\mathbf{x}_0, \delta)$ and $\mathbf{y} \in B(\mathbf{y}_0, \delta)$,

$$||\mathbf{T}(\mathbf{x}, \mathbf{y}) - \mathbf{T}(\mathbf{x}', \mathbf{y})|| \leq \frac{1}{2} ||\mathbf{x} - \mathbf{x}'||. \tag{10.25}$$

Solving (10.23) for $D_1\mathbf{f}(\mathbf{x},\mathbf{y})$,

$$D_1\mathbf{f}(\mathbf{x},\mathbf{y}) = D_1\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0)(I - D_1\mathbf{T}(\mathbf{x},\mathbf{y})).$$

By Lemma 10.26 and (10.24), $D_1\mathbf{f}(\mathbf{x},\mathbf{y})^{-1}$ exists and

$$\left|\left|D_1\mathbf{f}(\mathbf{x},\mathbf{y})^{-1}\right|\right| \leq 2\left|\left|D_1\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0)^{-1}\right|\right|. \tag{10.26}$$

Now we will restrict $\mathbf{y}$ some more. Let $0 < \eta < \min\left(\delta, \frac{\delta}{3M}\right)$. Then suppose $\mathbf{x} \in \overline{B(\mathbf{x}_0,\delta)}$ and $\mathbf{y} \in \overline{B(\mathbf{y}_0,\eta)}$. Consider $\mathbf{T}(\mathbf{x},\mathbf{y}) - \mathbf{x}_0 \equiv \mathbf{g}(\mathbf{x},\mathbf{y})$.

$$D_1\mathbf{g}(\mathbf{x},\mathbf{y}) = I - D_1\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0)^{-1} D_1\mathbf{f}(\mathbf{x},\mathbf{y}) = D_1\mathbf{T}(\mathbf{x},\mathbf{y}),$$

and

$$D_2\mathbf{g}(\mathbf{x},\mathbf{y}) = -D_1\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0)^{-1} D_2\mathbf{f}(\mathbf{x},\mathbf{y}).$$

Thus by (10.24), (10.21) saying that $\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0) = \mathbf{0}$, and Theorems 10.20 and (10.11), it follows that for such $(\mathbf{x},\mathbf{y})$,

$$||\mathbf{T}(\mathbf{x},\mathbf{y}) - \mathbf{x}_0|| = ||\mathbf{g}(\mathbf{x},\mathbf{y})|| = ||\mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{g}(\mathbf{x}_0,\mathbf{y}_0)||$$

$$\leq \frac{1}{2}||\mathbf{x} - \mathbf{x}_0|| + M||\mathbf{y} - \mathbf{y}_0|| < \frac{\delta}{2} + \frac{\delta}{3} = \frac{5\delta}{6} < \delta. \tag{10.27}$$

Also for such $(\mathbf{x},\mathbf{y}_i)$, $i = 1,2$, we can use Theorem 10.20 and (10.24) to obtain

$$||\mathbf{T}(\mathbf{x},\mathbf{y}_1) - \mathbf{T}(\mathbf{x},\mathbf{y}_2)|| = \left|\left|D_1\mathbf{f}(\mathbf{x}_0,\mathbf{y}_0)^{-1}(\mathbf{f}(\mathbf{x},\mathbf{y}_2) - \mathbf{f}(\mathbf{x},\mathbf{y}_1))\right|\right|$$

$$\leq M||\mathbf{y}_2 - \mathbf{y}_1||. \tag{10.28}$$

From now on we assume $||\mathbf{x} - \mathbf{x}_0|| < \delta$ and $||\mathbf{y} - \mathbf{y}_0|| < \eta$ so that (10.28), (10.26), (10.27), (10.25), and (10.24) all hold. By (10.28), (10.25), (10.27), and the uniform contraction principle, Theorem 10.27 applied to $X \equiv \overline{B\left(\mathbf{x}_0, \frac{5\delta}{6}\right)}$ and $Y \equiv \overline{B(\mathbf{y}_0,\eta)}$ implies that for each $\mathbf{y} \in B(\mathbf{y}_0,\eta)$, there exists a unique $\mathbf{x}(\mathbf{y}) \in B(\mathbf{x}_0,\delta)$ (actually in $\overline{B\left(\mathbf{x}_0, \frac{5\delta}{6}\right)}$) such that $\mathbf{T}(\mathbf{x}(\mathbf{y}),\mathbf{y}) = \mathbf{x}(\mathbf{y})$ which is equivalent to

$$\mathbf{f}(\mathbf{x}(\mathbf{y}),\mathbf{y}) = \mathbf{0}.$$

Furthermore,

$$||\mathbf{x}(\mathbf{y}) - \mathbf{x}(\mathbf{y}')|| \leq 2M||\mathbf{y} - \mathbf{y}'||. \tag{10.29}$$

This proves the implicit function theorem except for the verification that $\mathbf{y} \to \mathbf{x}(\mathbf{y})$ is C^1. This is shown next. Letting $\mathbf{v}$ be sufficiently small, Theorem 10.21 and Theorem 10.20 imply

$$\mathbf{0} = \mathbf{f}(\mathbf{x}(\mathbf{y}+\mathbf{v}),\mathbf{y}+\mathbf{v}) - \mathbf{f}(\mathbf{x}(\mathbf{y}),\mathbf{y}) =$$

$$D_1\mathbf{f}(\mathbf{x}(\mathbf{y}),\mathbf{y})(\mathbf{x}(\mathbf{y}+\mathbf{v}) - \mathbf{x}(\mathbf{y})) +$$

$$+D_2\mathbf{f}(\mathbf{x}(\mathbf{y}),\mathbf{y})\mathbf{v} + o((\mathbf{x}(\mathbf{y}+\mathbf{v}) - \mathbf{x}(\mathbf{y}),\mathbf{v})).$$

The last term in the above is $o(\mathbf{v})$ because of (10.29). Therefore, by (10.26), we can solve the above equation for $\mathbf{x}(\mathbf{y}+\mathbf{v})-\mathbf{x}(\mathbf{y})$ and obtain

$$\mathbf{x}(\mathbf{y}+\mathbf{v})-\mathbf{x}(\mathbf{y}) = -D_1(\mathbf{x}(\mathbf{y}),\mathbf{y})^{-1} D_2\mathbf{f}(\mathbf{x}(\mathbf{y}),\mathbf{y})\,\mathbf{v}+o(\mathbf{v})$$

Which shows that $\mathbf{y}\to\mathbf{x}(\mathbf{y})$ is differentiable on $B(\mathbf{y}_0,\eta)$ and

$$D\mathbf{x}(\mathbf{y}) = -D_1(\mathbf{x}(\mathbf{y}),\mathbf{y})^{-1} D_2\mathbf{f}(\mathbf{x}(\mathbf{y}),\mathbf{y}).$$

Now it follows from the continuity of $D_2\mathbf{f}$, $D_1\mathbf{f}$, the inverse map, (10.29), and this formula for $D\mathbf{x}(\mathbf{y})$that $\mathbf{x}(\cdot)$ is $C^1(B(\mathbf{y}_0,\eta))$. This proves the theorem.

The next theorem is a very important special case of the implicit function theorem known as the inverse function theorem. Actually one can also obtain the implicit function theorem from the inverse function theorem.

Theorem 10.29 *(Inverse Function Theorem) Let $\mathbf{x}_0\in U$, an open set in X , and let $\mathbf{f}:U\to Y$. Suppose*

$$\mathbf{f} \text{ is } C^1(U)\ ,\quad \text{and}\quad D\mathbf{f}(\mathbf{x}_0)^{-1}\in\mathcal{L}(Y,X). \tag{10.30}$$

Then there exist open sets, W, and V such that

$$\mathbf{x}_0\in W\subseteq U, \tag{10.31}$$

$$\mathbf{f}:W\to V \text{ is } 1-1 \text{ and onto}, \tag{10.32}$$

$$\mathbf{f}^{-1} \text{ is } C^1, \tag{10.33}$$

Proof: Apply the implicit function theorem to the function

$$\mathbf{F}(\mathbf{x},\mathbf{y})\equiv\mathbf{f}(\mathbf{x})-\mathbf{y}$$

where $\mathbf{y}_0\equiv\mathbf{f}(\mathbf{x}_0)$. Thus the function $\mathbf{y}\to\mathbf{x}(\mathbf{y})$ defined in that theorem is $\mathbf{f}^{-1}$. Now let

$$W\equiv B(\mathbf{x}_0,\delta)\cap\mathbf{f}^{-1}(B(\mathbf{y}_0,\eta))$$

and

$$V\equiv B(\mathbf{y}_0,\eta).$$

This proves the theorem.

Lemma 10.30 *Let*

$$O\equiv\{A\in\mathcal{L}(X,Y):A^{-1}\in\mathcal{L}(Y,X)\} \tag{10.34}$$

and let

$$\mathfrak{I}:O\to\mathcal{L}(Y,X),\ \mathfrak{I}A\equiv A^{-1}. \tag{10.35}$$

Then O is open and $\mathfrak{I}$ is in $C^m(O)$ for all $m=1,2,\cdots$ Also

$$D\mathfrak{I}(A)(B) = -\mathfrak{I}(A)(B)\,\mathfrak{I}(A). \tag{10.36}$$

Proof: Let $A \in O$ and let $B \in \mathcal{L}(X,Y)$ with

$$||B|| \leq \frac{1}{2}\left|\left|A^{-1}\right|\right|^{-1}.$$

Then

$$\left|\left|A^{-1}B\right|\right| \leq \left|\left|A^{-1}\right|\right| ||B|| \leq \frac{1}{2}$$

and so by Lemma 10.26,

$$\left(I + A^{-1}B\right)^{-1} \in \mathcal{L}(X,X).$$

Thus

$$(A+B)^{-1} = \left(I + A^{-1}B\right)^{-1} A^{-1} =$$

$$\sum_{n=0}^{\infty}(-1)^n \left(A^{-1}B\right)^n A^{-1} = \left[I - A^{-1}B + o(B)\right] A^{-1}$$

which shows that O is open and also,

$$\begin{aligned}
\mathfrak{I}(A+B) - \mathfrak{I}(A) &= \sum_{n=0}^{\infty}(-1)^n \left(A^{-1}B\right)^n A^{-1} - A^{-1} \\
&= -A^{-1}BA^{-1} + o(B) \\
&= -\mathfrak{I}(A)(B)\mathfrak{I}(A) + o(B)
\end{aligned}$$

which demonstrates (10.36). It follows from this that we can continue taking derivatives of $\mathfrak{I}$. For $||B_1||$ small,

$$-\left[D\mathfrak{I}(A+B_1)(B) - D\mathfrak{I}(A)(B)\right]$$

$$= \mathfrak{I}(A+B_1)(B)\mathfrak{I}(A+B_1) - \mathfrak{I}(A)(B)\mathfrak{I}(A)$$

$$\begin{aligned}
= \; & \mathfrak{I}(A+B_1)(B)\mathfrak{I}(A+B_1) - \mathfrak{I}(A)(B)\mathfrak{I}(A+B_1) + \\
& \mathfrak{I}(A)(B)\mathfrak{I}(A+B_1) - \mathfrak{I}(A)(B)\mathfrak{I}(A)
\end{aligned}$$

$$= \left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right](B)\mathfrak{I}(A+B_1)$$

$$+\mathfrak{I}(A)(B)\left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right]$$

$$\begin{aligned}
= \; & \left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right](B)\left[A^{-1} - A^{-1}B_1A^{-1}\right] + \\
& \mathfrak{I}(A)(B)\left[\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)\right]
\end{aligned}$$

$$= \mathfrak{I}(A)(B_1)\mathfrak{I}(A)(B)\mathfrak{I}(A) + \mathfrak{I}(A)(B)\mathfrak{I}(A)(B_1)\mathfrak{I}(A) + o(B_1)$$

and so

$$D^2\mathfrak{I}(A)(B_1)(B) =$$

$$\mathfrak{I}(A)(B_1)\mathfrak{I}(A)(B)\mathfrak{I}(A) + \mathfrak{I}(A)(B)\mathfrak{I}(A)(B_1)\mathfrak{I}(A)$$

which shows $\mathfrak{I}$ is $C^2(O)$. Clearly we can continue in this way, which shows $\mathfrak{I}$ is in $C^m(O)$ for all $m = 1, 2, \cdots$

Corollary 10.31 *In the inverse or implicit function theorems, assume*

$$\mathbf{f} \in C^m(U), m \geq 1.$$

Then

$$\mathbf{f}^{-1} \in C^m(V)$$

in the case of the inverse function theorem. In the implicit function theorem, the function

$$\mathbf{y} \to \mathbf{x}(\mathbf{y})$$

is C^m.

Proof: We consider the case of the inverse function theorem.

$$D\mathbf{f}^{-1}(\mathbf{y}) = \mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right).$$

Now by Lemma 10.30, and the chain rule,

$$D^2\mathbf{f}^{-1}(\mathbf{y})(B) =$$

$$-\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)(B)\,\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)D^2\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)D\mathbf{f}^{-1}(\mathbf{y})$$

$$= -\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)(B)\,\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)\cdot$$

$$D^2\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\mathfrak{I}\left(D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right).$$

Continuing in this way we see that it is possible to continue taking derivatives up to order m. Similar reasoning applies in the case of the implicit function theorem. This proves the corollary.

As an application of the implicit function theorem, we consider the method of Lagrange multipliers from calculus. Recall the problem is to maximize or minimize a function subject to equality constraints. Let

$$\mathbf{x} \in \mathbb{R}^n$$

and let $f : U \to \mathbb{R}$ be a C^1 function. Also let

$$g_i(\mathbf{x}) = 0,\ i = 1, \cdots, m \tag{10.37}$$

be a collection of equality constraints with $m < n$. Now consider the system of nonlinear equations

$$\begin{aligned} f(\mathbf{x}) &= a \\ g_i(\mathbf{x}) &= 0,\ i = 1, \cdots, m. \end{aligned}$$

We say $\mathbf{x}_0$ is a local maximum if $f(\mathbf{x}_0) \geq f(\mathbf{x})$ for all $\mathbf{x}$ near $\mathbf{x}_0$ which also satisfies the constraints (10.37). A local minimum is defined similarly. Let $\mathbf{F} : U \times \mathbb{R} \to \mathbb{R}^{m+1}$ be defined by

$$\mathbf{F}(\mathbf{x}, a) \equiv \begin{pmatrix} f(\mathbf{x}) - a \\ g_1(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{pmatrix}. \tag{10.38}$$

Now consider the $m+1\times n$ Jacobian matrix,

$$\left(\begin{array}{ccc} f_{x_1}(\mathbf{x}_0) & \cdots & f_{x_n}(\mathbf{x}_0) \\ g_{1x_1}(\mathbf{x}_0) & \cdots & g_{1x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ g_{mx_1}(\mathbf{x}_0) & \cdots & g_{mx_n}(\mathbf{x}_0) \end{array}\right).$$

If this matrix has rank $m+1$ then it follows from the implicit function theorem that we can select $m+1$ variables, $x_{i_1},\cdots,x_{i_{m+1}}$ such that the system

$$\mathbf{F}(\mathbf{x},a)=\mathbf{0} \tag{10.39}$$

specifies these $m+1$ variables as a function of the remaining $n-(m+1)$ variables and a in an open set of $\mathbb{R}^{n-m}$. Thus there is a solution $(\mathbf{x},a)$ to (10.39) for some $\mathbf{x}$ close to $\mathbf{x}_0$ whenever a is in some open interval. Therefore, $\mathbf{x}_0$ cannot be either a local minimum or a local maximum. It follows that if $\mathbf{x}_0$ is either a local maximum or a local minimum, then the above matrix must have rank less than $m+1$ which requires the rows to be linearly dependent. Thus, there exist m scalars,

$$\lambda_1,\cdots,\lambda_m,$$

and a scalar μ such that

$$\mu\left(\begin{array}{c} f_{x_1}(\mathbf{x}_0) \\ \vdots \\ f_{x_n}(\mathbf{x}_0) \end{array}\right)=\lambda_1\left(\begin{array}{c} g_{1x_1}(\mathbf{x}_0) \\ \vdots \\ g_{1x_n}(\mathbf{x}_0) \end{array}\right)+\cdots+\lambda_m\left(\begin{array}{c} g_{mx_1}(\mathbf{x}_0) \\ \vdots \\ g_{mx_n}(\mathbf{x}_0) \end{array}\right). \tag{10.40}$$

If the column vectors

$$\left(\begin{array}{c} g_{1x_1}(\mathbf{x}_0) \\ \vdots \\ g_{1x_n}(\mathbf{x}_0) \end{array}\right),\cdots\left(\begin{array}{c} g_{mx_1}(\mathbf{x}_0) \\ \vdots \\ g_{mx_n}(\mathbf{x}_0) \end{array}\right) \tag{10.41}$$

are linearly independent, then, $\mu\neq 0$ and dividing by μ yields an expression of the form

$$\left(\begin{array}{c} f_{x_1}(\mathbf{x}_0) \\ \vdots \\ f_{x_n}(\mathbf{x}_0) \end{array}\right)=\lambda_1\left(\begin{array}{c} g_{1x_1}(\mathbf{x}_0) \\ \vdots \\ g_{1x_n}(\mathbf{x}_0) \end{array}\right)+\cdots+\lambda_m\left(\begin{array}{c} g_{mx_1}(\mathbf{x}_0) \\ \vdots \\ g_{mx_n}(\mathbf{x}_0) \end{array}\right) \tag{10.42}$$

at every point $\mathbf{x}_0$ which is either a local maximum or a local minimum. This proves the following theorem.

Theorem 10.32 *Let U be an open subset of $\mathbb{R}^n$ and let $f:U\rightarrow\mathbb{R}$ be a C^1 function. Then if $\mathbf{x}_0\in U$ is either a local maximum or local minimum of f subject to the constraints (10.37), then (10.40) must hold for some scalars $\mu,\lambda_1,\cdots,\lambda_m$ not all equal to zero. If the vectors in (10.41) are linearly independent, it follows that an equation of the form (10.42) holds.*

10.5 Taylor's formula

First we recall the Taylor formula with the Lagrange form of the remainder. Since we will only need this on a specific interval, we will state it for this interval.

Theorem 10.33 *Let $h:(-\delta,1+\delta)\rightarrow\mathbb{R}$ have $m+1$ derivatives. Then there exists $t\in[0,1]$ such that*

$$h(1)=h(0)+\sum_{k=1}^{m}\frac{h^{(k)}(0)}{k!}+\frac{h^{(m+1)}(t)}{(m+1)!}.$$

Now let $f : U \to \mathbb{R}$ where $U \subseteq X$ a normed linear space and suppose $f \in C^m(U)$. Let $\mathbf{x} \in U$ and let $r > 0$ be such that

$$B(\mathbf{x},r) \subseteq U.$$

Then for $||\mathbf{v}|| < r$ we consider

$$f(\mathbf{x}+t\mathbf{v}) - f(\mathbf{x}) \equiv h(t)$$

for $t \in [0,1]$. Then

$$h'(t) = Df(\mathbf{x}+t\mathbf{v})(\mathbf{v}),\ h''(t) = D^2 f(\mathbf{x}+t\mathbf{v})(\mathbf{v})(\mathbf{v})$$

and continuing in this way, we see that

$$h^{(k)}(t) = D^{(k)} f(\mathbf{x}+t\mathbf{v})(\mathbf{v})(\mathbf{v}) \cdots (\mathbf{v}) \equiv D^{(k)} f(\mathbf{x}+t\mathbf{v})\mathbf{v}^k.$$

It follows from Taylor's formula for a function of one variable that

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + \sum_{k=1}^{m} \frac{D^{(k)} f(\mathbf{x})\mathbf{v}^k}{k!} + \frac{D^{(m+1)} f(\mathbf{x}+t\mathbf{v})\mathbf{v}^{m+1}}{(m+1)!}. \tag{10.43}$$

This proves the following theorem.

Theorem 10.34 *Let $f : U \to \mathbb{R}$ and let $f \in C^{m+1}(U)$. Then if*

$$B(\mathbf{x},r) \subseteq U,$$

and $||\mathbf{v}|| < r$, there exists $t \in (0,1)$ such that (10.43) holds.

Now we consider the case where $U \subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}$ is $C^2(U)$. Then from Taylor's theorem, if $\mathbf{v}$ is small enough, there exists $t \in (0,1)$ such that

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + Df(\mathbf{x})\mathbf{v} + \frac{D^2 f(\mathbf{x}+t\mathbf{v})\mathbf{v}^2}{2}.$$

Letting

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i,$$

where $\mathbf{e}_i$ are the usual basis vectors, the second derivative term reduces to

$$\frac{1}{2}\sum_{i,j} D^2 f(\mathbf{x}+t\mathbf{v})(\mathbf{e}_i)(\mathbf{e}_j) v_i v_j = \frac{1}{2}\sum_{i,j} H_{ij}(\mathbf{x}+t\mathbf{v}) v_i v_j$$

where

$$H_{ij}(\mathbf{x}+t\mathbf{v}) = D^2 f(\mathbf{x}+t\mathbf{v})(\mathbf{e}_i)(\mathbf{e}_j) = \frac{\partial^2 f(\mathbf{x}+t\mathbf{v})}{\partial x_j \partial x_i},$$

the Hessian matrix. From Theorem 10.25, this is a symmetric matrix. By the continuity of the second partial derivative and this,

$$f(\mathbf{x}+\mathbf{v}) = f(\mathbf{x}) + Df(\mathbf{x})\mathbf{v} + \frac{1}{2}\mathbf{v}^T H(\mathbf{x})\mathbf{v} +$$

$$\frac{1}{2}\left(\mathbf{v}^T\left(H\left(\mathbf{x}+t\mathbf{v}\right)-H\left(\mathbf{x}\right)\right)\mathbf{v}\right). \tag{10.44}$$

where the last two terms involve ordinary matrix multiplication and

$$\mathbf{v}^T = (v_1 \cdots v_n)$$

for v_i the components of $\mathbf{v}$ relative to the standard basis.

Theorem 10.35 *In the above situation, suppose $Df(\mathbf{x}) = 0$. Then if $H(\mathbf{x})$ has all positive eigenvalues, $\mathbf{x}$ is a local minimum. If $H(\mathbf{x})$ has all negative eigenvalues, then $\mathbf{x}$ is a local maximum. If $H(\mathbf{x})$ has a positive eigenvalue, then there exists a direction in which f has a local minimum at $\mathbf{x}$, while if $H(x)$ has a negative eigenvalue, there exists a direction in which $H(\mathbf{x})$ has a local maximum at $\mathbf{x}$.*

Proof: Since $Df(\mathbf{x}) = 0$, formula (10.44) holds and by continuity of the second derivative, we know $H(\mathbf{x})$ is a symmetric matrix. Thus $H(\mathbf{x})$ has all real eigenvalues. Suppose first that $H(\mathbf{x})$ has all positive eigenvalues and that all are larger than $\delta^2 > 0$. Then $H(\mathbf{x})$ has an orthonormal basis of eigenvectors, $\{\mathbf{v}_i\}_{i=1}^n$ and if $\mathbf{u}$ is an arbitrary vector, we can write $\mathbf{u} = \sum_{j=1}^n u_j\mathbf{v}_j$ where $u_j = \mathbf{u}\cdot\mathbf{v}_j$. Thus

$$\mathbf{u}^T H(\mathbf{x})\,\mathbf{u} = \sum_{j=1}^n u_j\mathbf{v}_j^T H(\mathbf{x}) \sum_{j=1}^n u_j\mathbf{v}_j$$

$$= \sum_{j=1}^n u_j^2\lambda_j \geq \delta^2 \sum_{j=1}^n u_j^2 = \delta^2 |\mathbf{u}|^2 .$$

From (10.44) and the continuity of H, if $\mathbf{v}$ is small enough,

$$f(\mathbf{x}+\mathbf{v}) \geq f(\mathbf{x}) + \frac{1}{2}\delta^2|\mathbf{v}|^2 - \frac{1}{4}\delta^2|\mathbf{v}|^2 = f(\mathbf{x}) + \frac{\delta^2}{4}|\mathbf{v}|^2 .$$

This shows the first claim of the theorem. The second claim follows from similar reasoning. Suppose $H(\mathbf{x})$ has a positive eigenvalue λ^2. Then let $\mathbf{v}$ be an eigenvector for this eigenvalue. Then from (10.44),

$$f(\mathbf{x}+t\mathbf{v}) = f(\mathbf{x}) + \frac{1}{2}t^2\mathbf{v}^T H(\mathbf{x})\,\mathbf{v} +$$

$$\frac{1}{2}t^2\left(\mathbf{v}^T\left(H\left(\mathbf{x}+t\mathbf{v}\right)-H\left(\mathbf{x}\right)\right)\mathbf{v}\right)$$

which implies

$$\begin{aligned} f(\mathbf{x}+t\mathbf{v}) &= f(\mathbf{x}) + \frac{1}{2}t^2\lambda^2|\mathbf{v}|^2 + \frac{1}{2}t^2\left(\mathbf{v}^T\left(H\left(\mathbf{x}+t\mathbf{v}\right)-H\left(\mathbf{x}\right)\right)\mathbf{v}\right) \\ &\geq f(\mathbf{x}) + \frac{1}{4}t^2\lambda^2|\mathbf{v}|^2 \end{aligned}$$

whenever t is small enough. Thus in the direction $\mathbf{v}$ the function has a local minimum at $\mathbf{x}$. The assertion about the local maximum in some direction follows similarly. This prove the theorem.

This theorem is an analogue of the second derivative test for higher dimensions. As in one dimension, when there is a zero eigenvalue, it may be impossible to determine from the Hessian matrix what the local qualitative behavior of the function is. For example, consider

$$f_1(x,y) = x^4 + y^2,\ f_2(x,y) = -x^4 + y^2.$$

Then $Df_i(0,0) = \mathbf{0}$ and for both functions, the Hessian matrix evaluated at $(0,0)$ equals

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

but the behavior of the two functions is very different near the origin. The second has a saddle point while the first has a minimum there.

10.6 Exercises

1. Suppose $L \in \mathcal{L}(X, Y)$ where X and Y are two finite dimensional vector spaces and suppose L is one to one. Show there exists $r > 0$ such that for all $\mathbf{x} \in X$,
$$|L\mathbf{x}| \geq r\,|\mathbf{x}|\,.$$
Hint: Define $|\mathbf{x}|_1 \equiv |L\mathbf{x}|$, observe that $|\cdot|_1$ is a norm and then use the theorem proved earlier that all norms are equivalent in a finite dimensional normed linear space.

2. Let U be an open subset of X, $\mathbf{f} : U \to Y$ where X, Y are finite dimensional normed linear spaces and suppose $\mathbf{f} \in C^1(U)$ and $D\mathbf{f}(\mathbf{x}_0)$ is one to one. Then show $\mathbf{f}$ is one to one near $\mathbf{x}_0$. **Hint:** Show using the assumption that $\mathbf{f}$ is C^1 that there exists $\delta > 0$ such that if
$$\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{x}_0, \delta)\,,$$
then
$$|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2) - D\mathbf{f}(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_2)| \leq \frac{r}{2}\,|\mathbf{x}_1 - \mathbf{x}_2|$$
then use Problem 1.

3. Suppose $M \in \mathcal{L}(X, Y)$ where X and Y are finite dimensional linear spaces and suppose M is onto. Show there exists $L \in \mathcal{L}(Y, X)$ such that
$$LM\mathbf{x} = P\mathbf{x}$$
where $P \in \mathcal{L}(X, X)$, and $P^2 = P$. **Hint:** Let $\{\mathbf{y}_1 \cdots \mathbf{y}_n\}$ be a basis of Y and let $M\mathbf{x}_i = \mathbf{y}_i$. Then define
$$L\mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i \text{ where } \mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{y}_i.$$
Show $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ is a linearly independent set and show you can obtain $\{\mathbf{x}_1, \cdots, \mathbf{x}_n, \cdots, \mathbf{x}_m\}$, a basis for X in which $M\mathbf{x}_j = \mathbf{0}$ for $j > n$. Then let
$$P\mathbf{x} \equiv \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$$
where
$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i.$$

4. Let $\mathbf{f} : U \to Y$, $\mathbf{f} \in C^1(U)$, and $D\mathbf{f}(\mathbf{x}_1)$ is onto. Show there exists $\delta, \epsilon > 0$ such that $\mathbf{f}(B(\mathbf{x}_1, \delta)) \supseteq B(\mathbf{f}(\mathbf{x}_1), \epsilon)$. **Hint:**Let
$$L \in \mathcal{L}(Y, X)\,,\ LD\mathbf{f}(\mathbf{x}_1)\,\mathbf{x} = P\mathbf{x},$$
and let $X_1 \equiv PX$ where $P^2 = P$, $\mathbf{x}_1 \in X_1$, and let $U_1 \equiv X_1 \cap U$. Now apply the inverse function theorem to $\mathbf{f}$ restricted to X_1.

5. Let $\mathbf{f} : U \to Y$, $\mathbf{f}$ is C^1, and $D\mathbf{f}(\mathbf{x})$ is onto for each $\mathbf{x} \in U$. Then show $\mathbf{f}$ maps open subsets of U onto open sets in Y.

6. Suppose $U \subseteq \mathbb{R}^2$ is an open set and $\mathbf{f} : U \to \mathbb{R}^3$ is C^1. Suppose $D\mathbf{f}(s_0, t_0)$ has rank two and
$$\mathbf{f}(s_0, t_0) = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}.$$
Show that for (s, t) near (s_0, t_0), the points $\mathbf{f}(s, t)$ may be realized in one of the following forms.
$$\{(x, y, \phi(x, y)) : (x, y) \text{ near } (x_0, y_0)\},$$
$$\{(\phi(y, z) y, z) : (y, z) \text{ near } (y_0, z_0)\},$$
or
$$\{(x, \phi(x, z), z,) : (x, z) \text{ near } (x_0, z_0)\}.$$

7. Suppose B is an open ball in X and $\mathbf{f} : B \to Y$ is differentiable. Suppose also there exists $L \in \mathcal{L}(X, Y)$ such that
$$||D\mathbf{f}(\mathbf{x}) - L|| < k$$
for all $\mathbf{x} \in B$. Show that if $\mathbf{x}_1, \mathbf{x}_2 \in B$,
$$||\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2) - L(\mathbf{x}_1 - \mathbf{x}_2)|| \leq k ||\mathbf{x}_1 - \mathbf{x}_2||.$$
Hint: Consider
$$||\mathbf{f}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)) - \mathbf{f}(\mathbf{x}_1) - tL(\mathbf{x}_2 - \mathbf{x}_1)||$$
and let
$$S \equiv \{t \in [0, 1] : ||\mathbf{f}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)) - \mathbf{f}(\mathbf{x}_1) - tL(\mathbf{x}_2 - \mathbf{x}_1)|| \leq$$
$$(k + \epsilon) t ||\mathbf{x}_2 - \mathbf{x}_1||\}.$$
Now imitate the proof of Theorem 10.20.

8. Let $\mathbf{f} : U \to Y$, $D\mathbf{f}(\mathbf{x})$ exists for all $\mathbf{x} \in U$, $B(\mathbf{x}_0, \delta) \subseteq U$, and there exists $L \in \mathcal{L}(X, Y)$, such that $L^{-1} \in \mathcal{L}(Y, X)$, and for all $\mathbf{x} \in B(\mathbf{x}_0, \delta)$
$$||D\mathbf{f}(\mathbf{x}) - L|| < \frac{r}{||L^{-1}||},\ r < 1.$$
Show that there exists $\epsilon > 0$ and an open subset of $B(\mathbf{x}_0, \delta), V$, such that $\mathbf{f} : V \to B(\mathbf{f}(\mathbf{x}_0), \epsilon)$ is one to one and onto. Also $D\mathbf{f}^{-1}(\mathbf{y})$ exists for each $\mathbf{y} \in B(f(\mathbf{x}_0), \epsilon)$ and is given by the formula
$$D\mathbf{f}^{-1}(\mathbf{y}) = \left[D\mathbf{f}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right]^{-1}.$$
Hint: Let
$$T_{\mathbf{y}}(\mathbf{x}) \equiv T(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x} - \mathbf{L}^{-1}(\mathbf{f}(\mathbf{x}) - \mathbf{y})$$
for $|\mathbf{y} - \mathbf{f}(\mathbf{x}_0)| < \frac{(1-r)\delta}{2||L^{-1}||}$, consider $\{T_{\mathbf{y}}^n(\mathbf{x}_0)\}$. This is a version of the inverse function theorem for $\mathbf{f}$ only differentiable, not C^1.

9. Denote by $C([0,T]:\mathbb{R}^n)$ the space of functions which are continuous having values in $\mathbb{R}^n$ and define a norm on this linear space as follows.

$$||\mathbf{f}||_\lambda \equiv \max\left\{|\mathbf{f}(t)|\, e^{\lambda t} : t \in [0,T]\right\}.$$

Show for each $\lambda \in \mathbb{R}$, this is a norm and that $C([0,T];\mathbb{R}^n)$ is a complete normed linear space with this norm.

10. Let $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be continuous and suppose $\mathbf{f}$ satisfies a Lipschitz condition,

$$|\mathbf{f}(t,\mathbf{x}) - \mathbf{f}(t,\mathbf{y})| \leq K\,|\mathbf{x}-\mathbf{y}|$$

and let $\mathbf{x}_0 \in \mathbb{R}^n$. Show there exists a unique solution to the Cauchy problem,

$$\mathbf{x}' = \mathbf{f}(t,\mathbf{x}),\ \mathbf{x}(0) = \mathbf{x}_0,$$

for $t \in [0,T]$. **Hint:** Consider the map

$$G : C([0,T];\mathbb{R}^n) \to C([0,T];\mathbb{R}^n)$$

defined by

$$G\mathbf{x}(t) \equiv \mathbf{x}_0 + \int_0^t \mathbf{f}(s,\mathbf{x}(s))\,ds,$$

where the integral is defined componentwise. Show G is a contraction map for $||\cdot||_\lambda$ given in Problem 9 for a suitable choice of λ and that therefore, it has a unique fixed point in $C([0,T];\mathbb{R}^n)$. Next argue, using the fundamental theorem of calculus, that this fixed point is the unique solution to the Cauchy problem.

11. Let (X,d) be a complete metric space and let $T : X \to X$ be a mapping which satisfies

$$d(T^n x, T^n y) \leq r d(x,y)$$

for some $r < 1$ whenever n is sufficiently large. Show T has a unique fixed point. Can you give another proof of Problem 10 using this result?

Change of variables for C^1 maps

In this chapter we will give theorems for the change of variables in Lebesgue integrals in which the mapping relating the two sets of variables is not linear. To begin with we will assume U is a nonempty open set and $\mathbf{h} : U \to V \equiv \mathbf{h}(U)$ is C^1, one to one, and $\det D\mathbf{h}(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in U$. Note that this implies by the inverse function theorem that V is also open. Using Theorem 3.32, there exist open sets, U_1 and U_2 which satisfy

$$\emptyset \neq U_1 \subseteq \overline{U_1} \subseteq U_2 \subseteq \overline{U_2} \subseteq U,$$

and $\overline{U_2}$ is compact. Then

$$0 < r \equiv dist\left(\overline{U_1}, U_2^C\right) \equiv \inf\left\{||\mathbf{x}-\mathbf{y}|| : \mathbf{x} \in \overline{U_1} \text{ and } \mathbf{y} \in U_2^C\right\}.$$

In this section $||\cdot||$ will be defined as

$$||\mathbf{x}|| \equiv \max\{|x_i|, i = 1, \cdots, n\}$$

where $\mathbf{x} \equiv (x_1, \cdots, x_n)^T$. We do this because with this definition, $B(\mathbf{x},r)$ is just an open n dimensional cube, $\prod_{i=1}^n (x_i - r, x_i + r)$ whose center is at $\mathbf{x}$ and whose sides are of length $2r$. This is not the most usual norm used and this norm has dreadful geometric properties but it is very convenient here because of the way we can fill open sets up with little n cubes of this sort. By Corollary 10.6 there are constants δ and Δ depending on n such that

$$\delta |\mathbf{x}| \leq ||\mathbf{x}|| \leq \Delta |\mathbf{x}|.$$

Therefore, letting $B_0(\mathbf{x},r)$ denote the ball taken with respect to the usual norm, $|\cdot|$, we obtain

$$B(\mathbf{x},r) \leq B_0\left(\mathbf{x},\delta^{-1}r\right) \leq B\left(\mathbf{x},\Delta\delta^{-1}r\right). \tag{11.1}$$

Thus we can use this norm in the definition of differentiability.

Recall that for $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$,

$$||A|| \equiv \sup\{||A\mathbf{x}|| : ||\mathbf{x}|| \leq 1\},$$

is called the operator norm of A taken with respect to the norm $||\cdot||$. Theorem 10.8 implied $||\cdot||$ is a norm on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ which satisfies

$$||A\mathbf{x}|| \leq ||A|| \, ||\mathbf{x}||$$

and is equivalent to the operator norm of A taken with respect to the usual norm, $|\cdot|$ because, by this theorem, $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a finite dimensional vector space and Corollary 10.6 implies any two norms on this finite dimensional vector space are equivalent.

We will also write dy or dx to denote the integration with respect to Lebesgue measure. This is done because the notation is standard and yields formulae which are easy to remember.

Lemma 11.1 *Let $\epsilon > 0$ be given. There exists $r_1 \in (0, r)$ such that whenever $\mathbf{x} \in \overline{U_1}$,*

$$\frac{||\mathbf{h}(\mathbf{x}+\mathbf{v}) - \mathbf{h}(\mathbf{x}) - D\mathbf{h}(\mathbf{x})\mathbf{v}||}{||\mathbf{v}||} < \epsilon$$

for all $||\mathbf{v}|| < r_1$.

Proof: The above expression equals

$$\frac{\left|\left|\int_0^1 (D\mathbf{h}(\mathbf{x}+t\mathbf{v})\,\mathbf{v} - D\mathbf{h}(\mathbf{x})\,\mathbf{v})\,dt\right|\right|}{||\mathbf{v}||}$$

which is no larger than

$$\frac{\int_0^1 ||D\mathbf{h}(\mathbf{x}+t\mathbf{v}) - D\mathbf{h}(\mathbf{x})||\,dt\,||\mathbf{v}||}{||\mathbf{v}||} =$$

$$\int_0^1 ||D\mathbf{h}(\mathbf{x}+t\mathbf{v}) - D\mathbf{h}(\mathbf{x})||\,dt. \tag{11.2}$$

Now $\mathbf{x} \to D\mathbf{h}(\mathbf{x})$ is uniformly continuous on the compact set $\overline{U_2}$. Therefore, there exists $\delta_1 > 0$ such that if $||\mathbf{x}-\mathbf{y}|| < \delta_1$, then

$$||D\mathbf{h}(\mathbf{x}) - D\mathbf{h}(\mathbf{y})|| < \epsilon.$$

Let $0 < r_1 < \min(\delta_1, r)$. Then if $||\mathbf{v}|| < r_1$, the expression in (11.2) is smaller than ϵ whenever $\mathbf{x} \in \overline{U_1}$.

Now let $\mathcal{D}_p$ consist of all rectangles

$$\prod_{i=1}^n (a_i, b_i] \cap (-\infty, \infty)$$

where $a_i = l2^{-p}$ or $\pm\infty$ and $b_i = (l+1)\,2^{-p}$ or $\pm\infty$ for k, l, p integers with $p > 0$. The following lemma is the key result in establishing a change of variable theorem for the above mappings, $\mathbf{h}$.

Lemma 11.2 *Let $R \in \mathcal{D}_p$. Then*

$$\int \mathcal{X}_{\mathbf{h}(R\cap U_1)}(\mathbf{y})\,dy \le \int \mathcal{X}_{R\cap U_1}(\mathbf{x})\,|\det D\mathbf{h}(\mathbf{x})|\,dx \tag{11.3}$$

and also the integral on the left is well defined because the integrand is measurable.

Proof: The set, U_1 is the countable union of disjoint sets, $\left\{\widetilde{R_i}\right\}_{i=1}^\infty$ of rectangles, $\widetilde{R_j} \in \mathcal{D}_q$ where $q > p$ is chosen large enough that for r_1 described in Lemma 11.1,

$$2^{-q} < r_1, ||\det D\mathbf{h}(\mathbf{x})| - |\det D\mathbf{h}(\mathbf{y})|| < \epsilon$$

if $||\mathbf{x}-\mathbf{y}|| \le 2^{-q}, \mathbf{x}, \mathbf{y} \in \overline{U_1}$. Each of these sets, $\left\{\widetilde{R_i}\right\}_{i=1}^\infty$ is either a subset of R or has empty intersection with R. Denote by $\{R_i\}_{i=1}^\infty$ those which are contained in R. Thus

$$R \cap U_1 = \cup_{i=1}^\infty R_i.$$

The set, $\mathbf{h}(R_i)$, is a Borel set because for

$$R_i \equiv \prod_{i=1}^{n}(a_i, b_i],$$

$\mathbf{h}(R_i)$ equals

$$\cup_{k=1}^{\infty}\mathbf{h}\left(\prod_{i=1}^{n}\left[a_i + k^{-1}, b_i\right]\right),$$

a countable union of compact sets. Then $\mathbf{h}(R \cap U_1) = \cup_{i=1}^{\infty}\mathbf{h}(R_i)$ and so it is also a Borel set. Therefore, the integrand in the integral on the left of (11.3) is measurable.

$$\int \mathcal{X}_{\mathbf{h}(R\cap U_1)}(\mathbf{y})\, dy = \int \sum_{i=1}^{\infty} \mathcal{X}_{\mathbf{h}(R_i)}(\mathbf{y})\, dy = \sum_{i=1}^{\infty} \int \mathcal{X}_{\mathbf{h}(R_i)}(\mathbf{y})\, dy.$$

Now by Lemma 11.1 if $\mathbf{x}_i$ is the center of interior$(R_i) = B(\mathbf{x}_i, \delta)$, then since all these rectangles are chosen so small, we have for $||\mathbf{v}|| \leq \delta$,

$$\mathbf{h}(\mathbf{x}_i + \mathbf{v}) - \mathbf{h}(\mathbf{x}_i) = D\mathbf{h}(\mathbf{x}_i)\left(\mathbf{v} + D\mathbf{h}(\mathbf{x}_i)^{-1} o(\mathbf{v})\right)$$

where $||o(\mathbf{v})|| < \epsilon\, ||\mathbf{v}||$. Therefore,

$$\mathbf{h}(R_i) \subseteq \mathbf{h}(\mathbf{x}_i) + D\mathbf{h}(\mathbf{x}_i)\left(\overline{B\left(\mathbf{0}, \delta\left(1 + \epsilon\left|\left|Dh(\mathbf{x}_i)^{-1}\right|\right|\right)\right)}\right)$$

$$\subseteq \mathbf{h}(\mathbf{x}_i) + D\mathbf{h}(\mathbf{x}_i)\left(\overline{B(\mathbf{0}, \delta(1 + \epsilon C))}\right)$$

where

$$C = \max\left\{\left|\left|D\mathbf{h}(\mathbf{x})^{-1}\right|\right| : \mathbf{x} \in \overline{U_1}\right\}.$$

It follows from Theorem 7.21 that

$$m_n(\mathbf{h}(R_i)) \leq |\det D\mathbf{h}(\mathbf{x}_i)|\, m_n(R_i)(1 + \epsilon C)^n$$

and hence

$$\int_{\mathbf{h}(R\cap U_1)} dy \leq \sum_{i=1}^{\infty} |\det D\mathbf{h}(\mathbf{x}_i)|\, m_n(R_i)(1 + \epsilon C)^n$$

$$= \sum_{i=1}^{\infty} \int_{R_i} |\det D\mathbf{h}(\mathbf{x}_i)|\, dx\, (1 + \epsilon C)^n$$

$$\leq \sum_{i=1}^{\infty} \int_{R_i} (|\det D\mathbf{h}(\mathbf{x})| + \epsilon)\, dx\, (1 + \epsilon C)^n$$

$$= \left(\sum_{i=1}^{\infty} \int_{R_i} |\det D\mathbf{h}(\mathbf{x})|\, dx\right)(1 + \epsilon C)^n + \epsilon m_n(U_1)(1 + \epsilon C)^n$$

$$= \left(\left(\int \mathcal{X}_{R \cap U_1} (\mathbf{x}) \left| \det D\mathbf{h} (\mathbf{x}) \right| dx \right) + \epsilon m_n (U_1) \right) (1 + \epsilon C)^n .$$

Since ϵ is arbitrary, this proves the Lemma.

Borel measurable functions have a very nice feature. When composed with a continuous map, the result is still Borel measurable. This is a general result described in the following lemma which we will use shortly.

Lemma 11.3 *Let* $\mathbf{h}$ *be a continuous map from* U *to* $\mathbb{R}^n$ *and let* $g : \mathbb{R}^n \to (X, \tau)$ *be Borel measurable. Then* $g \circ \mathbf{h}$ *is also Borel measurable.*

Proof: Let $V \in \tau$. Then

$$(g \circ \mathbf{h})^{-1} (V) = \mathbf{h}^{-1} \left(g^{-1} (V) \right) = \mathbf{h}^{-1} (\text{Borel set}) .$$

The conclusion of the lemma will follow if we show that $\mathbf{h}^{-1}$ (Borel set) = Borel set. Let

$$\mathcal{S} \equiv \left\{ E \in \text{Borel sets} : \mathbf{h}^{-1} (E) \text{ is Borel} \right\} .$$

Then $\mathcal{S}$ is a σ algebra which contains the open sets and so it coincides with the Borel sets. This proves the lemma.

Let $\mathcal{E}_q$ consist of all finite disjoint unions of sets of $\cup \{ \mathcal{D}_p : p \geq q \}$. By Lemma 5.6 and Corollary 5.7, $\mathcal{E}_q$ is an algebra. Let $\mathcal{E} \equiv \cup_{q=1}^{\infty} \mathcal{E}_q$. Then $\mathcal{E}$ is also an algebra.

Lemma 11.4 *Let* $E \in \sigma (\mathcal{E})$. *Then*

$$\int \mathcal{X}_{\mathbf{h}(E \cap U_1)} (\mathbf{y}) \, dy \leq \int \mathcal{X}_{E \cap U_1} (\mathbf{x}) \left| \det D\mathbf{h} (\mathbf{x}) \right| dx \tag{11.4}$$

and the integrand on the left of the inequality is measurable.

Proof: Let

$$\mathcal{M} \equiv \{ E \in \sigma (\mathcal{E}) : (11.4) \text{ holds} \} .$$

Then from the monotone and dominated convergence theorems, $\mathcal{M}$ is a monotone class. By Lemma 11.2, $\mathcal{M}$ contains $\mathcal{E}$, it follows $\mathcal{M}$ equals $\sigma (\mathcal{E})$ by the theorem on monotone classes. This proves the lemma.

Note that by Lemma 7.3 $\sigma (\mathcal{E})$ contains the open sets and so it contains the Borel sets also. Now let F be any Borel set and let $V_1 = \mathbf{h} (U_1)$. By the inverse function theorem, V_1 is open. Thus for F a Borel set, $F \cap V_1$ is also a Borel set. Therefore, since $F \cap V_1 = \mathbf{h} \left(\mathbf{h}^{-1} (F) \cap U_1 \right)$, we can apply (11.4) in the first inequality below and write

$$\int_{V_1} \mathcal{X}_F (\mathbf{y}) \, dy \equiv \int \mathcal{X}_{F \cap V_1} (\mathbf{y}) \, dy \leq \int \mathcal{X}_{\mathbf{h}^{-1}(F) \cap U_1} (\mathbf{x}) \left| \det D\mathbf{h} (\mathbf{x}) \right| dx$$

$$= \int_{U_1} \mathcal{X}_{\mathbf{h}^{-1}(F)} (\mathbf{x}) \left| \det D\mathbf{h} (\mathbf{x}) \right| dx = \int_{U_1} \mathcal{X}_F (\mathbf{h} (\mathbf{x})) \left| \det D\mathbf{h} (\mathbf{x}) \right| dx \tag{11.5}$$

It follows that (11.5) holds for $\mathcal{X}_F$ replaced with an arbitrary nonnegative Borel measurable simple function. Now if g is a nonnegative Borel measurable function, we may obtain g as the pointwise limit of an increasing sequence of nonnegative simple functions. Therefore, the monotone convergence theorem applies and we may write the following inequality for all g a nonnegative Borel measurable function.

$$\int_{V_1} g (\mathbf{y}) \, dy \leq \int_{U_1} g (\mathbf{h} (\mathbf{x})) \left| \det D\mathbf{h} (\mathbf{x}) \right| dx. \tag{11.6}$$

With this preparation, we are ready for the main result in this section, the change of variables formula.

Theorem 11.5 *Let U be an open set and let $\mathbf{h}$ be one to one, C^1, and $\det D\mathbf{h}(\mathbf{x}) \neq 0$ on U. For $V = \mathbf{h}(U)$ and $g \geq 0$ and Borel measurable,*

$$\int_V g(\mathbf{y})\, dy = \int_U g(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| dx. \tag{11.7}$$

Proof: By the inverse function theorem, $\mathbf{h}^{-1}$ is also a C^1 function. Therefore, using (11.6) on $\mathbf{x} = \mathbf{h}^{-1}(\mathbf{y})$, along with the chain rule and the property of determinants which states that the determinant of the product equals the product of the determinants,

$$\int_{V_1} g(\mathbf{y})\, dy \leq \int_{U_1} g(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| dx$$

$$\leq \int_{V_1} g(\mathbf{y}) \left|\det D\mathbf{h}\left(\mathbf{h}^{-1}(\mathbf{y})\right)\right| \left|\det D\mathbf{h}^{-1}(\mathbf{y})\right| dy$$

$$= \int_{V_1} g(\mathbf{y}) \left|\det D\left(\mathbf{h} \circ \mathbf{h}^{-1}\right)(\mathbf{y})\right| dy = \int_{V_1} g(\mathbf{y})\, dy$$

which shows the formula (11.7) holds for U_1 replacing U. To verify the theorem, let U_k be an increasing sequence of open sets whose union is U and whose closures are compact as in Theorem 3.32. Then from the above, (11.7) holds for U replaced with U_k and V replaced with V_k. Now (11.7) follows from the monotone convergence theorem. This proves the theorem.

11.1 Generalizations

In this section we give some generalizations of the theorem of the last section. The first generalization will be to remove the assumption that $\det D\mathbf{h}(\mathbf{x}) \neq 0$. This is accomplished through the use of the following fundamental lemma known as Sard's lemma. Actually the following lemma is a special case of Sard's lemma.

Lemma 11.6 *(Sard) Let U be an open set in $\mathbb{R}^n$ and let $\mathbf{h} : U \to \mathbb{R}^n$ be C^1. Let*

$$Z \equiv \{\mathbf{x} \in U : \det D\mathbf{h}(x) = 0\}.$$

Then $m_n(\mathbf{h}(Z)) = 0$.

Proof: Let $\{U_k\}_{k=1}^{\infty}$ be the increasing sequence of open sets whose closures are compact and whose union equals U which exists by Theorem 3.32 and let $Z_k \equiv Z \cap \overline{U_k}$. We will show that $\mathbf{h}(Z_k)$ has measure zero. Let W be an open set contained in U_{k+1} which contains Z_k and satisfies

$$m_n(Z_k) + \epsilon > m_n(W)$$

where we will always assume $\epsilon < 1$. Let

$$r = dist\left(\overline{U_k}, U_{k+1}^C\right)$$

and let $r_1 > 0$ be the constant of Lemma 11.1 such that whenever $\mathbf{x} \in \overline{U_k}$ and $0 < ||\mathbf{v}|| \leq r_1$,

$$||\mathbf{h}(\mathbf{x}+\mathbf{v}) - \mathbf{h}(\mathbf{x}) - D\mathbf{h}(\mathbf{x})\mathbf{v}|| < \epsilon\, ||\mathbf{v}||. \tag{11.8}$$

Now let

$$W = \cup_{i=1}^{\infty} \widetilde{R_i}$$

where the $\left\{\widetilde{R_i}\right\}$ are disjoint half open cubes from $\mathcal{D}_q$ where q is chosen so large that for each i,

$$diam\left(\widetilde{R_i}\right) \equiv \sup\left\{||\mathbf{x}-\mathbf{y}|| : \mathbf{x},\mathbf{y}\in\widetilde{R_i}\right\} < r_1.$$

Denote by $\{R_i\}$ those cubes which have nonempty intersection with Z_k, let d_i be the diameter of R_i, and let $\mathbf{z}_i$ be a point in $R_i\cap Z_k$. Since $\mathbf{z}_i \in Z_k$, it follows $D\mathbf{h}(\mathbf{z}_i)\,B(\mathbf{0},d_i) = D_i$ where D_i is contained in a subspace of $\mathbb{R}^n$ having dimension $p \leq n-1$ and the diameter of D_i is no larger than $C_k d_i$ where

$$C_k \geq \max\{||D\mathbf{h}(\mathbf{x})|| : \mathbf{x}\in Z_k\}$$

Then by (11.8), if $\mathbf{z}\in R_i$,

$$\mathbf{h}(\mathbf{z}) - \mathbf{h}(\mathbf{z}_i) \in D_i + B(\mathbf{0},\epsilon d_i)$$

$$\subseteq D_i + B_0\left(\mathbf{0},\epsilon\delta^{-1}d_i\right).$$

(Recall B_0 refers to the ball taken with respect to the usual norm.) Therefore, by Theorem 2.24, there exists an orthogonal linear transformation, Q, which preserves distances measured with respect to the usual norm and $QD_i \subseteq \mathbb{R}^{n-1}$. Therefore, for $\mathbf{z}\in R_i$,

$$Q\mathbf{h}(\mathbf{z}) - Q\mathbf{h}(\mathbf{z}_i) \in QD_i + B_0\left(\mathbf{0},\epsilon\delta^{-1}d_i\right).$$

Refering to (11.1), it follows that

$$m_n(\mathbf{h}(R_i)) = |\det(Q)|\, m_n(\mathbf{h}(R_i)) = m_n(Q\mathbf{h}(R_i))$$

$$\leq m_n\left(QD_i + B_0\left(\mathbf{0},\epsilon\delta^{-1}d_i\right)\right) \leq m_n\left(QD_i + B\left(\mathbf{0},\epsilon\delta^{-1}\Delta d_i\right)\right)$$

$$\leq \left(C_k d_i + 2\epsilon\delta^{-1}\Delta d_i\right)^{n-1} 2\epsilon\delta^{-1}\Delta d_i \leq C_{k,n} m_n(R_i)\,\epsilon.$$

Therefore,

$$m_n(\mathbf{h}(Z_k)) \leq \sum_{i=1}^{\infty} m(\mathbf{h}(R_i)) \leq C_{k,n}\epsilon\sum_{i=1}^{\infty} m_n(R_i) \leq C_{k,n}\epsilon m_n(W)$$

$$\leq C_{k,n}\epsilon\,(m_n(Z_k)+\epsilon).$$

Since $\epsilon > 0$ is arbitrary, this shows $m_n(\mathbf{h}(Z_k)) = 0$. Now this implies

$$m_n(\mathbf{h}(Z)) = \lim_{k\to\infty} m_n(\mathbf{h}(Z_k)) = 0$$

and this proves the lemma.

With Sard's lemma it is easy to remove the assumption that $\det D\mathbf{h}(\mathbf{x}) \neq 0$.

Theorem 11.7 *Let U be an open set in $\mathbb{R}^n$ and let $\mathbf{h}$ be a one to one mapping from U to $\mathbb{R}^n$ which is C^1. Then if $g \geq 0$ is a Borel measurable function,*

$$\int_{\mathbf{h}(U)} g(\mathbf{y})\,dy = \int_U g(\mathbf{h}(\mathbf{x}))\,|\det D\mathbf{h}(\mathbf{x})|\,dx$$

and every function which needs to be measurable, is. In particular the integrand of the second integral is Lebesgue measurable.

Proof: We observe first that $\mathbf{h}(U)$ is measurable because, thanks to Theorem 3.32,

$$U = \cup_{i=1}^{\infty} \overline{U_i}$$

where $\overline{U_i}$ is compact. Therefore, $\mathbf{h}(U) = \cup_{i=1}^{\infty} \mathbf{h}\left(\overline{U_i}\right)$, a Borel set. The function, $\mathbf{x} \rightarrow g(\mathbf{h}(\mathbf{x}))$ is also measurable from Lemma 11.3. Letting Z be described above, it follows by continuity of the derivative, Z is a closed set. Thus the inverse function theorem applies and we may say that $\mathbf{h}(U \setminus Z)$ is an open set. Therefore, if $g \geq 0$ is a Borel measurable function,

$$\int_{\mathbf{h}(U)} g(\mathbf{y})\, dy = \int_{\mathbf{h}(U \setminus Z)} g(\mathbf{y})\, dy =$$

$$\int_{U \setminus Z} g(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| dx = \int_U g(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| dx,$$

the middle equality holding by Theorem 11.5 and the first equality holding by Sard's lemma. This proves the theorem.

It is possible to extend this theorem to the case when g is only assumed to be Lebesgue measurable. We do this next using the following lemma.

Lemma 11.8 *Suppose $0 \leq f \leq g$, g is measurable with respect to the σ algebra of Lebesgue measurable sets, $\mathcal{S}$, and $g = 0$ a.e. Then f is also Lebesgue measurable.*

Proof: Let $a \geq 0$. Then

$$f^{-1}((a, \infty]) \subseteq g^{-1}((a, \infty]) \subseteq \{\mathbf{x} : g(\mathbf{x}) > 0\},$$

a set of measure zero. Therefore, by completeness of Lebesgue measure, it follows $f^{-1}((a, \infty])$ is also Lebesgue measurable. This proves the lemma.

To extend Theorem 11.7 to the case where g is only Lebesgue measurable, first suppose F is a Lebesgue measurable set which has measure zero. Then by regularity of Lebesgue measure, there exists a Borel set, $G \supseteq F$ such that $m_n(G) = 0$ also. By Theorem 11.7,

$$0 = \int_{\mathbf{h}(U)} \mathcal{X}_G(\mathbf{y})\, dy = \int_U \mathcal{X}_G(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| dx$$

and the integrand of the second integral is Lebesgue measurable. Therefore, this measurable function,

$$\mathbf{x} \rightarrow \mathcal{X}_G(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right|,$$

is equal to zero a.e. Thus,

$$0 \leq \mathcal{X}_F(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| \leq \mathcal{X}_G(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right|$$

and Lemma 11.8 implies the function $\mathbf{x} \rightarrow \mathcal{X}_F(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right|$ is also Lebesgue measurable and

$$\int_U \mathcal{X}_F(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right| dx = 0.$$

Now let F be an arbitrary bounded Lebesgue measurable set and use the regularity of the measure to obtain $G \supseteq F \supseteq E$ where G and E are Borel sets with the property that $m_n(G \setminus E) = 0$. Then from what was just shown,

$$\mathbf{x} \rightarrow \mathcal{X}_{F \setminus E}(\mathbf{h}(\mathbf{x})) \left|\det D\mathbf{h}(\mathbf{x})\right|$$

is Lebesgue measurable and

$$\int_U \mathcal{X}_{F\setminus E}(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right| dx = 0.$$

It follows since

$$\mathcal{X}_F(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right| = \mathcal{X}_{F\setminus E}(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right| + \mathcal{X}_E(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right|$$

and $\mathbf{x} \to \mathcal{X}_E(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right|$ is measurable, that

$$\mathbf{x} \to \mathcal{X}_F(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right|$$

is measurable and so

$$\int_U \mathcal{X}_F(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right| dx = \int_U \mathcal{X}_E(\mathbf{h}(\mathbf{x}))\left|\det D\mathbf{h}(\mathbf{x})\right| dx$$

$$= \int_{\mathbf{h}(U)} \mathcal{X}_E(\mathbf{y})\, dy = \int_{\mathbf{h}(U)} \mathcal{X}_F(\mathbf{y})\, dy. \tag{11.9}$$

To obtain the result in the case where F is an arbitrary Lebesgue measurable set, let

$$F_k \equiv F \cap B(0,k)$$

and apply (11.9) to F_k and then let $k \to \infty$ and use the monotone convergence theorem to obtain (11.9) for a general Lebesgue measurable set, F. It follows from this formula that we may replace $\mathcal{X}_F$ with an arbitrary nonnegative simple function. Now if g is an arbitrary nonnegative Lebesgue measurable function, we obtain g as a pointwise limit of an increasing sequence of nonnegative simple functions and use the monotone convergence theorem again to obtain (11.9) with $\mathcal{X}_F$ replaced with g. This proves the following corollary.

Corollary 11.9 *The conclusion of Theorem 11.7 hold if g is only assumed to be Lebesgue measurable.*

Corollary 11.10 *If $g \in L^1(\mathbf{h}(U), \mathcal{S}, m_n)$, then the conclusion of Theorem 11.7 holds.*

Proof: Apply Corollary 11.9 to the positive and negative parts of the real and complex parts of g.

There are many other ways to prove the above results. To see some alternative presentations, see [24], [19], or [11].

Next we give a version of this theorem which considers the case where $\mathbf{h}$ is only C^1, not necessarily 1-1. For

$$U_+ \equiv \{\mathbf{x} \in U : \left|\det D\mathbf{h}(x)\right| > 0\}$$

and Z the set where $\left|\det D\mathbf{h}(\mathbf{x})\right| = 0$, Lemma 11.6 implies $m(\mathbf{h}(Z)) = 0$. For $\mathbf{x} \in U_+$, the inverse function theorem implies there exists an open set $B_{\mathbf{x}}$ such that

$$\mathbf{x} \in B_{\mathbf{x}} \subseteq U_+, \ \mathbf{h} \text{ is } 1-1 \text{ on } B_{\mathbf{x}}.$$

Let $\{B_i\}$ be a countable subset of $\{B_{\mathbf{x}}\}_{\mathbf{x}\in U_+}$ such that

$$U_+ = \cup_{i=1}^{\infty} B_i.$$

Let $E_1 = B_1$. If $E_1, \cdots, E_k$ have been chosen, $E_{k+1} = B_{k+1} \setminus \cup_{i=1}^{k} E_i$. Thus

$$\cup_{i=1}^{\infty} E_i = U_+, \ \ \mathbf{h} \text{ is } 1-1 \text{ on } E_i, \ \ E_i \cap E_j = \emptyset,$$

and each E_i is a Borel set contained in the open set B_i. Now we define

$$n(\mathbf{y}) = \sum_{i=1}^{\infty} \mathcal{X}_{\mathbf{h}(E_i)}(\mathbf{y}) + \mathcal{X}_{\mathbf{h}(Z)}(\mathbf{y}).$$

Thus $n(\mathbf{y}) \geq 0$ and is Borel measurable.

Lemma 11.11 *Let $F \subseteq \mathbf{h}(U)$ be measurable. Then*

$$\int_{\mathbf{h}(U)} n(\mathbf{y})\mathcal{X}_F(\mathbf{y})dy = \int_U \mathcal{X}_F(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx.$$

Proof: Using Lemma 11.6 and the Monotone Convergence Theorem or Fubini's Theorem,

$$\begin{aligned}
\int_{\mathbf{h}(U)} n(\mathbf{y})\mathcal{X}_F(\mathbf{y})dy &= \int_{\mathbf{h}(U)} \left(\sum_{i=1}^{\infty} \mathcal{X}_{\mathbf{h}(E_i)}(\mathbf{y}) + \mathcal{X}_{\mathbf{h}(Z)}(\mathbf{y})\right)\mathcal{X}_F(\mathbf{y})dy \\
&= \sum_{i=1}^{\infty} \int_{\mathbf{h}(U)} \mathcal{X}_{\mathbf{h}(E_i)}(\mathbf{y})\mathcal{X}_F(\mathbf{y})dy \\
&= \sum_{i=1}^{\infty} \int_{\mathbf{h}(B_i)} \mathcal{X}_{\mathbf{h}(E_i)}(\mathbf{y})\mathcal{X}_F(\mathbf{y})dy \\
&= \sum_{i=1}^{\infty} \int_{B_i} \mathcal{X}_{E_i}(\mathbf{x})\mathcal{X}_F(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx \\
&= \sum_{i=1}^{\infty} \int_{U} \mathcal{X}_{E_i}(\mathbf{x})\mathcal{X}_F(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx \\
&= \int_{U} \sum_{i=1}^{\infty} \mathcal{X}_{E_i}(\mathbf{x})\mathcal{X}_F(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx
\end{aligned}$$

$$= \int_{U_+} \mathcal{X}_F(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx = \int_U \mathcal{X}_F(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx.$$

This proves the lemma.

Definition 11.12 *For $\mathbf{y} \in \mathbf{h}(U)$,*

$$\#(\mathbf{y}) \equiv |\mathbf{h}^{-1}(\mathbf{y})|.$$

Thus $\#(\mathbf{y}) \equiv$ number of elements in $\mathbf{h}^{-1}(\mathbf{y})$.

We observe that

$$\#(\mathbf{y}) = n(\mathbf{y}) \quad \text{a.e.} \tag{11.10}$$

And thus $\#$ is a measurable function. This follows because $n(\mathbf{y}) = \#(\mathbf{y})$ if $\mathbf{y} \notin \mathbf{h}(Z)$, a set of measure 0.

Theorem 11.13 *Let $g \geq 0$, g measurable, and let $\mathbf{h}$ be $C^1(U)$. Then*

$$\int_{\mathbf{h}(U)} \#(\mathbf{y})g(\mathbf{y})dy = \int_U g(\mathbf{h}(\mathbf{x}))|\det D\mathbf{h}(\mathbf{x})|dx. \tag{11.11}$$

Proof: From (11.10) and Lemma 11.11, (11.11) holds for all g, a nonnegative simple function. Approximating an arbitrary $g \geq 0$ with an increasing pointwise convergent sequence of simple functions yields (11.11) for $g \geq 0$, g measurable. This proves the theorem.

11.2 Exercises

1. The gamma function is defined by

$$\Gamma(\alpha) \equiv \int_0^\infty e^{-t} t^{\alpha-1}$$

for $\alpha > 0$. Show this is a well defined function of these values of α and verify that $\Gamma(\alpha+1) = \Gamma(\alpha)\alpha$. What is $\Gamma(n+1)$ for n a nonnegative integer?

2. Show that $\int_{-\infty}^{\infty} e^{\frac{-s^2}{2}} ds = \sqrt{2\pi}$. **Hint:** Denote this integral by I and observe that $I^2 = \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dxdy$. Then change variables to polar coordinates, $x = r\cos(\theta)$, $y = r\sin\theta$.

3. $\uparrow$ Now that you know what the gamma function is, consider in the formula for $\Gamma(\alpha+1)$ the following change of variables. $t = \alpha + \alpha^{1/2}s$. Then in terms of the new variable, s, the formula for $\Gamma(\alpha+1)$ is

$$e^{-\alpha}\alpha^{\alpha+\frac{1}{2}} \int_{-\sqrt{\alpha}}^{\infty} e^{-\sqrt{\alpha}s}\left(1 + \frac{s}{\sqrt{\alpha}}\right)^{\alpha} ds = e^{-\alpha}\alpha^{\alpha+\frac{1}{2}} \int_{-\sqrt{\alpha}}^{\infty} e^{\alpha\left[\ln\left(1+\frac{s}{\sqrt{\alpha}}\right) - \frac{s}{\sqrt{\alpha}}\right]} ds$$

Show the integrand converges to $e^{-\frac{s^2}{2}}$. Show that then

$$\lim_{\alpha\to\infty} \frac{\Gamma(\alpha+1)}{e^{-\alpha}\alpha^{\alpha+(1/2)}} = \int_{-\infty}^{\infty} e^{\frac{-s^2}{2}} ds = \sqrt{2\pi}.$$

Hint: You will need to obtain a dominating function for the integral so that you can use the dominated convergence theorem. You might try considering $s \in (-\sqrt{\alpha}, \sqrt{\alpha})$ first and consider something like $e^{1-(s^2/4)}$ on this interval. Then look for another function for $s > \sqrt{\alpha}$. This formula is known as Stirling's formula.

The L^p Spaces

12.1 Basic inequalities and properties

The Lebesgue integral makes it possible to define and prove theorems about the space of functions described below. These L^p spaces are very useful in applications of real analysis and this chapter is about these spaces. In what follows $(\Omega, \mathcal{S}, \mu)$ will be a measure space.

Definition 12.1 *Let $1 \leq p < \infty$. We define*

$$L^p(\Omega) \equiv \{f \ : f \ \ \text{is measurable and} \int_\Omega |f(\omega)|^p d\mu < \infty\}$$

and

$$||f||_{L^p} \equiv \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}} \equiv ||f||_p.$$

In fact $|| \ ||_p$ is a norm if things are interpreted correctly. First we need to obtain Holder's inequality. We will always use the following convention for each $p > 1$.

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Often one uses p' instead of q in this context.

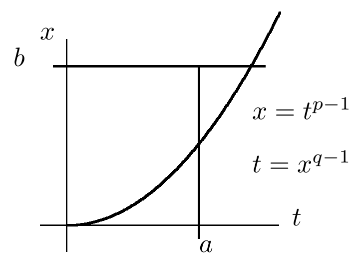
Theorem 12.2 *(Holder's inequality) If f and g are measurable functions, then if $p > 1$,*

$$\int |f| \ |g| \ d\mu \leq \left(\int |f|^p d\mu\right)^{\frac{1}{p}} \left(\int |g|^q d\mu\right)^{\frac{1}{q}}. \tag{12.1}$$

Proof: To begin with, we prove Young's inequality.

Lemma 12.3 *If $0 \leq a, b$ then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.*

Proof: Consider the following picture:



$$ab \leq \int_0^a t^{p-1} dt + \int_0^b x^{q-1} dx = \frac{a^p}{p} + \frac{b^q}{q}.$$

Note equality occurs when $a^p = b^q$.

If either $\int |f|^p d\mu$ or $\int |g|^p d\mu$ equals ∞ or 0, the inequality (12.1) is obviously valid. Therefore assume both of these are less than ∞ and not equal to 0. By the lemma,

$$\int \frac{|f|}{||f||_p} \frac{|g|}{||g||_q} d\mu \leq \frac{1}{p} \int \frac{|f|^p}{||f||_p^p} d\mu + \frac{1}{q} \int \frac{|g|^q}{||g||_q^q} d\mu = 1.$$

Hence,

$$\int |f| \, |g| \, d\mu \leq ||f||_p \, ||g||_q.$$

This proves Holder's inequality.

Corollary 12.4 *(Minkowski inequality) Let* $1 \leq p < \infty$. *Then*

$$||f + g||_p \leq ||f||_p + ||g||_p. \tag{12.2}$$

Proof: If $p = 1$, this is obvious. Let $p > 1$. We can assume $||f||_p$ and $||g||_p < \infty$ and $||f + g||_p \neq 0$ or there is nothing to prove. Therefore,

$$\int |f + g|^p d\mu \leq 2^{p-1} \left(\int |f|^p + |g|^p d\mu \right) < \infty.$$

Now

$$\int |f + g|^p d\mu \leq$$

$$\begin{aligned}
& \int |f + g|^{p-1} |f| d\mu + \int |f + g|^{p-1} |g| d\mu \\
= \quad & \int |f + g|^{\frac{p}{q}} |f| d\mu + \int |f + g|^{\frac{p}{q}} |g| d\mu \\
\leq \quad & (\int |f + g|^p d\mu)^{\frac{1}{q}} (\int |f|^p d\mu)^{\frac{1}{p}} + (\int |f + g|^p d\mu)^{\frac{1}{q}} (\int |g|^p d\mu)^{\frac{1}{p}}.
\end{aligned}$$

Dividing both sides by $(\int |f + g|^p d\mu)^{\frac{1}{q}}$ yields (12.2). This proves the corollary.

This shows that if $f, g \in L^p$, then $f+g \in L^p$. Also, it is clear that if a is a constant and $f \in L^p$, then $af \in L^p$. Hence L^p is a vector space. Also we have the following from the Minkowski inequality and the definition of $|| \, ||_p$.

a.) $||f||_p \geq 0, ||f||_p = 0$ if and only if $f = 0$ a.e.

b.) $||af||_p = |a| \, ||f||_p$ if a is a scalar.

c.) $||f+g||_p \leq ||f||_p + ||g||_p$.

We see that $|| \, ||_p$ would be a norm if $||f||_p = 0$ implied $f = 0$. If we agree to identify all functions in L^p that differ only on a set of measure zero, then $|| \, ||_p$ is a norm and L^p is a normed vector space. We will do so from now on.

Definition 12.5 *A complete normed linear space is called a Banach space.*

Next we show that L^p is a Banach space.

Theorem 12.6 *The following hold for $L^p(\Omega)$*

a.) $L^p(\Omega)$ is complete.

b.) If $\{f_n\}$ is a Cauchy sequence in $L^p(\Omega)$, then there exists $f \in L^p(\Omega)$ and a subsequence which converges a.e. to $f \in L^p(\Omega)$, and $||f_n - f||_p \to 0$.

Proof: Let $\{f_n\}$ be a Cauchy sequence in $L^p(\Omega)$. This means that for every $\varepsilon > 0$ there exists N such that if $n, m \geq N$, then $||f_n - f_m||_p < \varepsilon$. Now we will select a subsequence. Let n_1 be such that $||f_n - f_m||_p < 2^{-1}$ whenever $n, m \geq n_1$. Let n_2 be such that $n_2 > n_1$ and $||f_n - f_m||_p < 2^{-2}$ whenever $n, m \geq n_2$. If $n_1, \cdots, n_k$ have been chosen, let $n_{k+1} > n_k$ and whenever $n, m \geq n_{k+1}, ||f_n - f_m||_p < 2^{-(k+1)}$. The subsequence will be $\{f_{n_k}\}$. Thus, $||f_{n_k} - f_{n_{k+1}}||_p < 2^{-k}$. Let

$$g_{k+1} = f_{n_{k+1}} - f_{n_k}.$$

Then by the Minkowski inequality,

$$\infty > \sum_{k=1}^{\infty} ||g_{k+1}||_p \geq \sum_{k=1}^{m} ||g_{k+1}||_p \geq \left|\left| \sum_{k=1}^{m} |g_{k+1}| \right|\right|_p$$

for all m. It follows that

$$\int \left(\sum_{k=1}^{m} |g_{k+1}| \right)^p d\mu \leq \left(\sum_{k=1}^{\infty} ||g_{k+1}||_p \right)^p < \infty \tag{12.3}$$

for all m and so the monotone convergence theorem implies that the sum up to m in (12.3) can be replaced by a sum up to ∞. Thus,

$$\sum_{k=1}^{\infty} |g_{k+1}(x)| < \infty \text{ a.e. } x.$$

Therefore, $\sum_{k=1}^{\infty} g_{k+1}(x)$ converges for a.e. x and for such x, let

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} g_{k+1}(x).$$

Note that $\sum_{k=1}^{m} g_{k+1}(x) = f_{n_{m+1}}(x) - f_{n_1}(x)$. Therefore $\lim_{k\to\infty} f_{n_k}(x) = f(x)$ for all $x \notin E$ where $\mu(E) = 0$. If we redefine f_{n_k} to equal 0 on E and let $f(x) = 0$ for $x \in E$, it then follows that $\lim_{k\to\infty} f_{n_k}(x) = f(x)$ for all x. By Fatou's lemma,

$$||f - f_{n_k}||_p \leq \liminf_{m\to\infty} ||f_{n_m} - f_{n_k}||_p \leq \sum_{i=k}^{\infty} \left|\left| f_{n_{i+1}} - f_{n_i} \right|\right|_p \leq 2^{-(k-1)}.$$

Therefore, $f \in L^p(\Omega)$ because

$$||f||_p \leq ||f - f_{n_k}||_p + ||f_{n_k}||_p < \infty,$$

and $\lim_{k\to\infty} ||f_{n_k} - f||_p = 0$. This proves b.). To see that the original Cauchy sequence converges to f in $L^p(\Omega)$, we write

$$||f - f_n||_p \leq ||f - f_{n_k}||_p + ||f_{n_k} - f_n||_p.$$

If $\varepsilon > 0$ is given, let $2^{-(k-1)} < \frac{\varepsilon}{2}$. Then if $n > n_k$,

$$||f - f_n|| < 2^{-(k-1)} + 2^{-k} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves part a.) and completes the proof of the theorem.

In working with the L^p spaces, the following inequality also known as Minkowski's inequality is very useful.

Theorem 12.7 *Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{F}, \lambda)$ be σ-finite measure spaces and let f be product measurable. Then the following inequality is valid for $p \geq 1$.*

$$\int_X \left(\int_Y |f(x,y)|^p \, d\lambda\right)^{\frac{1}{p}} d\mu \geq \left(\int_Y (\int_X |f(x,y)| \, d\mu)^p d\lambda\right)^{\frac{1}{p}}. \tag{12.4}$$

Proof: Let $X_n \uparrow X,\ Y_n \uparrow Y,\ \lambda(Y_n) < \infty,\ \mu(X_n) < \infty$, and let

$$f_m(x,y) = \begin{cases} f(x,y) & \text{if } |f(x,y)| \leq m, \\ m & \text{if } |f(x,y)| > m. \end{cases}$$

Thus

$$\left(\int_{Y_n} (\int_{X_k} |f_m(x,y)| d\mu)^p d\lambda\right)^{\frac{1}{p}} < \infty.$$

Let

$$J(y) = \int_{X_k} |f_m(x,y)| d\mu.$$

Then

$$\begin{aligned} \int_{Y_n} \left(\int_{X_k} |f_m(x,y)| d\mu\right)^p d\lambda &= \int_{Y_n} J(y)^{p-1} \int_{X_k} |f_m(x,y)| d\mu \, d\lambda \\ &= \int_{X_k} \int_{Y_n} J(y)^{p-1} |f_m(x,y)| d\lambda \, d\mu \end{aligned}$$

by Fubini's theorem. Now we apply Holder's inequality and recall $p - 1 = \frac{p}{q}$. This yields

$$\begin{aligned} &\int_{Y_n} \left(\int_{X_k} |f_m(x,y)| d\mu\right)^p d\lambda \\ \leq\ &\int_{X_k} \left(\int_{Y_n} J(y)^p d\lambda\right)^{\frac{1}{q}} \left(\int_{Y_n} |f_m(x,y)|^p d\lambda\right)^{\frac{1}{p}} d\mu \\ =\ &\left(\int_{Y_n} J(y)^p d\lambda\right)^{\frac{1}{q}} \int_{X_k} \left(\int_{Y_n} |f_m(x,y)|^p d\lambda\right)^{\frac{1}{p}} d\mu \end{aligned}$$

$$= \left(\int_{Y_n} (\int_{X_k} |f_m(x,y)| d\mu)^p d\lambda\right)^{\frac{1}{q}} \int_{X_k} \left(\int_{Y_n} |f_m(x,y)|^p d\lambda\right)^{\frac{1}{p}} d\mu.$$

Therefore,

$$\left(\int_{Y_n} \left(\int_{X_k} |f_m(x,y)| d\mu\right)^p d\lambda\right)^{\frac{1}{p}} \leq \int_{X_k} \left(\int_{Y_n} |f_m(x,y)|^p d\lambda\right)^{\frac{1}{p}} d\mu. \tag{12.5}$$

To obtain (12.4) let $m \to \infty$ and use the Monotone Convergence theorem to replace f_m by f in (12.5). Next let $k \to \infty$ and use the same theorem to replace X_k with X. Finally let $n \to \infty$ and use the Monotone Convergence theorem again to replace Y_n with Y. This yields (12.4).

Next, we develop some properties of the L^p spaces.

12.2 Density of simple functions

Theorem 12.8 *Let $p \geq 1$ and let $(\Omega, \mathcal{S}, \mu)$ be a measure space. Then the simple functions are dense in $L^p(\Omega)$.*

Proof: By breaking an arbitrary function into real and imaginary parts and then considering the positive and negative parts of these, we see that there is no loss of generality in assuming f has values in $[0, \infty]$. By Theorem 5.31, there is an increasing sequence of simple functions, $\{s_n\}$, converging pointwise to $f(x)^p$. Let $t_n(x) = s_n(x)^{\frac{1}{p}}$. Thus, $t_n(x) \uparrow f(x)$. Now

$$|f(x) - t_n(x)| \leq |f(x)|.$$

By the Dominated Convergence theorem, we may conclude

$$0 = \lim_{n \to \infty} \int |f(x) - t_n(x)|^p d\mu.$$

Thus simple functions are dense in L^p.

Recall that for Ω a topological space, $C_c(\Omega)$ is the space of continuous functions with compact support in Ω. Also recall the following definition.

Definition 12.9 *Let $(\Omega, \mathcal{S}, \mu)$ be a measure space and suppose (Ω, τ) is also a topological space. Then $(\Omega, \mathcal{S}, \mu)$ is called a regular measure space if the σ-algebra of Borel sets is contained in $\mathcal{S}$ and for all $E \in \mathcal{S}$,*

$$\mu(E) = \inf\{\mu(V) : V \supseteq E \text{ and } V \text{ open}\}$$

and

$$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ and } K \text{ is compact }\}.$$

For example Lebesgue measure is an example of such a measure.

Lemma 12.10 *Let Ω be a locally compact metric space and let K be a compact subset of V, an open set. Then there exists a continuous function $f : \Omega \to [0,1]$ such that $f(x) = 1$ for all $x \in K$ and $spt(f)$ is a compact subset of V.*

Proof: Let $K \subseteq W \subseteq \overline{W} \subseteq V$ and $\overline{W}$ is compact. Define f by

$$f(x) = \frac{dist(x, W^C)}{dist(x, K) + dist(x, W^C)}.$$

It is not necessary to be in a metric space to do this. You can accomplish the same thing using Urysohn's lemma.

Theorem 12.11 *Let $(\Omega, \mathcal{S}, \mu)$ be a regular measure space as in Definition 12.9 where the conclusion of Lemma 12.10 holds. Then $C_c(\Omega)$ is dense in $L^p(\Omega)$.*

Proof: Let $f \in L^p(\Omega)$ and pick a simple function, s, such that $||s - f||_p < \frac{\varepsilon}{2}$ where $\varepsilon > 0$ is arbitrary. Let

$$s(x) = \sum_{i=1}^{m} c_i \mathcal{X}_{E_i}(x)$$

where $c_1, \cdots, c_m$ are the distinct nonzero values of s. Thus the E_i are disjoint and $\mu(E_i) < \infty$ for each i. Therefore there exist compact sets, K_i and open sets, V_i, such that $K_i \subseteq E_i \subseteq V_i$ and

$$\sum_{i=1}^{m} |c_i| \mu(V_i \setminus K_i)^{\frac{1}{p}} < \frac{\varepsilon}{2}.$$

Let $h_i \in C_c(\Omega)$ satisfy

$$\begin{aligned} h_i(x) &= 1 \text{ for } x \in K_i \\ spt(h_i) &\subseteq V_i. \end{aligned}$$

Let

$$g = \sum_{i=1}^{m} c_i h_i.$$

Then by Minkowski's inequality,

$$\begin{aligned} ||g - s||_p &\leq \left(\int_{\Omega} (\sum_{i=1}^{m} |c_i| \; |h_i(x) - \mathcal{X}_{E_i}(x)|)^p d\mu \right)^{\frac{1}{p}} \\ &\leq \sum_{i=1}^{m} \left(\int_{\Omega} |c_i|^p |h_i(x) - \mathcal{X}_{E_i}(x)|^p d\mu \right)^{\frac{1}{p}} \\ &\leq \sum_{i=1}^{m} |c_i| \mu(V_i \setminus K_i)^{\frac{1}{p}} < \frac{\varepsilon}{2}. \end{aligned}$$

Therefore,

$$||f - g||_p \leq ||f - s||_p + ||s - g||_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves the theorem.

12.3 Continuity of translation

Definition 12.12 *Let f be a function defined on $U \subseteq \mathbb{R}^n$ and let $\mathbf{w} \in \mathbb{R}^n$. Then $f_{\mathbf{w}}$ will be the function defined on $\mathbf{w} + U$ by*

$$f_{\mathbf{w}}(\mathbf{x}) = f(\mathbf{x} - \mathbf{w}).$$

Theorem 12.13 *(Continuity of translation in L^p) Let $f \in L^p(\mathbb{R}^n)$ with $\mu = m$, Lebesgue measure. Then*

$$\lim_{||\mathbf{w}|| \to 0} ||f_{\mathbf{w}} - f||_p = 0.$$

Proof: Let $\varepsilon > 0$ be given and let $g \in C_c(\mathbb{R}^n)$ with $||g - f||_p < \frac{\varepsilon}{3}$. Since Lebesgue measure is translation invariant $(m(\mathbf{w} + E) = m(E))$, $||g_{\mathbf{w}} - f_{\mathbf{w}}||_p = ||g - f||_p < \frac{\varepsilon}{3}$. Therefore

$$\begin{aligned} ||f - f_{\mathbf{w}}||_p &\leq ||f - g||_p + ||g - g_{\mathbf{w}}||_p + ||g_{\mathbf{w}} - f_{\mathbf{w}}|| \\ &< \frac{2\varepsilon}{3} + ||g - g_{\mathbf{w}}||_p. \end{aligned}$$

But $\lim_{|\mathbf{w}| \to 0} g_{\mathbf{w}}(\mathbf{x}) = g(\mathbf{x})$ uniformly in $\mathbf{x}$ because g is uniformly continuous. Therefore, whenever $|\mathbf{w}|$ is small enough, $||g - g_{\mathbf{w}}||_p < \frac{\varepsilon}{3}$. Thus $||f - f_{\mathbf{w}}||_p < \varepsilon$ whenever $|\mathbf{w}|$ is sufficiently small. This proves the theorem.

12.4 Separability

Theorem 12.14 *For $p \geq 1$, $L^p(\mathbb{R}^n, m)$ is separable. This means there exists a countable set, $\mathcal{D}$, such that if $f \in L^p(\mathbb{R}^n)$ and $\varepsilon > 0$, there exists $g \in \mathcal{D}$ such that $||f - g||_p < \varepsilon$.*

Proof: Let Q be all functions of the form $c\mathcal{X}_{[\mathbf{a},\mathbf{b})}$ where

$$[\mathbf{a}, \mathbf{b}) \equiv [a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_n, b_n),$$

and both a_i, b_i are rational, while c has rational real and imaginary parts. Let $\mathcal{D}$ be the set of all finite sums of functions in Q. Thus, $\mathcal{D}$ is countable. We now show that $\mathcal{D}$ is dense in $L^p(\mathbb{R}^n, m)$. To see this, we need to show that for every $f \in L^p(\mathbb{R}^n)$, there exists an element of $\mathcal{D}$, s such that $||s - f||_p < \varepsilon$. By Theorem 12.11 we can assume without loss of generality that $f \in C_c(\mathbb{R}^n)$. Let $\mathcal{P}_m$consist of all sets of the form $[\mathbf{a}, \mathbf{b})$ where $a_i = j2^{-m}$and $b_i = (j+1)2^{-m}$ for j an integer. Thus $\mathcal{P}_m$consists of a tiling of $\mathbb{R}^n$ into half open rectangles having diameters $2^{-m}n^{\frac{1}{2}}$. There are countably many of these rectangles; so, let $\mathcal{P}_m = \{[\mathbf{a}_i, \mathbf{b}_i)\}$ and $\mathbb{R}^n = \cup_{i=1}^{\infty}[\mathbf{a}_i, \mathbf{b}_i)$. Let c_i^mbe complex numbers with rational real and imaginary parts satisfying

$$|f(\mathbf{a}_i) - c_i^m| < 5^{-m},$$

$$|c_i^m| \leq |f(\mathbf{a}_i)|. \tag{12.6}$$

Let $s_m(x) = \sum_{i=1}^{\infty} c_i^m \mathcal{X}_{[\mathbf{a}_i, \mathbf{b}_i)}$. Since $f(\mathbf{a}_i) = 0$ except for finitely many values of i, (12.6) implies $s_m \in \mathcal{D}$. It is also clear that, since f is uniformly continuous, $\lim_{m\to\infty} s_m(x) = f(x)$ uniformly in x. Hence $\lim_{x\to 0} ||s_m - f||_p = 0$.

Corollary 12.15 *Let Ω be any Lebesgue measurable subset of $\mathbb{R}^n$. Then $L^p(\Omega)$ is separable. Here the σ algebra of measurable sets will consist of all intersections of Lebesgue measurable sets with Ω and the measure will be m_n restricted to these sets.*

Proof: Let $\widetilde{\mathcal{D}}$ be the restrictions of $\mathcal{D}$ to Ω. If $f \in L^p(\Omega)$, let F be the zero extension of f to all of $\mathbb{R}^n$. Let $\varepsilon > 0$ be given. By Theorem 12.14 there exists $s \in \mathcal{D}$ such that $||F - s||_p < \varepsilon$. Thus

$$||s - f||_{L^p(\Omega)} \leq ||s - F||_{L^p(\mathbb{R}^n)} < \varepsilon$$

and so the countable set $\widetilde{\mathcal{D}}$ is dense in $L^p(\Omega)$.

12.5 Mollifiers and density of smooth functions

Definition 12.16 *Let U be an open subset of $\mathbb{R}^n$. $C_c^\infty(U)$ is the vector space of all infinitely differentiable functions which equal zero for all $\mathbf{x}$ outside of some compact set contained in U.*

Example 12.17 *Let* $U = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < 2\}$

$$\psi(x) = \begin{cases} \exp\left[\left(|\mathbf{x}|^2 - 1\right)^{-1}\right] & \text{if } |\mathbf{x}| < 1, \\ 0 \text{ if } |\mathbf{x}| \geq 1. \end{cases}$$

Then a little work shows $\psi \in C_c^\infty(U)$. *The following also is easily obtained.*

Lemma 12.18 *Let* U *be any open set. Then* $C_c^\infty(U) \neq \emptyset$.

Definition 12.19 *Let* $U = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < 1\}$. *A sequence* $\{\psi_m\} \subseteq C_c^\infty(U)$ *is called a mollifier (sometimes an approximate identity) if*

$$\psi_m(\mathbf{x}) \geq 0, \ \psi_m(\mathbf{x}) = 0, \ \text{if } |\mathbf{x}| \geq \frac{1}{m},$$

and $\int \psi_m(\mathbf{x}) = 1$.

As before, $\int f(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{y})$ will mean $\mathbf{x}$ is fixed and the function $\mathbf{y} \to f(\mathbf{x}, \mathbf{y})$ is being integrated. We may also write dx for $dm(x)$ in the case of Lebesgue measure.

Example 12.20 *Let*

$$\psi \in C_c^\infty(B(0,1)) \ \ (B(0,1) = \{\mathbf{x} : |\mathbf{x}| < 1\})$$

with $\psi(\mathbf{x}) \geq 0$ *and* $\int \psi dm = 1$. *Let* $\psi_m(\mathbf{x}) = c_m \psi(m\mathbf{x})$ *where* c_m *is chosen in such a way that* $\int \psi_m dm = 1$. *By the change of variables theorem we see that* $c_m = m^n$.

Definition 12.21 *A function,* f, *is said to be in* $L^1_{loc}(\mathbb{R}^n, \mu)$ *if* f *is* μ *measurable and if* $|f|\mathcal{X}_K \in L^1(\mathbb{R}^n, \mu)$ *for every compact set,* K. *Here* μ *is a Radon measure on* $\mathbb{R}^n$. *Usually* $\mu = m$, *Lebesgue measure. When this is so, we write* $L^1_{loc}(\mathbb{R}^n)$ *or* $L^p(\mathbb{R}^n)$, *etc. If* $f \in L^1_{loc}(\mathbb{R}^n)$, *and* $g \in C_c(\mathbb{R}^n)$,

$$f * g(\mathbf{x}) \equiv \int f(\mathbf{x} - \mathbf{y}) g(\mathbf{y}) dm = \int f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) dm.$$

The following lemma will be useful in what follows.

Lemma 12.22 *Let* $f \in L^1_{loc}(\mathbb{R}^n)$, *and* $g \in C_c^\infty(\mathbb{R}^n)$. *Then* $f * g$ *is an infinitely differentiable function.*

Proof: We look at the difference quotient for calculating a partial derivative of $f * g$.

$$\frac{f * g(\mathbf{x} + t\mathbf{e}_j) - f * g(\mathbf{x})}{t} = \frac{1}{t} \int f(\mathbf{y}) \frac{g(\mathbf{x} + t\mathbf{e}_j - \mathbf{y}) - g(\mathbf{x} - \mathbf{y})}{t} dm(y).$$

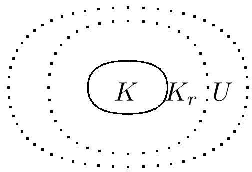
Using the fact that $g \in C_c^\infty(\mathbb{R}^n)$, we can argue that the quotient, $\frac{g(\mathbf{x} + t\mathbf{e}_j - \mathbf{y}) - g(\mathbf{x} - \mathbf{y})}{t}$ is uniformly bounded. Therefore, there exists a dominating function for the integrand of the above integral which is of the form $C|f(\mathbf{y})| \mathcal{X}_K$ where K is a compact set containing the support of g. It follows we can take the limit of the difference quotient above inside the integral and write

$$\frac{\partial}{\partial x_j}(f * g)(\mathbf{x}) = \int f(\mathbf{y}) \frac{\partial}{\partial x_j} g(\mathbf{x} - \mathbf{y}) dm(y).$$

Now letting $\frac{\partial}{\partial x_j} g$ play the role of g in the above argument, we can continue taking partial derivatives of all orders. This proves the lemma.

Theorem 12.23 *Let K be a compact subset of an open set, U. Then there exists a function, $h \in C_c^\infty(U)$, such that $h(\mathbf{x}) = 1$ for all $\mathbf{x} \in K$ and $h(\mathbf{x}) \in [0,1]$ for all $\mathbf{x}$.*

Proof: Let $r > 0$ be small enough that $K + B(0, 3r) \subseteq U$. Let $K_r = K + B(0, r)$.



Consider $\mathcal{X}_{K_r} * \psi_m$ where ψ_m is a mollifier. Let m be so large that $\frac{1}{m} < r$. Then using Lemma 12.22 it is straightforward to verify that $h = \mathcal{X}_{K_r} * \psi_m$ satisfies the desired conclusions.

Although we will not use the following corollary till later, it follows as an easy consequence of the above theorem and is useful. Therefore, we state it here.

Corollary 12.24 *Let K be a compact set in $\mathbb{R}^n$ and let $\{U_i\}_{i=1}^\infty$ be an open cover of K. Then there exists functions, $\psi_k \in C_c^\infty(U_i)$ such that $\psi_i \prec U_i$ and*

$$\sum_{i=1}^{\infty} \psi_i(\mathbf{x}) = 1.$$

If K_1 is a compact subset of U_1 we may also take ψ_1 such that $\psi_1(\mathbf{x}) = 1$ for all $\mathbf{x} \in K_1$.

Proof: This follows from a repeat of the proof of Theorem 6.11, replacing the lemma used in that proof with Theorem 12.23.

Theorem 12.25 *For each $p \geq 1$, $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.*

Proof: Let $f \in L^p(\mathbb{R}^n)$ and let $\varepsilon > 0$ be given. Choose $g \in C_c(\mathbb{R}^n)$ such that $||f - g||_p < \frac{\varepsilon}{2}$. Now let $g_m = g * \psi_m$ where $\{\psi_m\}$ is a mollifier.

$$[g_m(\mathbf{x} + h\mathbf{e}_i) - g_m(\mathbf{x})]h^{-1}$$

$$= h^{-1} \int g(\mathbf{y})[\psi_m(\mathbf{x} + h\mathbf{e}_i - \mathbf{y}) - \psi_m(\mathbf{x} - \mathbf{y})]dm.$$

The integrand is dominated by $C|g(\mathbf{y})|h$ for some constant C depending on

$$\max\{|\partial\psi_m(\mathbf{x})/\partial x_j| : \mathbf{x} \in \mathbb{R}^n, j \in \{1, 2, \cdots, n\}\}.$$

By the Dominated Convergence theorem, the limit as $h \to 0$ exists and yields

$$\frac{\partial g_m(\mathbf{x})}{\partial x_i} = \int g(\mathbf{y}) \frac{\partial \psi_m(\mathbf{x} - \mathbf{y})}{\partial x_i} dy.$$

Similarly, all other partial derivatives exist and are continuous as are all higher order derivatives. Consequently, $g_m \in C_c^\infty(\mathbb{R}^n)$. It vanishes if $\mathbf{x} \notin \operatorname{spt}(g) + B(0, \frac{1}{m})$.

$$\begin{aligned}
||g - g_m||_p &= \left(\int \left|g(\mathbf{x}) - \int g(\mathbf{x}-\mathbf{y})\psi_m(\mathbf{y})dm(\mathbf{y})\right|^p dm(\mathbf{x})\right)^{\frac{1}{p}} \\
&\leq \left(\int \left(\int |g(\mathbf{x}) - g(\mathbf{x}-\mathbf{y})|\psi_m(\mathbf{y})dm(\mathbf{y})\right)^p dm(\mathbf{x})\right)^{\frac{1}{p}} \\
&\leq \int \left(\int |g(\mathbf{x}) - g(\mathbf{x}-\mathbf{y})|^p dm(\mathbf{x})\right)^{\frac{1}{p}} \psi_m(\mathbf{y})dm(\mathbf{y}) \\
&= \int_{B(0,\frac{1}{m})} ||g - g_{\mathbf{y}}||_p \psi_m(\mathbf{y})dm(\mathbf{y}) \\
&< \frac{\varepsilon}{2}
\end{aligned}$$

whenever m is large enough. This follows from Theorem 12.13. Theorem 12.7 was used to obtain the third inequality. There is no measurability problem because the function

$$(\mathbf{x}, \mathbf{y}) \to |g(\mathbf{x}) - g(\mathbf{x}-\mathbf{y})|\psi_m(\mathbf{y})$$

is continuous so it is surely Borel measurable, hence product measurable. Thus when m is large enough,

$$||f - g_m||_p \leq ||f - g||_p + ||g - g_m||_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves the theorem.

12.6 Exercises

1. Let E be a Lebesgue measurable set in $\mathbb{R}$. Suppose $m(E) > 0$. Consider the set

$$E - E = \{x - y : x \in E, y \in E\}.$$

Show that $E - E$ contains an interval. **Hint:** Let

$$f(x) = \int \mathcal{X}_E(t)\mathcal{X}_E(x+t)dt.$$

Note f is continuous at 0 and $f(0) > 0$. Remember continuity of translation in L^p.

2. Give an example of a sequence of functions in $L^p(\mathbb{R})$ which converges to zero in L^p but does not converge pointwise to 0. Does this contradict the proof of the theorem that L^p is complete?

3. Let $\phi_m \in C_c^\infty(\mathbb{R}^n), \phi_m(x) \geq 0$,and $\int_{\mathbb{R}^n} \phi_m(y)dy = 1$ with $\lim_{m\to\infty} \sup\{|x| : x \in \sup(\phi_m)\} = 0$. Show if $f \in L^p(\mathbb{R}^n)$, $\lim_{m\to\infty} f * \phi_m = f$ in $L^p(\mathbb{R}^n)$.

4. Let $\phi : \mathbb{R} \to \mathbb{R}$ be convex. This means

$$\phi(\lambda x + (1-\lambda)y) \leq \lambda\phi(x) + (1-\lambda)\phi(y)$$

whenever $\lambda \in [0, 1]$. Show that if ϕ is convex, then ϕ is continuous. Also verify that if $x < y < z$, then $\frac{\phi(y)-\phi(x)}{y-x} \leq \frac{\phi(z)-\phi(y)}{z-y}$ and that $\frac{\phi(z)-\phi(x)}{z-x} \leq \frac{\phi(z)-\phi(y)}{z-y}$.

5. ↑ Prove Jensen's inequality. If $\phi : \mathbb{R} \to \mathbb{R}$ is convex, $\mu(\Omega) = 1$, and $f : \Omega \to \mathbb{R}$ is in $L^1(\Omega)$, then $\phi(\int_\Omega f\ du) \leq \int_\Omega \phi(f) d\mu$. **Hint:** Let $s = \int_\Omega f\ d\mu$ and show there exists λ such that $\phi(s) \leq \phi(t) + \lambda(s-t)$ for all t.

6. Let $\frac{1}{p} + \frac{1}{p'} = 1,\ p > 1,$ let $f \in L^p(\mathbb{R}),\ g \in L^{p'}(\mathbb{R})$. Show $f * g$ is uniformly continuous on $\mathbb{R}$ and $|(f*g)(x)| \leq ||f||_{L^p} ||g||_{L^{p'}}$.

7. $B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx, \Gamma(p) = \int_0^\infty e^{-t}t^{p-1}dt$ for $p, q > 0$. The first of these is called the beta function, while the second is the gamma function. Show a.) $\Gamma(p+1) = p\Gamma(p)$; b.) $\Gamma(p)\Gamma(q) = B(p,q)\Gamma(p+q)$.

8. Let $f \in C_c(0,\infty)$ and define $F(x) = \frac{1}{x}\int_0^x f(t)dt$. Show

$$||F||_{L^p(0,\infty)} \leq \frac{p}{p-1} ||f||_{L^p(0,\infty)} \text{ whenever } p > 1.$$

Hint: Use $xF' = f - F$ and integrate $\int_0^\infty |F(x)|^p dx$ by parts.

9. ↑ Now suppose $f \in L^p(0,\infty),\ p > 1$, and f not necessarily in $C_c(0,\infty)$. Note that $F(x) = \frac{1}{x}\int_0^x f(t)dt$ still makes sense for each $x > 0$. Is the inequality of Problem 8 still valid? Why? This inequality is called Hardy's inequality.

10. When does equality hold in Holder's inequality? **Hint:** First suppose $f, g \geq 0$. This isolates the most interesting aspect of the question.

11. ↑ Consider Hardy's inequality of Problems 8 and 9. Show equality holds only if $f = 0$ a.e. **Hint:** If equality holds, we can assume $f \geq 0$. Why? You might then establish $(p-1)\int_0^\infty F^p dx = p\int_0^\infty F^{\frac{p}{q}} f\ dx$ and use Problem 10.

12. ↑ In Hardy's inequality, show the constant $p(p-1)^{-1}$ cannot be improved. Also show that if $f > 0$ and $f \in L^1$, then $F \notin L^1$ so it is important that $p > 1$. **Hint:** Try $f(x) = x^{-\frac{1}{p}} \mathcal{X}_{[A^{-1},A]}$.

13. A set of functions, $\Phi \subseteq L^1$, is uniformly integrable if for all $\varepsilon > 0$ there exists a $\sigma > 0$ such that $\left|\int_E f\ du\right| < \varepsilon$ whenever $\mu(E) < \sigma$. Prove Vitali's Convergence theorem: Let $\{f_n\}$ be uniformly integrable, $\mu(\Omega) < \infty,\ f_n(x) \to f(x)$ a.e. where f is measurable and $|f(x)| < \infty$ a.e. Then $f \in L^1$ and $\lim_{n\to\infty} \int_\Omega |f_n - f| d\mu = 0$. **Hint:** Use Egoroff's theorem to show $\{f_n\}$ is a Cauchy sequence in $L^1(\Omega)$. This yields a different and easier proof than what was done earlier.

14. ↑ Show the Vitali Convergence theorem implies the Dominated Convergence theorem for finite measure spaces.

15. ↑ Suppose $\mu(\Omega) < \infty,\ \{f_n\} \subseteq L^1(\Omega)$, and

$$\int_\Omega h(|f_n|)\, d\mu < C$$

for all n where h is a continuous, nonnegative function satisfying

$$\lim_{t\to\infty} \frac{h(t)}{t} = \infty.$$

Show $\{f_n\}$ is uniformly integrable.

16. ↑ Give an example of a situation in which the Vitali Convergence theorem applies, but the Dominated Convergence theorem does not.

17. $\uparrow$ Sometimes, especially in books on probability, a different definition of uniform integrability is used than that presented here. A set of functions, $\mathfrak{S}$, defined on a finite measure space, $(\Omega, \mathcal{S}, \mu)$ is said to be uniformly integrable if for all $\epsilon > 0$ there exists $\alpha > 0$ such that for all $f \in \mathfrak{S}$,

$$\int_{[|f| \geq \alpha]} |f| \, d\mu \leq \epsilon.$$

Show that this definition is equivalent to the definition of uniform integrability given earlier with the addition of the condition that there is a constant, $C < \infty$ such that

$$\int |f| \, d\mu \leq C$$

for all $f \in \mathfrak{S}$. If this definition of uniform integrability is used, show that if $f_n(\omega) \to f(\omega)$ a.e., then it is automatically the case that $|f(\omega)| < \infty$ a.e. so it is not necessary to check this condition in the hypotheses for the Vitali convergence theorem.

18. We say $f \in L^\infty(\Omega, \mu)$ if there exists a set of measure zero, E, and a constant $C < \infty$ such that $|f(x)| \leq C$ for all $x \notin E$.

$$||f||_\infty \equiv \inf\{C : |f(x)| \leq C \text{ } a.e.\}.$$

Show $|| \, ||_\infty$ is a norm on $L^\infty(\Omega, \mu)$ provided we identify f and g if $f(x) = g(x)$ a.e. Show $L^\infty(\Omega, \mu)$ is complete.

19. Suppose $f \in L^\infty \cap L^1$. Show $\lim_{p\to\infty} ||f||_{L^p} = ||f||_\infty$.

20. Suppose $\phi : \mathbb{R} \to \mathbb{R}$ and $\phi(\int_0^1 f(x)dx) \leq \int_0^1 \phi(f(x))dx$ for every real bounded measurable f. Can it be concluded that ϕ is convex?

21. Suppose $\mu(\Omega) < \infty$. Show that if $1 \leq p < q$, then $L^q(\Omega) \subseteq L^p(\Omega)$.

22. Show $L^1(\mathbb{R}) \nsubseteq L^2(\mathbb{R})$ and $L^2(\mathbb{R}) \nsubseteq L^1(\mathbb{R})$ if Lebesgue measure is used.

23. Show that if $x \in [0, 1]$ and $p \geq 2$, then

$$\left(\frac{1+x}{2}\right)^p + \left(\frac{1-x}{2}\right)^p \leq \frac{1}{2}(1 + x^p).$$

Note this is obvious if $p = 2$. Use this to conclude the following inequality valid for all $z, w \in \mathbb{C}$ and $p \geq 2$.

$$\left|\frac{z+w}{2}\right|^p + \left|\frac{z-w}{2}\right|^p \leq \frac{|z|^p}{2} + \frac{|w|^p}{2}.$$

Hint: For the first part, divide both sides by x^p, let $y = \frac{1}{x}$ and show the resulting inequality is valid for all $y \geq 1$. If $|z| \geq |w| > 0$, this takes the form

$$|\frac{1}{2}(1 + re^{i\theta})|^p + |\frac{1}{2}(1 - re^{i\theta})|^p \leq \frac{1}{2}(1 + r^p)$$

whenever $0 \leq \theta < 2\pi$ and $r \in [0, 1]$. Show the expression on the left is maximized when $\theta = 0$ and use the first part.

24. $\uparrow$ If $p \geq 2$, establish Clarkson's inequality. Whenever $f, g \in L^p$,

$$\left|\left|\frac{1}{2}(f+g)\right|\right|_p^p + \left|\left|\frac{1}{2}(f-g)\right|\right|_p^p \leq \frac{1}{2}\,||f||^p + \frac{1}{2}\,||g||_p^p.$$

For more on Clarkson inequalities (there are others), see Hewitt and Stromberg [15] or Ray [22].

25. ↑ Show that for $p \geq 2$, L^p is uniformly convex. This means that if $\{f_n\}, \{g_n\} \subseteq L^p$, $||f_n||_p$, $||g_n||_p \leq 1$, and $||f_n + g_n||_p \to 2$, then $||f_n - g_n||_p \to 0$.

26. Suppose that $\theta \in [0,1]$ and $r, s, q > 0$ with

$$\frac{1}{q} = \frac{\theta}{r} + \frac{1-\theta}{s}.$$

show that

$$(\int |f|^q d\mu)^{1/q} \leq ((\int |f|^r d\mu)^{1/r})^\theta ((\int |f|^s d\mu)^{1/s})^{1-\theta}.$$

If $q, r, s \geq 1$ this says that

$$||f||_q \leq ||f||_r^\theta ||f||_s^{1-\theta}.$$

Hint:

$$\int |f|^q d\mu = \int |f|^{q\theta} |f|^{q(1-\theta)} d\mu.$$

Now note that $1 = \frac{\theta q}{r} + \frac{q(1-\theta)}{s}$ and use Holder's inequality.

27. Generalize Theorem 12.7 as follows. Let $0 \leq p_1 \leq p_2 < \infty$. Then

$$\left(\int_Y \left(\int_X |f(x,y)|^{p_1} d\mu \right)^{p_2/p_1} d\lambda \right)^{1/p_2}$$

$$\leq \left(\int_X \left(\int_Y |f(x,y)|^{p_2} d\lambda \right)^{p_1/p_2} d\mu \right)^{1/p_1}.$$

Fourier Transforms

13.1 The Schwartz class

The Fourier transform of a function in $L^1(\mathbb{R}^n)$ is given by

$$Ff(\mathbf{t}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x}) dx.$$

However, we want to take the Fourier transform of many other kinds of functions. In particular we want to take the Fourier transform of functions in $L^2(\mathbb{R}^n)$ which is not a subset of $L^1(\mathbb{R}^n)$. Thus the above integral may not make sense. In defining what is meant by the Fourier Transform of more general functions, it is convenient to use a special class of functions known as the Schwartz class which is a subset of $L^p(\mathbb{R}^n)$ for all $p \geq 1$. The procedure is to define the Fourier transform of functions in the Schwartz class and then use this to define the Fourier transform of a more general function in terms of what happens when it is multiplied by the Fourier transform of functions in the Schwartz class.

The functions in the Schwartz class are infinitely differentiable and they vanish very rapidly as $|x| \to \infty$ along with all their partial derivatives. To describe precisely what we mean by this, we need to present some notation.

Definition 13.1 *$\alpha = (\alpha_1, \cdots, \alpha_n)$ for $\alpha_1 \cdots \alpha_n$ positive integers is called a multi-index. For α a multi-index, $|\alpha| \equiv \alpha_1 + \cdots + \alpha_n$ and if $\mathbf{x} \in \mathbb{R}^n$,*

$$\mathbf{x} = (x_1, \cdots, x_n),$$

and f a function, we define

$$\mathbf{x}^\alpha \equiv x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n},\ D^\alpha f(\mathbf{x}) \equiv \frac{\partial^{|\alpha|} f(\mathbf{x})}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}.$$

Definition 13.2 *We say $f \in \mathfrak{S}$, the Schwartz class, if $f \in C^\infty(\mathbb{R}^n)$ and for all positive integers N,*

$$\rho_N(f) < \infty$$

where

$$\rho_N(f) = \sup\{(1 + |\mathbf{x}|^2)^N |D^\alpha f(\mathbf{x})| : \mathbf{x} \in \mathbb{R}^n,\ |\alpha| \leq N\}.$$

Thus $f \in \mathfrak{S}$ if and only if $f \in C^\infty(\mathbb{R}^n)$ and

$$\sup\{|\mathbf{x}^\beta D^\alpha f(\mathbf{x})| : \mathbf{x} \in \mathbb{R}^n\} < \infty \tag{13.1}$$

for all multi indices α and β.

Also note that if $f \in \mathfrak{S}$, then $p(f) \in \mathfrak{S}$ for any polynomial, p with $p(0) = 0$ and that

$$\mathfrak{S} \subseteq L^p(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$$

for any $p \geq 1$.

Definition 13.3 *(Fourier transform on $\mathfrak{S}$) For $f \in \mathfrak{S}$,*

$$Ff(\mathbf{t}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x}) dx,$$

$$F^{-1}f(\mathbf{t}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x}) dx.$$

Here $\mathbf{x} \cdot \mathbf{t} = \sum_{i=1}^n x_i t_i$.

It will follow from the development given below that $(F \circ F^{-1})(f) = f$ and $(F^{-1} \circ F)(f) = f$ whenever $f \in \mathfrak{S}$, thus justifying the above notation.

Theorem 13.4 *If $f \in \mathfrak{S}$, then Ff and $F^{-1}f$ are also in $\mathfrak{S}$.*

Proof: To begin with, let $\alpha = \mathbf{e}_j = (0, 0, \cdots, 1, 0, \cdots, 0)$, the 1 in the jth slot.

$$\frac{F^{-1}f(\mathbf{t} + h\mathbf{e}_j) - F^{-1}f(\mathbf{t})}{h} = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x}) (\frac{e^{ihx_j} - 1}{h}) dx. \tag{13.2}$$

Consider the integrand in (13.2).

$$\begin{aligned} \left| e^{i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x}) (\frac{e^{ihx_j} - 1}{h}) \right| &= |f(\mathbf{x})| \left| (\frac{e^{i(h/2)x_j} - e^{-i(h/2)x_j}}{h}) \right| \\ &= |f(\mathbf{x})| \left| \frac{i \sin((h/2) x_j)}{(h/2)} \right| \\ &\leq |f(\mathbf{x})| |x_j| \end{aligned}$$

and this is a function in $L^1(\mathbb{R}^n)$ because $f \in \mathfrak{S}$. Therefore by the Dominated Convergence Theorem,

$$\begin{aligned} \frac{\partial F^{-1}f(\mathbf{t})}{\partial t_j} &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} i x_j f(\mathbf{x}) dx \\ &= i(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} \mathbf{x}^{\mathbf{e}_j} f(\mathbf{x}) dx. \end{aligned}$$

Now $\mathbf{x}^{\mathbf{e}_j} f(\mathbf{x}) \in \mathfrak{S}$ and so we may continue in this way and take derivatives indefinitely. Thus $F^{-1}f \in C^\infty(\mathbb{R}^n)$ and from the above argument,

$$D^\alpha F^{-1}f(\mathbf{t}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} (i\mathbf{x})^\alpha f(\mathbf{x}) dx.$$

To complete showing $F^{-1}f \in \mathfrak{S}$,

$$\mathbf{t}^\beta D^\alpha F^{-1}f(\mathbf{t}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} \mathbf{t}^\beta (i\mathbf{x})^a f(\mathbf{x}) dx.$$

Integrate this integral by parts to get

$$\mathbf{t}^\beta D^\alpha F^{-1}f(\mathbf{t}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} i^{|\beta|} e^{i\mathbf{t}\cdot\mathbf{x}} D^\beta ((i\mathbf{x})^a f(\mathbf{x})) dx. \tag{13.3}$$

Here is how this is done.

$$\begin{aligned}\int_{\mathbb{R}} e^{it_j x_j} t_j^{\beta_j} (i\mathbf{x})^\alpha f(\mathbf{x}) dx_j &= \frac{e^{it_j x_j}}{it_j} t_j^{\beta_j} (i\mathbf{x})^\alpha f(\mathbf{x}) \mid_{-\infty}^{\infty} + \\ &\quad i \int_{\mathbb{R}} e^{it_j x_j} t_j^{\beta_j - 1} D^{\mathbf{e}_j} ((i\mathbf{x})^\alpha f(\mathbf{x})) dx_j\end{aligned}$$

where the boundary term vanishes because $f \in \mathfrak{S}$. Returning to (13.3), we use (13.1), and the fact that $|e^{ia}| = 1$ to conclude

$$|\mathbf{t}^\beta D^\alpha F^{-1} f(\mathbf{t})| \leq C \int_{\mathbb{R}^n} |D^\beta ((i\mathbf{x})^a f(\mathbf{x}))| dx < \infty.$$

It follows $F^{-1} f \in \mathfrak{S}$. Similarly $Ff \in \mathfrak{S}$ whenever $f \in \mathfrak{S}$.

Theorem 13.5 *$F \circ F^{-1}(f) = f$ and $F^{-1} \circ F(f) = f$ whenever $f \in \mathfrak{S}$.*

Before proving this theorem, we need a lemma.

Lemma 13.6

$$(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{(-1/2)\mathbf{u}\cdot\mathbf{u}} du = 1, \tag{13.4}$$

$$(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{(-1/2)(\mathbf{u}-i\mathbf{a})\cdot(\mathbf{u}-i\mathbf{a})} du = 1. \tag{13.5}$$

Proof:

$$\begin{aligned}(\int_{\mathbb{R}} e^{-x^2/2} dx)^2 &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-(x^2+y^2)/2} dx dy \\ &= \int_0^\infty \int_0^{2\pi} r e^{-r^2/2} d\theta dr = 2\pi.\end{aligned}$$

Therefore

$$\int_{\mathbb{R}^n} e^{(-1/2)\mathbf{u}\cdot\mathbf{u}} du = \prod_{i=1}^n \int_{\mathbb{R}} e^{-x_j^2/2} dx_j = (2\pi)^{n/2}.$$

This proves (13.4). To prove (13.5) it suffices to show that

$$\int_{\mathbb{R}} e^{(-1/2)(x-ia)^2} dx = (2\pi)^{1/2}. \tag{13.6}$$

Define $h(a)$ to be the left side of (13.6). Thus

$$\begin{aligned}h(a) &= (\int_{\mathbb{R}} e^{(-1/2)x^2} (\cos(ax) + i \sin ax)\, dx) e^{a^2/2} \\ &= (\int_{\mathbb{R}} e^{(-1/2)x^2} \cos(ax) dx) e^{a^2/2}\end{aligned}$$

because sin is an odd function. We know $h(0) = (2\pi)^{1/2}$.

$$h'(a) = ah(a) + e^{a^2/2} \frac{d}{da} (\int_{\mathbb{R}} e^{-x^2/2} \cos(ax) dx). \tag{13.7}$$

Forming difference quotients and using the Dominated Convergence Theorem, we can take $\frac{d}{da}$ inside the integral in (13.7) to obtain

$$-\int_{\mathbb{R}} xe^{(-1/2)x^2}\sin(ax)dx.$$

Integrating this by parts yields

$$\frac{d}{da}(\int_{\mathbb{R}} e^{-x^2/2}\cos(ax)dx) = -a(\int_{\mathbb{R}} e^{-x^2/2}\cos(ax)dx).$$

Therefore

$$\begin{aligned} h'(a) &= ah(a) - ae^{a^2/2}\int_{\mathbb{R}} e^{-x^2/2}\cos(ax)dx \\ &= ah(a) - ah(a) = 0. \end{aligned}$$

This proves the lemma since $h(0) = (2\pi)^{1/2}$.

Proof of Theorem 13.5 Let

$$g_\varepsilon(\mathbf{x}) = e^{(-\varepsilon^2/2)\mathbf{x}\cdot\mathbf{x}}.$$

Thus $0 \leq g_\varepsilon(\mathbf{x}) \leq 1$ and $\lim_{\varepsilon\to 0+} g_\varepsilon(\mathbf{x}) = 1$. By the Dominated Convergence Theorem,

$$(F\circ F^{-1})f(\mathbf{x}) = \lim_{\varepsilon\to 0}(2\pi)^{\frac{-n}{2}}\int_{\mathbb{R}^n} F^{-1}f(\mathbf{t})g_\varepsilon(\mathbf{t})e^{-i\mathbf{t}\cdot\mathbf{x}}dt.$$

Therefore,

$$(F\circ F^{-1})f(\mathbf{x}) =$$

$$\begin{aligned} &= \lim_{\varepsilon\to 0}(2\pi)^{-n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i\mathbf{y}\cdot\mathbf{t}}f(\mathbf{y})g_\varepsilon(\mathbf{t})e^{-i\mathbf{x}\cdot\mathbf{t}}dydt \\ &= \lim_{\varepsilon\to 0}(2\pi)^{-n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i(\mathbf{y}-\mathbf{x})\cdot\mathbf{t}}f(\mathbf{y})g_\varepsilon(\mathbf{t})dtdy \\ &= \lim_{\varepsilon\to 0}(2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} f(\mathbf{y})[(2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} e^{i(\mathbf{y}-\mathbf{x})\cdot\mathbf{t}}g_\varepsilon(\mathbf{t})dt]dy. \end{aligned} \tag{13.8}$$

Consider [ ] in (13.8). This equals

$$(2\pi)^{-\frac{n}{2}}\varepsilon^{-n}\int_{\mathbb{R}^n} e^{-\frac{1}{2}(\mathbf{u}-i\mathbf{a})\cdot(\mathbf{u}-i\mathbf{a})}e^{-\frac{1}{2}|\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}|^2}du,$$

where $\mathbf{a} = \varepsilon^{-1}(\mathbf{y}-\mathbf{x})$, and $|\mathbf{z}| = (\mathbf{z}\cdot\mathbf{z})^{\frac{1}{2}}$. Applying Lemma 13.6

$$\begin{aligned} (2\pi)^{-\frac{n}{2}}[\ ] &= (2\pi)^{-\frac{n}{2}}\varepsilon^{-n}e^{-\frac{1}{2}|\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}|^2} \\ &\equiv m_\varepsilon(\mathbf{y}-\mathbf{x}) = m_\varepsilon(\mathbf{x}-\mathbf{y}) \end{aligned}$$

and by Lemma 13.6,

$$\int_{\mathbb{R}^n} m_\varepsilon(\mathbf{y})dy = 1. \tag{13.9}$$

Thus from (13.8),

$$(F \circ F^{-1})f(\mathbf{x}) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} f(\mathbf{y}) m_\varepsilon(\mathbf{x} - \mathbf{y}) dy = \lim_{\varepsilon \to 0} f * m_\varepsilon(\mathbf{x}). \tag{13.10}$$

$$\int_{|\mathbf{y}| \geq \delta} m_\varepsilon(\mathbf{y}) dy = (2\pi)^{-\frac{n}{2}} (\int_{|\mathbf{y}| \geq \delta} e^{-\frac{1}{2}|\frac{\mathbf{y}}{\varepsilon}|^2} dy) \varepsilon^{-n}.$$

Using polar coordinates,

$$= (2\pi)^{-n/2} \int_\delta^\infty \int_{S^{n-1}} e^{(-1/(2\varepsilon^2))\rho^2} \rho^{n-1} d\omega d\rho \varepsilon^{-n}$$

$$= (2\pi)^{-n/2} (\int_{\delta/\varepsilon}^\infty e^{(-1/2)\rho^2} \rho^{n-1} d\rho) C_n.$$

This clearly converges to 0 as $\varepsilon \to 0+$ because of the Dominated Convergence Theorem and the fact that $\rho^{n-1} e^{-\rho^2/2}$ is in $L^1(\mathbb{R})$. Hence

$$\lim_{\varepsilon \to 0} \int_{|\mathbf{y}| \geq \delta} m_\varepsilon(\mathbf{y}) dy = 0.$$

Let δ be small enough that $|f(\mathbf{x}) - f(\mathbf{x} - \mathbf{y})| < \eta$ whenever $|\mathbf{y}| \leq \delta$. Therefore, from Formulas (13.9) and (13.10),

$$|f(\mathbf{x}) - (F \circ F^{-1})f(\mathbf{x})| = \lim_{\varepsilon \to 0} |f(\mathbf{x}) - f * m_\varepsilon(\mathbf{x})|$$

$$\leq \limsup_{\varepsilon \to 0} \int_{\mathbb{R}^n} |f(\mathbf{x}) - f(\mathbf{x} - \mathbf{y})| m_\varepsilon(\mathbf{y}) dy$$

$$\leq \limsup_{\varepsilon \to 0} \left(\int_{|\mathbf{y}| > \delta} |f(\mathbf{x}) - f(\mathbf{x} - \mathbf{y})| m_\varepsilon(\mathbf{y}) dy + \right.$$

$$\left. \int_{|\mathbf{y}| \leq \delta} |f(\mathbf{x}) - f(\mathbf{x} - \mathbf{y})| m_\varepsilon(\mathbf{y}) dy \right)$$

$$\leq \limsup_{\varepsilon \to 0} ((\int_{|\mathbf{y}| > \delta} m_\varepsilon(\mathbf{y}) dy) 2||f||_\infty + \eta) = \eta.$$

Since $\eta > 0$ is arbitrary, $f(\mathbf{x}) = (F \circ F^{-1})f(\mathbf{x})$ whenever $f \in \mathfrak{S}$. This proves Theorem 13.5 and justifies the notation in Definition 13.3.

13.2 Fourier transforms of functions in $L^2(\mathbb{R}^n)$

With this preparation, we are ready to begin the consideration of Ff and $F^{-1}f$ for $f \in L^2(\mathbb{R}^n)$. First note that the formula given for Ff and $F^{-1}f$ when $f \in \mathfrak{S}$ will not work for $f \in L^2(\mathbb{R}^n)$ unless f is also in $L^1(\mathbb{R}^n)$. The following theorem will make possible the definition of Ff and $F^{-1}f$ for arbitrary $f \in L^2(\mathbb{R}^n)$.

Theorem 13.7 *For $\phi \in \mathfrak{S}$, $||F\phi||_2 = ||F^{-1}\phi||_2 = ||\phi||_2$.*

Proof: First note that for $\psi \in \mathfrak{S}$,

$$F(\bar{\psi}) = \overline{F^{-1}(\psi)}\ ,\ F^{-1}(\bar{\psi}) = \overline{F(\psi)}. \tag{13.11}$$

This follows from the definition. Let $\phi, \psi \in \mathfrak{S}$.

$$\begin{aligned} \int_{\mathbb{R}^n} (F\phi(\mathbf{t}))\psi(\mathbf{t})dt &= (2\pi)^{-n/2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \psi(\mathbf{t})\phi(\mathbf{x})e^{-i\mathbf{t}\cdot\mathbf{x}}dxdt \\ &= \int_{\mathbb{R}^n} \phi(\mathbf{x})(F\psi(\mathbf{x}))dx. \end{aligned} \tag{13.12}$$

Similarly,

$$\int_{\mathbb{R}^n} \phi(\mathbf{x})(F^{-1}\psi(\mathbf{x}))dx = \int_{\mathbb{R}^n} (F^{-1}\phi(\mathbf{t}))\psi(\mathbf{t})dt. \tag{13.13}$$

Now, (13.11) - (13.13) imply

$$\begin{aligned} \int_{\mathbb{R}^n} |\phi(\mathbf{x})|^2 dx &= \int_{\mathbb{R}^n} \phi(\mathbf{x})\overline{F^{-1}(F\phi(\mathbf{x}))}dx \\ &= \int_{\mathbb{R}^n} \phi(\mathbf{x})F(\overline{F\phi(\mathbf{x})})dx \\ &= \int_{\mathbb{R}^n} F\phi(\mathbf{x})(\overline{F\phi(\mathbf{x})})dx \\ &= \int_{\mathbb{R}^n} |F\phi|^2 dx. \end{aligned}$$

Similarly

$$||\phi||_2 = ||F^{-1}\phi||_2.$$

This proves the theorem.

With this theorem we are now able to define the Fourier transform of a function in $L^2(\mathbb{R}^n)$.

Definition 13.8 *Let $f \in L^2(\mathbb{R}^n)$ and let $\{\phi_k\}$ be a sequence of functions in $\mathfrak{S}$ converging to f in $L^2(\mathbb{R}^n)$. We know such a sequence exists because $\mathfrak{S}$ is dense in $L^2(\mathbb{R}^n)$. (Recall that even $C_c^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$.) Then $Ff \equiv \lim_{k\to\infty} F\phi_k$ where the limit is taken in $L^2(\mathbb{R}^n)$. A similar definition holds for $F^{-1}f$.*

Lemma 13.9 *The above definition is well defined.*

Proof: We need to verify two things. First we need to show that $\lim_{k\to\infty} F(\phi_k)$ exists in $L^2(\mathbb{R}^n)$ and next we need to verify that this limit is independent of the sequence of functions in $\mathfrak{S}$ which is converging to f.

To verify the first claim, note that since $\lim_{k\to\infty} ||f - \phi_k||_2 = 0$, it follows that $\{\phi_k\}$ is a Cauchy sequence. Therefore, by Theorem 13.7 $\{F\phi_k\}$ and $\{F^{-1}\phi_k\}$ are also Cauchy sequences in $L^2(\mathbb{R}^n)$. Therefore, they both converge in $L^2(\mathbb{R}^n)$ to a unique element of $L^2(\mathbb{R}^n)$. This verifies the first part of the lemma.

Now suppose $\{\phi_k\}$ and $\{\psi_k\}$ are two sequences of functions in $\mathfrak{S}$ which converge to $f \in L^2(\mathbb{R}^n)$. Do $\{F\phi_k\}$ and $\{F\psi_k\}$ converge to the same element of $L^2(\mathbb{R}^n)$? We know that for k large, $||\phi_k - \psi_k||_2 \leq ||\phi_k - f||_2 + ||f - \psi_k||_2 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Therefore, for large enough k, we have $||F\phi_k - F\psi_k||_2 = ||\phi_k - \psi_k||_2 < \varepsilon$ and so, since $\varepsilon > 0$ is arbitrary, it follows that $\{F\phi_k\}$ and $\{F\psi_k\}$ converge to the same element of $L^2(\mathbb{R}^n)$.

We leave as an easy exercise the following identity for $\phi, \psi \in \mathfrak{S}$.

$$\int_{\mathbb{R}^n} \psi(\mathbf{x}) F\phi(\mathbf{x}) dx = \int_{\mathbb{R}^n} F\psi(\mathbf{x})\phi(\mathbf{x}) dx$$

and

$$\int_{\mathbb{R}^n} \psi(\mathbf{x}) F^{-1}\phi(\mathbf{x}) dx = \int_{\mathbb{R}^n} F^{-1}\psi(\mathbf{x})\phi(\mathbf{x}) dx$$

Theorem 13.10 *If $f \in L^2(\mathbb{R}^n)$, Ff and $F^{-1}f$ are the unique elements of $L^2(\mathbb{R}^n)$ such that for all $\phi \in \mathfrak{S}$,*

$$\int_{\mathbb{R}^n} Ff(\mathbf{x})\phi(\mathbf{x}) dx = \int_{\mathbb{R}^n} f(\mathbf{x}) F\phi(\mathbf{x}) dx, \tag{13.14}$$

$$\int_{\mathbb{R}^n} F^{-1}f(\mathbf{x})\phi(\mathbf{x}) dx = \int_{\mathbb{R}^n} f(\mathbf{x}) F^{-1}\phi(\mathbf{x}) dx. \tag{13.15}$$

Proof: Let $\{\phi_k\}$ be a sequence in $\mathfrak{S}$ converging to f in $L^2(\mathbb{R}^n)$ so $F\phi_k$ converges to Ff in $L^2(\mathbb{R}^n)$. Therefore, by Holder's inequality or the Cauchy Schwartz inequality,

$$\begin{aligned}
\int_{\mathbb{R}^n} Ff(\mathbf{x})\phi(\mathbf{x}) dx &= \lim_{k\to\infty} \int_{\mathbb{R}^n} F\phi_k(\mathbf{x})\phi(\mathbf{x}) dx \\
&= \lim_{k\to\infty} \int_{\mathbb{R}^n} \phi_k(\mathbf{x}) F\phi(\mathbf{x}) dx \\
&= \int_{\mathbb{R}^n} f(\mathbf{x}) F\phi(\mathbf{x}) dx.
\end{aligned}$$

A similar formula holds for F^{-1}. It only remains to verify uniqueness. Suppose then that for some $G \in L^2(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} G(\mathbf{x})\,\phi(\mathbf{x})\,dx = \int_{\mathbb{R}^n} Ff(\mathbf{x})\phi(\mathbf{x}) dx \tag{13.16}$$

for all $\phi \in \mathfrak{S}$. Does it follow that $G = Ff$ in $L^2(\mathbb{R}^n)$? Let $\{\phi_k\}$ be a sequence of elements of $\mathfrak{S}$ converging in $L^2(\mathbb{R}^n)$ to $\overline{G - Ff}$. Then from (13.16),

$$0 = \lim_{k\to\infty} \int_{\mathbb{R}^n} (G(\mathbf{x}) - Ff(\mathbf{x}))\,\phi_k(\mathbf{x})\,dx = \int_{\mathbb{R}^n} |G(\mathbf{x}) - Ff(\mathbf{x})|^2\,dx.$$

Thus $G = Ff$ and this proves uniqueness. A similar argument applies for F^{-1}.

Theorem 13.11 *(Plancherel) For $f \in L^2(\mathbb{R}^n)$.*

$$(F^{-1} \circ F)(f) = f = (F \circ F^{-1})(f), \tag{13.17}$$

$$||f||_2 = ||Ff||_2 = ||F^{-1}f||_2. \tag{13.18}$$

Proof: Let $\phi \in \mathfrak{S}$.

$$\begin{aligned}\int_{\mathbb{R}^n} (F^{-1} \circ F)(f)\phi dx &= \int_{\mathbb{R}^n} Ff\ F^{-1}\phi dx = \int_{\mathbb{R}^n} f\ F(F^{-1}(\phi))dx \\ &= \int f\phi dx.\end{aligned}$$

Thus $(F^{-1} \circ F)(f) = f$ because $\mathfrak{S}$ is dense in $L^2(\mathbb{R}^n)$. Similarly $(F \circ F^{-1})(f) = f$. This proves (13.17). To show (13.18), we use the density of $\mathfrak{S}$ to obtain a sequence, $\{\phi_k\}$ converging to f in $L^2(\mathbb{R}^n)$. Then

$$||Ff||_2 = \lim_{k\to\infty} ||F\phi_k||_2 = \lim_{k\to\infty} ||\phi_k||_2 = ||f||_2 .$$

Similarly,

$$||f||_2 = ||F^{-1}f||_2.$$

This proves the theorem.

The following corollary is a simple generalization of this. To prove this corollary, we use the following simple lemma which comes as a consequence of the Cauchy Schwartz inequality.

Lemma 13.12 *Suppose $f_k \to f$ in $L^2(\mathbb{R}^n)$ and $g_k \to g$ in $L^2(\mathbb{R}^n)$. Then*

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k g_k dx = \int_{\mathbb{R}^n} fg dx$$

Proof:

$$\left|\int_{\mathbb{R}^n} f_k g_k dx - \int_{\mathbb{R}^n} fg dx\right| \leq \left|\int_{\mathbb{R}^n} f_k g_k dx - \int_{\mathbb{R}^n} f_k g dx\right| +$$

$$\left|\int_{\mathbb{R}^n} f_k g dx - \int_{\mathbb{R}^n} fg dx\right|$$

$$\leq ||f_k||_2 \, ||g - g_k||_2 + ||g||_2 \, ||f_k - f||_2 .$$

Now $||f_k||_2$ is a Cauchy sequence and so it is bounded independent of k. Therefore, the above expression is smaller than ε whenever k is large enough. This proves the lemma.

Corollary 13.13 *For $f, g \in L^2(\mathbb{R}^n)$,*

$$\int_{\mathbb{R}^n} f\overline{g} dx = \int_{\mathbb{R}^n} Ff\ \overline{Fg} dx = \int_{\mathbb{R}^n} F^{-1}f\ \overline{F^{-1}g} dx.$$

Proof: First note the above formula is obvious if $f, g \in \mathfrak{S}$. To see this, note

$$\begin{aligned}\int_{\mathbb{R}^n} Ff\ \overline{Fg} dx &= \int_{\mathbb{R}^n} Ff(x)\, \overline{\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-i\mathbf{x}\cdot\mathbf{t}} g(t)\, dt} dx \\ &= \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i\mathbf{x}\cdot\mathbf{t}} Ff(x)\, dx \overline{g(t)} dt \\ &= \int_{\mathbb{R}^n} \left(F^{-1} \circ F\right) f(t)\, \overline{g(t)} dt \\ &= \int_{\mathbb{R}^n} f(t)\, \overline{g(t)} dt.\end{aligned}$$

The formula with F^{-1} is exactly similar.

Now to verify the corollary, let $\phi_k \to f$ in $L^2(\mathbb{R}^n)$ and let $\psi_k \to g$ in $L^2(\mathbb{R}^n)$. Then

$$\begin{aligned}\int_{\mathbb{R}^n} Ff\,\overline{Fg}dx &= \lim_{k\to\infty}\int_{\mathbb{R}^n} F\phi_k\,\overline{F\psi_k}dx \\ &= \lim_{k\to\infty}\int_{\mathbb{R}^n} \phi_k\overline{\psi_k}dx \\ &= \int_{\mathbb{R}^n} f\overline{g}dx\end{aligned}$$

A similar argument holds for F^{-1}.This proves the corollary.

How do we compute Ff and $F^{-1}f$?

Theorem 13.14 *For $f \in L^2(\mathbb{R}^n)$, let $f_r = f\mathcal{X}_{E_r}$ where E_r is a bounded measurable set with $E_r \uparrow \mathbb{R}^n$. Then the following limits hold in $L^2(\mathbb{R}^n)$.*

$$Ff = \lim_{r\to\infty} Ff_r\ ,\ F^{-1}f = \lim_{r\to\infty} F^{-1}f_r.$$

Proof: $||f - f_r||_2 \to 0$ and so $||Ff - Ff_r||_2 \to 0$ and $||F^{-1}f - F^{-1}f_r||_2 \to 0$ by Plancherel's Theorem. This proves the theorem.

What are Ff_r and $F^{-1}f_r$? Let $\phi \in \mathfrak{S}$

$$\begin{aligned}\int_{\mathbb{R}^n} Ff_r\phi dx &= \int_{\mathbb{R}^n} f_r F\phi dx \\ &= (2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f_r(\mathbf{x})e^{-i\mathbf{x}\cdot\mathbf{y}}\phi(\mathbf{y})dydx \\ &= \int_{\mathbb{R}^n}[(2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} f_r(\mathbf{x})e^{-i\mathbf{x}\cdot\mathbf{y}}dx]\phi(\mathbf{y})dy.\end{aligned}$$

Since this holds for all $\phi \in \mathfrak{S}$, a dense subset of $L^2(\mathbb{R}^n)$, it follows that

$$Ff_r(\mathbf{y}) = (2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} f_r(\mathbf{x})e^{-i\mathbf{x}\cdot\mathbf{y}}dx.$$

Similarly

$$F^{-1}f_r(\mathbf{y}) = (2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} f_r(\mathbf{x})e^{i\mathbf{x}\cdot\mathbf{y}}dx.$$

This shows that to take the Fourier transform of a function in $L^2(\mathbb{R}^n)$, it suffices to take the limit as $r \to \infty$ in $L^2(\mathbb{R}^n)$ of $(2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} f_r(\mathbf{x})e^{-i\mathbf{x}\cdot\mathbf{y}}dx$. A similar procedure works for the inverse Fourier transform.

Definition 13.15 *For $f \in L^1(\mathbb{R}^n)$, define*

$$Ff(\mathbf{x}) \equiv (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i\mathbf{x}\cdot\mathbf{y}}f(\mathbf{y})\,dy,$$

$$F^{-1}f(\mathbf{x}) \equiv (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{i\mathbf{x}\cdot\mathbf{y}}f(\mathbf{y})\,dy.$$

Thus, for $f \in L^1(\mathbb{R}^n)$, Ff and $F^{-1}f$ are uniformly bounded.

Theorem 13.16 *Let $h \in L^2(\mathbb{R}^n)$ and let $f \in L^1(\mathbb{R}^n)$. Then $h * f \in L^2(\mathbb{R}^n)$,*

$$F^{-1}(h * f) = (2\pi)^{n/2} F^{-1}hF^{-1}f,$$

$$F(h * f) = (2\pi)^{n/2} FhFf,$$

and

$$||h * f||_2 \leq ||h||_2 \, ||f||_1 \,. \tag{13.19}$$

Proof: Without loss of generality, we may assume h and f are both Borel measurable. Then an application of Minkowski's inequality yields

$$\left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |h(\mathbf{x}-\mathbf{y})| \, |f(\mathbf{y})| \, dy\right)^2 dx\right)^{1/2} \leq ||f||_1 \, ||h||_2 \,. \tag{13.20}$$

Hence $\int |h(\mathbf{x}-\mathbf{y})| \, |f(\mathbf{y})| \, dy < \infty$ a.e. $\mathbf{x}$ and

$$\mathbf{x} \to \int h(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) \, dy$$

is in $L^2(\mathbb{R}^n)$. Let $E_r \uparrow \mathbb{R}^n$, $m(E_r) < \infty$. Thus,

$$h_r \equiv \mathcal{X}_{E_r} h \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n),$$

and letting $\phi \in \mathfrak{S}$,

$$\int F(h_r * f)(\phi) \, dx$$

$$\begin{aligned}
&\equiv \int (h_r * f)(F\phi) \, dx \\
&= (2\pi)^{-n/2} \int\int\int h_r(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) e^{-i\mathbf{x}\cdot\mathbf{t}} \phi(\mathbf{t}) \, dt dy dx \\
&= (2\pi)^{-n/2} \int\int\left(\int h_r(\mathbf{x}-\mathbf{y}) e^{-i(\mathbf{x}-\mathbf{y})\cdot\mathbf{t}} dx\right) f(\mathbf{y}) e^{-i\mathbf{y}\cdot\mathbf{t}} dy \phi(\mathbf{t}) \, dt \\
&= \int (2\pi)^{n/2} Fh_r(\mathbf{t}) Ff(\mathbf{t}) \phi(\mathbf{t}) \, dt.
\end{aligned}$$

Since ϕ is arbitrary and $\mathfrak{S}$ is dense in $L^2(\mathbb{R}^n)$,

$$F(h_r * f) = (2\pi)^{n/2} Fh_r Ff.$$

Now by Minkowski's Inequality, $h_r * f \to h * f$ in $L^2(\mathbb{R}^n)$ and also it is clear that $h_r \to h$ in $L^2(\mathbb{R}^n)$; so, by Plancherel's theorem, we may take the limit in the above and conclude

$$F(h * f) = (2\pi)^{n/2} FhFf.$$

The assertion for F^{-1} is similar and (13.19) follows from (13.20).

13.3 Tempered distributions

In this section we give an introduction to the general theory of Fourier transforms. Recall that $\mathfrak{S}$ is the set of all $\phi \in C^{\infty}(\mathbb{R}^n)$ such that for $N = 1, 2, \cdots$,

$$\rho_N(\phi) \equiv \sup_{|\alpha| \leq N, \mathbf{x} \in \mathbb{R}^n} \left(1 + |\mathbf{x}|^2\right)^N |D^{\alpha}\phi(\mathbf{x})| < \infty.$$

The set, $\mathfrak{S}$ is a vector space and we can make it into a topological space by letting sets of the form be a basis for the topology.

$$B^N(\phi, r) \equiv \{\psi \in \mathfrak{S} \text{ such that } \rho_N(\psi - \phi) < r\}.$$

Note the functions, ρ_N are increasing in N. Then $\mathfrak{S}'$, the space of continuous linear functions defined on $\mathfrak{S}$ mapping $\mathfrak{S}$ to $\mathbb{C}$, are called tempered distributions. Thus,

Definition 13.17 *$f \in \mathfrak{S}'$ means $f : \mathfrak{S} \to \mathbb{C}$, f is linear, and f is continuous.*

How can we verify $f \in \mathfrak{S}'$? The following lemma is about this question along with the similar question of when a linear map from $\mathfrak{S}$ to $\mathfrak{S}$ is continuous.

Lemma 13.18 *Let $f : \mathfrak{S} \to \mathbb{C}$ be linear and let $L : \mathfrak{S} \to \mathfrak{S}$ be linear. Then f is continuous if and only if*

$$|f(\phi)| \leq C\rho_N(\phi) \tag{a.}$$

for some N. Also, L is continuous if and only if for each N, there exists M such that

$$\rho_N(L\phi) \leq C\rho_M(\phi) \tag{b.}$$

for some C independent of ϕ.

Proof: It suffices to verify continuity at 0 because f and L are linear. We verify $(b.)$. Let $0 \in U$ where U is an open set. Then $0 \in B^N(0, r) \subseteq U$ for some $r > 0$ and N. Then if M and C are as described in $(b.)$, and $\psi \in B^M\left(0, C^{-1}r\right)$, we have

$$\rho_N(L\psi) \leq C\rho_M(\psi) < r;$$

so, this shows

$$B^M\left(0, C^{-1}r\right) \subseteq L^{-1}\left(B^N(0, r)\right) \subseteq L^{-1}(U),$$

which shows that L is continuous at 0. The argument for f and the only if part of the proof is left for the reader.

The key to extending the Fourier transform to $\mathfrak{S}'$ is the following theorem which states that F and F^{-1} are continuous. This is analogous to the procedure in defining the Fourier transform for functions in $L^2(\mathbb{R}^n)$. Recall we proved these mappings preserved the L^2 norms of functions in $\mathfrak{S}$.

Theorem 13.19 *For F and F^{-1} the Fourier and inverse Fourier transforms,*

$$F\phi(\mathbf{x}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\mathbf{x}\cdot\mathbf{y}}\phi(\mathbf{y})\, dy,$$

$$F^{-1}\phi(\mathbf{x}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{x}\cdot\mathbf{y}}\phi(\mathbf{y})\, dy,$$

F and F^{-1} are continuous linear maps from $\mathfrak{S}$ to $\mathfrak{S}$.

Proof: Let $|\alpha| \leq N$ where $N > 0$, and let $\mathbf{x} \neq \mathbf{0}$. Then

$$\left(1+|\mathbf{x}|^2\right)^N \left|D^\alpha F^{-1}\phi(\mathbf{x})\right| \equiv C_n \left(1+|\mathbf{x}|^2\right)^N \left|\int_{\mathbb{R}^n} e^{i\mathbf{x}\cdot\mathbf{y}} \mathbf{y}^\alpha \phi(\mathbf{y})\, dy\right|. \tag{13.21}$$

Suppose $x_j^2 \geq 1$ for $j \in \{i_1, \cdots, i_r\}$ and $x_j^2 < 1$ if $j \notin \{i_1, \cdots, i_r\}$. Then after integrating by parts in the integral of (13.21), we obtain the following for the right side of (13.21):

$$C_n \left(1+|\mathbf{x}|^2\right)^N \left|\int_{\mathbb{R}^n} e^{i\mathbf{x}\cdot\mathbf{y}} D^\beta \left(\mathbf{y}^\alpha \phi(\mathbf{y})\right) dy\right| \prod_{j\in\{i_1,\cdots,i_r\}} x_j^{-2N}, \tag{13.22}$$

where

$$\beta \equiv 2N \sum_{k=1}^r \mathbf{e}_{i_k},$$

the vector with 0 in the jth slot if $j \notin \{i_1, \cdots, i_r\}$ and a $2N$ in the jth slot for $j \in \{i_1, \cdots, i_r\}$. Now letting $C(n,N)$ denote a generic constant depending only on n and N, the product rule and a little estimating yields

$$\left|D^\beta \left(\mathbf{y}^\alpha \phi(\mathbf{y})\right)\right| \left(1+|\mathbf{y}|^2\right)^{nN} \leq \rho_{2Nn}(\phi)\, C(n,N).$$

Also the function $\mathbf{y} \to \left(1+|\mathbf{y}|^2\right)^{-nN}$ is integrable. Therefore, the expression in (13.22) is no larger than

$$C(n,N) \left(1+|\mathbf{x}|^2\right)^N \prod_{j\in\{i_1,\cdots,i_r\}} x_j^{-2N} \rho_{2Nn}(\phi)$$

$$\leq C(n,N) \left(1+n+\sum_{j\in\{i_1,\cdots,i_r\}} |x_j|^2\right)^N \prod_{j\in\{i_1,\cdots,i_r\}} x_j^{-2N} \rho_{2Nn}(\phi)$$

$$\leq C(n,N) \sum_{j\in\{i_1,\cdots,i_r\}} |x_j|^{2N} \prod_{j\in\{i_1,\cdots,i_r\}} |x_j|^{-2N} \rho_{2Nn}(\phi)$$

$$\leq C(n,N)\, \rho_{2Nn}(\phi).$$

Therefore, if $x_j^2 \geq 1$ for some j,

$$\left(1+|\mathbf{x}|^2\right)^N \left|D^\alpha F^{-1}\phi(\mathbf{x})\right| \leq C(n,N)\, \rho_{2Nn}(\phi).$$

If $x_j^2 < 1$ for all j, we can use (13.21) to obtain

$$\left(1+|\mathbf{x}|^2\right)^N \left|D^\alpha F^{-1}\phi(\mathbf{x})\right| \leq C_n (1+n)^N \left|\int_{\mathbb{R}^n} e^{i\mathbf{x}\cdot\mathbf{y}} \mathbf{y}^\alpha \phi(\mathbf{y})\, dy\right|$$

$$\leq C(n,N)\, \rho_R(\phi)$$

for some R depending on N and n. Let $M \geq \max(R, 2nN)$ and we see

$$\rho_N\left(F^{-1}\phi\right) \leq C \rho_M(\phi)$$

where M and C depend only on N and n. By Lemma 13.18, F^{-1} is continuous. Similarly, F is continuous. This proves the theorem.

Definition 13.20 *For $f \in \mathfrak{S}'$, we define Ff and $F^{-1}f$ in $\mathfrak{S}'$ by*

$$Ff(\phi) \equiv f(F\phi),\ F^{-1}f(\phi) \equiv f\left(F^{-1}\phi\right).$$

To see this is a good definition, consider the following.

$$|Ff(\phi)| \equiv |f(F\phi)| \leq C\rho_N(F\phi) \leq C\rho_M(\phi).$$

Also note that F and F^{-1} are both one to one and onto. This follows from the fact that these mappings map $\mathfrak{S}$ one to one onto $\mathfrak{S}$.

What are some examples of things in $\mathfrak{S}'$? In answering this question, we will use the following lemma.

Lemma 13.21 *If $f \in L^1_{loc}(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} f\phi dx = 0$ for all $\phi \in C_c^\infty(\mathbb{R}^n)$, then $f = 0$ a.e.*

Proof: It is enough to verify this for $f \geq 0$. Let

$$E \equiv \{\mathbf{x} : f(\mathbf{x}) \geq r\},\ E_R \equiv E \cap B(\mathbf{0},R).$$

Let K_n be an increasing sequence of compact sets and let V_n be a decreasing sequence of open sets satisfying

$$K_n \subseteq E_R \subseteq V_n,\ m(V_n \setminus K_n) \leq 2^{-n}, V_1 \text{ is bounded.}$$

Let

$$\phi_n \in C_c^\infty(V_n),\ K_n \prec \phi_n \prec V_n.$$

Then $\phi_n(\mathbf{x}) \to \mathcal{X}_{E_R}(\mathbf{x})$ a.e. because the set where $\phi_n(\mathbf{x})$ fails to converge is contained in the set of all $\mathbf{x}$ which are in infinitely many of the sets $V_n \setminus K_n$. This set has measure zero and so, by the dominated convergence theorem,

$$0 = \lim_{n\to\infty} \int_{\mathbb{R}^n} f\phi_n dx = \lim_{n\to\infty} \int_{V_1} f\phi_n dx = \int_{E_R} f dx \geq rm(E_R).$$

Thus, $m(E_R) = 0$ and therefore $m(E) = 0$. Since $r > 0$ is arbitrary, it follows

$$m([\mathbf{x} : f(\mathbf{x}) > 0]) = 0.$$

This proves the lemma.

Theorem 13.22 *Let f be a measurable function with polynomial growth,*

$$|f(\mathbf{x})| \leq C\left(1 + |\mathbf{x}|^2\right)^N \quad \text{for some } N,$$

or let $f \in L^p(\mathbb{R}^n)$ for some $p \in [1,\infty]$. Then $f \in \mathfrak{S}'$ if we define

$$f(\phi) \equiv \int f\phi dx.$$

Proof: Let f have polynomial growth first. Then

$$\begin{aligned}
\int |f| |\phi| dx &\leq C\int \left(1 + |\mathbf{x}|^2\right)^{nN} |\phi| dx \\
&\leq C\int \left(1 + |\mathbf{x}|^2\right)^{nN} \left(1 + |\mathbf{x}|^2\right)^{-2nN} dx\rho_{2nN}(\phi) \\
&\leq C(N,n)\rho_{2nN}(\phi) < \infty.
\end{aligned}$$

Therefore we can define

$$f(\phi) \equiv \int f\phi dx$$

and it follows that

$$|f(\phi)| \leq C(N,n)\,\rho_{2nN}(\phi).$$

By Lemma 13.18, $f \in \mathfrak{S}'$. Next suppose $f \in L^p(\mathbb{R}^n)$.

$$\int |f|\,|\phi|\,dx \leq \int |f|\left(1+|\mathbf{x}|^2\right)^{-M} dx\rho_M(\phi)$$

where we choose M large enough that $\left(1+|\mathbf{x}|^2\right)^{-M} \in L^{p'}$. Then by Holder's Inequality,

$$|f(\phi)| \leq ||f||_p\, C_n \rho_M(\phi).$$

By Lemma 13.18, $f \in \mathfrak{S}'$. This proves the theorem.

Definition 13.23 *If $f \in \mathfrak{S}'$ and $\phi \in \mathfrak{S}$, then $\phi f \in \mathfrak{S}'$ if we define*

$$\phi f(\psi) \equiv f(\phi\psi).$$

We need to verify that with this definition, $\phi f \in \mathfrak{S}'$. It is clearly linear. There exist constants C and N such that

$$\begin{aligned}
|\phi f(\psi)| &\equiv |f(\phi\psi)| \leq C\rho_N(\phi\psi) \\
&= C \sup_{\mathbf{x}\in\mathbb{R}^n,\ |\alpha|\leq N} \left(1+|\mathbf{x}|^2\right)^N |D^\alpha(\phi\psi)| \\
&\leq C(\phi,n,N)\,\rho_N(\psi).
\end{aligned}$$

Thus by Lemma 13.18, $\phi f \in \mathfrak{S}'$.

The next topic is that of convolution. This was discussed in Corollary 13.16 but we review it here. This corollary implies that if $f \in L^2(\mathbb{R}^n) \subseteq \mathfrak{S}'$ and $\phi \in \mathfrak{S}$, then

$$f * \phi \in L^2(\mathbb{R}^n),\ \ ||f*\phi||_2 \leq ||f||_2\,||\phi||_1,$$

and also

$$F(f*\phi)(\mathbf{x}) = F\phi(\mathbf{x})\,Ff(\mathbf{x})\,(2\pi)^{n/2}, \tag{13.23}$$

$$F^{-1}(f*\phi)(\mathbf{x}) = F^{-1}\phi(\mathbf{x})\,F^{-1}f(\mathbf{x})\,(2\pi)^{n/2}.$$

By Definition 13.23,

$$F(f*\phi) = (2\pi)^{n/2}\,F\phi Ff$$

whenever $f \in L^2(\mathbb{R}^n)$ and $\phi \in \mathfrak{S}$.

Now it is easy to see the proper way to define $f*\phi$ when f is only in $\mathfrak{S}'$ and $\phi \in \mathfrak{S}$.

Definition 13.24 *Let $f \in \mathfrak{S}'$ and let $\phi \in \mathfrak{S}$. Then we define*

$$f*\phi \equiv (2\pi)^{n/2}\,F^{-1}(F\phi Ff).$$

Theorem 13.25 *Let $f \in \mathfrak{S}'$ and let $\phi \in \mathfrak{S}$.*

$$F\left(f * \phi\right) = (2\pi)^{n/2} F\phi F f, \tag{13.24}$$

$$F^{-1}\left(f * \phi\right) = (2\pi)^{n/2} F^{-1}\phi F^{-1} f.$$

Proof: Note that (13.24) follows from Definition 13.24 and both assertions hold for $f \in \mathfrak{S}$. Next we write for $\psi \in \mathfrak{S}$,

$$\left(\psi * F^{-1}F^{-1}\phi\right)(\mathbf{x})$$

$$\begin{aligned}
&= \left(\int\int\int \psi\left(\mathbf{x}-\mathbf{y}\right) e^{i\mathbf{y}\cdot\mathbf{y}_1} e^{i\mathbf{y}_1\cdot\mathbf{z}} \phi\left(\mathbf{z}\right) dz dy_1 dy\right) (2\pi)^n \\
&= \left(\int\int\int \psi\left(\mathbf{x}-\mathbf{y}\right) e^{-i\mathbf{y}\cdot\tilde{\mathbf{y}}_1} e^{-i\tilde{\mathbf{y}}_1\cdot\mathbf{z}} \phi\left(\mathbf{z}\right) dz d\tilde{y}_1 dy\right) (2\pi)^n \\
&= \left(\psi * FF\phi\right)(\mathbf{x}).
\end{aligned}$$

Now for $\psi \in \mathfrak{S}$,

$$(2\pi)^{n/2} F\left(F^{-1}\phi F^{-1} f\right)(\psi) \equiv (2\pi)^{n/2}\left(F^{-1}\phi F^{-1} f\right)(F\psi) \equiv$$

$$(2\pi)^{n/2} F^{-1} f\left(F^{-1}\phi F\psi\right) \equiv (2\pi)^{n/2} f\left(F^{-1}\left(F^{-1}\phi F\psi\right)\right) =$$

$$f\left((2\pi)^{n/2} F^{-1}\left(\left(FF^{-1}F^{-1}\phi\right)(F\psi)\right)\right) \equiv$$

$$f\left(\psi * F^{-1}F^{-1}\phi\right) = f\left(\psi * FF\phi\right)$$

$$(2\pi)^{n/2} F^{-1}\left(F\phi F f\right)(\psi) \equiv (2\pi)^{n/2}\left(F\phi F f\right)\left(F^{-1}\psi\right) \equiv$$

$$(2\pi)^{n/2} F f\left(F\phi F^{-1}\psi\right) \equiv (2\pi)^{n/2} f\left(F\left(F\phi F^{-1}\psi\right)\right) =$$

by (13.23),

$$= f\left(F\left((2\pi)^{n/2}\left(F\phi F^{-1}\psi\right)\right)\right)$$

$$= f\left(F\left(F^{-1}\left(FF\phi * \psi\right)\right)\right) = f\left(\psi * FF\phi\right).$$

Comparing the above shows

$$(2\pi)^{n/2} F\left(F^{-1}\phi F^{-1} f\right) = (2\pi)^{n/2} F^{-1}\left(F\phi F f\right) \equiv f * \phi$$

and ; so,

$$(2\pi)^{n/2}\left(F^{-1}\phi F^{-1} f\right) = F^{-1}\left(f * \phi\right)$$

which proves the theorem.

13.4 Exercises

1. Suppose $f \in L^1(\mathbb{R}^n)$ and $||g_k - f||_1 \to 0$. Show Fg_k and $F^{-1}g_k$ converge uniformly to Ff and $F^{-1}f$ respectively.

2. Let $f \in L^1(\mathbb{R}^n)$,

$$Ff(\mathbf{t}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x})dx,$$

$$F^{-1}f(\mathbf{t}) \equiv (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}} f(\mathbf{x})dx.$$

Show that $F^{-1}f$ and Ff are both continuous and bounded. Show also that

$$\lim_{|\mathbf{x}|\to\infty} F^{-1}f(\mathbf{x}) = \lim_{|\mathbf{x}|\to\infty} Ff(\mathbf{x}) = 0.$$

Are the Fourier and inverse Fourier transforms of a function in $L^1(\mathbb{R}^n)$ uniformly continuous?

3. Suppose $F^{-1}f \in L^1(\mathbb{R}^n)$. Observe that just as in Theorem 13.5,

$$(F \circ F^{-1})f(\mathbf{x}) = \lim_{\varepsilon\to 0} f * m_\varepsilon(\mathbf{x}).$$

Use this to argue that if f and $F^{-1}f \in L^1(\mathbb{R}^n)$, then

$$\left(F \circ F^{-1}\right) f(\mathbf{x}) = f(\mathbf{x}) \text{ a.e. } x.$$

Similarly

$$(F^{-1} \circ F)f(\mathbf{x}) = f(\mathbf{x}) \text{ a.e.}$$

if f and $Ff \in L^1(\mathbb{R}^n)$. **Hint**: Show $f * m_\varepsilon \to f$ in $L^1(\mathbb{R}^n)$. Thus there is a subsequence $\varepsilon_k \to 0$ such that $f * m_{\varepsilon_k}(\mathbf{x}) \to f(\mathbf{x})$ a.e.

4. $\uparrow$ Show that if $F^{-1}f \in L^1$ or $Ff \in L^1$, then f equals a continuous bounded function a.e.

5. Let $f, g \in L^1(\mathbb{R}^n)$. Show $f * g \in L^1$ and $F(f * g) = (2\pi)^{n/2} FfFg$.

6. $\uparrow$ Suppose $f, g \in L^1(\mathbb{R})$ and $Ff = Fg$. Show $f = g$ a.e.

7. $\uparrow$ Suppose $f * f = f$ or $f * f = 0$ and $f \in L^1(\mathbb{R})$. Show $f = 0$.

8. For this problem define $\int_a^\infty f(t)\,dt \equiv \lim_{r\to\infty} \int_a^r f(t)\,dt$. Note this coincides with the Lebesgue integral when $f \in L^1(a, \infty)$. Show

   (a) $\int_0^\infty \frac{\sin(u)}{u} du = \frac{\pi}{2}$

   (b) $\lim_{r\to\infty} \int_\delta^\infty \frac{\sin(ru)}{u} du = 0$ whenever $\delta > 0$.

   (c) If $f \in L^1(\mathbb{R})$, then $\lim_{r\to\infty} \int_{\mathbb{R}} \sin(ru) f(u)\,du = 0$.

   Hint: For the first two, use $\frac{1}{u} = \int_0^\infty e^{-ut}dt$ and apply Fubini's theorem to $\int_0^R \sin u \int_{\mathbb{R}} e^{-ut}dtdu$. For the last part, first establish it for $f \in C_c^\infty(\mathbb{R})$ and then use the density of this set in $L^1(\mathbb{R})$ to obtain the result. This is sometimes called the Riemann Lebesgue lemma.

9. $\uparrow$Suppose that $g \in L^1(\mathbb{R})$ and that at some $x > 0$ we have that g is locally Holder continuous from the right and from the left. By this we mean

$$\lim_{r \to 0+} g(x+r) \equiv g(x+)$$

exists,

$$\lim_{r \to 0+} g(x-r) \equiv g(x-)$$

exists and there exist constants $K, \delta > 0$ and $r \in (0,1]$ such that for $|x-y| < \delta$,

$$|g(x+) - g(y)| < K|x-y|^r$$

for $y > x$ and

$$|g(x-) - g(y)| < K|x-y|^r$$

for $y < x$. Show that under these conditions,

$$\lim_{r \to \infty} \frac{2}{\pi} \int_0^\infty \frac{\sin(ur)}{u} \left(\frac{g(x-u) + g(x+u)}{2} \right) du$$

$$= \frac{g(x+) + g(x-)}{2}.$$

10. $\uparrow$ Let $g \in L^1(\mathbb{R})$ and suppose g is locally Holder continuous from the right and from the left at x. Show that then

$$\lim_{R \to \infty} \frac{1}{2\pi} \int_{-R}^{R} e^{ixt} \int_{-\infty}^{\infty} e^{-ity} g(y)\, dy dt = \frac{g(x+) + g(x-)}{2}.$$

This is very interesting. If $g \in L^2(\mathbb{R})$, this shows $F^{-1}(Fg)(x) = \frac{g(x+)+g(x-)}{2}$, the midpoint of the jump in g at the point, x. In particular, if $g \in \mathfrak{S}$, $F^{-1}(Fg) = g$. **Hint:** Show the left side of the above equation reduces to

$$\frac{2}{\pi} \int_0^\infty \frac{\sin(ur)}{u} \left(\frac{g(x-u) + g(x+u)}{2} \right) du$$

and then use Problem 9 to obtain the result.

11. $\uparrow$ We say a measurable function g defined on $(0, \infty)$ has exponential growth if $|g(t)| \le Ce^{\eta t}$ for some η. For $\operatorname{Re}(s) > \eta$, we can define the Laplace Transform by

$$Lg(s) \equiv \int_0^\infty e^{-su} g(u)\, du.$$

Assume that g has exponential growth as above and is Holder continuous from the right and from the left at t. Pick $\gamma > \eta$. Show that

$$\lim_{R \to \infty} \frac{1}{2\pi} \int_{-R}^{R} e^{\gamma t} e^{iyt} Lg(\gamma + iy)\, dy = \frac{g(t+) + g(t-)}{2}.$$

This formula is sometimes written in the form

$$\frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{st} Lg(s)\, ds$$

and is called the complex inversion integral for Laplace transforms. It can be used to find inverse Laplace transforms. **Hint:**

$$\frac{1}{2\pi}\int_{-R}^{R} e^{\gamma t}e^{iyt}Lg\left(\gamma+iy\right)dy =$$

$$\frac{1}{2\pi}\int_{-R}^{R} e^{\gamma t}e^{iyt}\int_{0}^{\infty} e^{-(\gamma+iy)u}g\left(u\right)dudy.$$

Now use Fubini's theorem and do the integral from $-R$ to R to get this equal to

$$\frac{e^{\gamma t}}{\pi}\int_{-\infty}^{\infty} e^{-\gamma u}\overline{g}\left(u\right)\frac{\sin\left(R\left(t-u\right)\right)}{t-u}du$$

where $\overline{g}$ is the zero extension of g off $[0,\infty)$. Then this equals

$$\frac{e^{\gamma t}}{\pi}\int_{-\infty}^{\infty} e^{-\gamma(t-u)}\overline{g}\left(t-u\right)\frac{\sin\left(Ru\right)}{u}du$$

which equals

$$\frac{2e^{\gamma t}}{\pi}\int_{0}^{\infty}\frac{\overline{g}\left(t-u\right)e^{-\gamma(t-u)}+\overline{g}\left(t+u\right)e^{-\gamma(t+u)}}{2}\frac{\sin\left(Ru\right)}{u}du$$

and then apply the result of Problem 9.

12. Several times in the above chapter we used an argument like the following. Suppose $\int_{\mathbb{R}^n} f\left(x\right)\phi\left(x\right)dx = 0$ for all $\phi\in\mathfrak{S}$. Therefore, $f=0$ in $L^2\left(\mathbb{R}^n\right)$. Prove the validity of this argument.

13. Suppose $f\in\mathfrak{S}$, $f_{x_j}\in L^1(\mathbb{R}^n)$. Show $F(f_{x_j})(t)=it_jFf(t)$.

14. Let $f\in\mathfrak{S}$ and let k be a positive integer.

$$||f||_{k,2}\equiv(||f||_2^2+\sum_{|\alpha|\leq k}||D^\alpha f||_2^2)^{1/2}.$$

One could also define

$$|||f|||_{k,2}\equiv(\int_{R^n}|Ff(\mathbf{x})|^2(1+|\mathbf{x}|^2)^k dx)^{1/2}.$$

Show both $||\,||_{k,2}$ and $|||\,|||_{k,2}$ are norms on $\mathfrak{S}$ and that they are equivalent. These are Sobolev space norms. What are some possible advantages of the second norm over the first? **Hint:** For which values of k do these make sense?

15. $\uparrow$ Define $H^k(\mathbb{R}^n)$ by $f\in L^2(\mathbb{R}^n)$ such that

$$(\int|Ff(x)|^2(1+|x|^2)^k dx)^{\frac{1}{2}}<\infty,$$

$$|||f|||_{k,2}\equiv(\int|Ff(x)|^2(1+|x|^2)^k dx)^{\frac{1}{2}}.$$

Show $H^k(\mathbb{R}^n)$ is a Banach space, and that if k is a positive integer, $H^k(\mathbb{R}^n)=\{\, f\in L^2(\mathbb{R}^n):$ there exists $\{u_j\}\subseteq\mathfrak{S}$ with $||u_j-f||_2\to 0$ and $\{u_j\}$ is a Cauchy sequence in $||\,||_{k,2}$ of Problem 14$\}$. This is one way to define Sobolev Spaces. **Hint:** One way to do the second part of this is to let $g_s\to Ff$ in $L^2((1+|x|^2)^k dx)$ where $g_s\in C_c(\mathbb{R}^n)$. We can do this because $(1+|x|^2)^k dx$ is a Radon measure. By convolving with a mollifier, we can, without loss of generality, assume $g_s\in C_c^\infty(\mathbb{R}^n)$. Thus $g_s=Ff_s$, $f_s\in\mathfrak{S}$. Then by Problem 14, f_s is Cauchy in the norm $||\,||_{k,2}$.

16. $\uparrow$ If $2k > n$, show that if $f \in H^k(\mathbb{R}^n)$, then f equals a bounded continuous function a.e. **Hint:** Show $Ff \in L^1(\mathbb{R}^n)$, and then use Problem 4. To do this, write

$$|Ff(x)| = |Ff(x)|(1+|x|^2)^{\frac{k}{2}}(1+|x|^2)^{\frac{-k}{2}}.$$

So

$$\int |Ff(x)|dx = \int |Ff(x)|(1+|x|^2)^{\frac{k}{2}}(1+|x|^2)^{\frac{-k}{2}}dx.$$

Use Holder's Inequality. This is an example of a Sobolev Embedding Theorem.

17. Let $u \in \mathfrak{S}$. Then we know $Fu \in \mathfrak{S}$ and so, in particular, it makes sense to form the integral,

$$\int_{\mathbb{R}} Fu\,(x', x_n)\, dx_n$$

where $(x', x_n) = x \in \mathbb{R}^n$. For $u \in \mathfrak{S}$, define $\gamma u\,(x') \equiv u\,(x', 0)$. Find a constant such that $F\,(\gamma u)\,(x')$ equals this constant times the above integral. **Hint:** By the dominated convergence theorem

$$\int_{\mathbb{R}} Fu\,(x', x_n)\, dx_n = \lim_{\varepsilon \to 0} \int_{\mathbb{R}} e^{-(\varepsilon x_n)^2} Fu\,(x', x_n)\, dx_n.$$

Now use the definition of the Fourier transform and Fubini's theorem as required in order to obtain the desired relationship.

18. Recall from the chapter on Fourier series that the Fourier series of a function in $L^2\,(-\pi, \pi)$ converges to the function in the mean square sense. Prove a similar theorem with $L^2\,(-\pi, \pi)$ replaced by $L^2\,(-m\pi, m\pi)$ and the functions $\left\{(2\pi)^{-(1/2)}\, e^{inx}\right\}_{n\in\mathbb{Z}}$ used in the Fourier series replaced with $\left\{(2m\pi)^{-(1/2)}\, e^{i\frac{n}{m}x}\right\}_{n\in\mathbb{Z}}$. Now suppose f is a function in $L^2\,(\mathbb{R})$ satisfying $Ff\,(t) = 0$ if $|t| > m\pi$. Show that if this is so, then

$$f\,(x) = \frac{1}{\pi} \sum_{n\in\mathbb{Z}} f\left(\frac{-n}{m}\right) \frac{\sin\left(\pi\,(mx+n)\right)}{mx+\pi}.$$

Here m is a positive integer. This is sometimes called the Shannon sampling theorem.

Banach Spaces

14.1 Baire category theorem

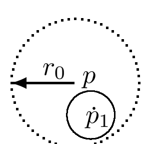
Functional analysis is the study of various types of vector spaces which are also topological spaces and the linear operators defined on these spaces. As such, it is really a generalization of linear algebra and calculus. The vector spaces which are of interest in this subject include the usual spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ but also many which are infinite dimensional such as the space $C(X;\mathbb{R}^n)$ discussed in Chapter 4 in which we think of a function as a point or a vector. When the topology comes from a norm, the vector space is called a normed linear space and this is the case of interest here. A normed linear space is called real if the field of scalars is $\mathbb{R}$ and complex if the field of scalars is $\mathbb{C}$. We will assume a linear space is complex unless stated otherwise. A normed linear space may be considered as a metric space if we define $d(x,y) \equiv ||x-y||$. As usual, if every Cauchy sequence converges, the metric space is called complete.

Definition 14.1 *A complete normed linear space is called a Banach space.*

The purpose of this chapter is to prove some of the most important theorems about Banach spaces. The next theorem is called the Baire category theorem and it will be used in the proofs of many of the other theorems.

Theorem 14.2 *Let (X,d) be a complete metric space and let $\{U_n\}_{n=1}^{\infty}$ be a sequence of open subsets of X satisfying $\overline{U_n} = X$ (U_n is dense). Then $D \equiv \cap_{n=1}^{\infty} U_n$ is a dense subset of X.*

Proof: Let $p \in X$ and let $r_0 > 0$. We need to show $D \cap B(p,r_0) \neq \emptyset$. Since U_1 is dense, there exists $p_1 \in U_1 \cap B(p,r_0)$, an open set. Let $p_1 \in B(p_1,r_1) \subseteq \overline{B(p_1,r_1)} \subseteq U_1 \cap B(p,r_0)$ and $r_1 < 2^{-1}$. We are using Theorem 3.14.



There exists $p_2 \in U_2 \cap B(p_1,r_1)$ because U_2 is dense. Let

$$p_2 \in B(p_2,r_2) \subseteq \overline{B(p_2,r_2)} \subseteq U_2 \cap B(p_1,r_1) \subseteq U_1 \cap U_2 \cap B(p,r_0).$$

and let $r_2 < 2^{-2}$. Continue in this way. Thus

$$r_n < 2^{-n},$$

$$\overline{B(p_n,r_n)} \subseteq U_1 \cap U_2 \cap ... \cap U_n \cap B(p,r_0),$$

$$\overline{B(p_n, r_n)} \subseteq B(p_{n-1}, r_{n-1}).$$

Consider the Cauchy sequence, $\{p_n\}$. Since X is complete, let

$$\lim_{n\to\infty} p_n = p_\infty.$$

Since all but finitely many terms of $\{p_n\}$ are in $\overline{B(p_m, r_m)}$, it follows that $p_\infty \in \overline{B(p_m, r_m)}$. Since this holds for every m,

$$p_\infty \in \cap_{m=1}^\infty \overline{B(p_m, r_m)} \subseteq \cap_{i=1}^\infty U_i \cap B(p, r_0).$$

This proves the theorem.

Corollary 14.3 *Let X be a complete metric space and suppose $X = \cup_{i=1}^\infty F_i$ where each F_i is a closed set. Then for some i, interior $F_i \neq \emptyset$.*

The set D of Theorem 14.2 is called a G_δ set because it is the countable intersection of open sets. Thus D is a dense G_δ set.

Recall that a norm satisfies:

a.) $||x|| \geq 0$, $||x|| = 0$ if and only if $x = 0$.

b.) $||x + y|| \leq ||x|| + ||y||$.

c.) $||cx|| = |c|\ ||x||$ if c is a scalar and $x \in X$.

We also recall the following lemma which gives a simple way to tell if a function mapping a metric space to a metric space is continuous.

Lemma 14.4 *If $(X, d), (Y, p)$ are metric spaces, f is continuous at x if and only if*

$$\lim_{n\to\infty} x_n = x$$

implies

$$\lim_{n\to\infty} f(x_n) = f(x).$$

The proof is left to the reader and follows quickly from the definition of continuity. See Problem 5 of Chapter 3. For the sake of simplicity, we will write $x_n \to x$ sometimes instead of $\lim_{n\to\infty} x_n = x$.

Theorem 14.5 *Let X and Y be two normed linear spaces and let $L : X \to Y$ be linear ($L(ax + by) = aL(x) + bL(y)$ for a, b scalars and $x, y \in X$). The following are equivalent*

a.) L is continuous at 0

b.) L is continuous

c.) There exists $K > 0$ such that $||Lx||_Y \leq K\ ||x||_X$ for all $x \in X$ (L is bounded).

Proof: a.)$\Rightarrow$b.) Let $x_n \to x$. Then $(x_n - x) \to 0$. It follows $Lx_n - Lx \to 0$ so $Lx_n \to Lx$. b.)$\Rightarrow$c.) Since L is continuous, L is continuous at 0. Hence $||Lx||_Y < 1$ whenever $||x||_X \leq \delta$ for some δ. Therefore, suppressing the subscript on the $||\ ||$,

$$||L\left(\frac{\delta x}{||x||}\right)|| \leq 1.$$

Hence

$$||Lx|| \leq \frac{1}{\delta}||x||.$$

c.)$\Rightarrow$a.) is obvious.

Definition 14.6 *Let $L : X \to Y$ be linear and continuous where X and Y are normed linear spaces. We denote the set of all such continuous linear maps by $\mathcal{L}(X, Y)$ and define*

$$||L|| = \sup\{||Lx|| : ||x|| \leq 1\}. \tag{14.1}$$

The proof of the next lemma is left to the reader.

Lemma 14.7 *With $||L||$ defined in (14.1), $\mathcal{L}(X, Y)$ is a normed linear space. Also $||Lx|| \leq ||L|| \, ||x||$.*

For example, we could consider the space of linear transformations defined on $\mathbb{R}^n$ having values in $\mathbb{R}^m$, and the above gives a way to measure the distance between two linear transformations. In this case, the linear transformations are all continuous because if L is such a linear transformation, and $\{\mathbf{e}_k\}_{k=1}^n$ and $\{\mathbf{e}_i\}_{i=1}^m$ are the standard basis vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, there are scalars l_{ik} such that

$$L(\mathbf{e}_k) = \sum_{i=1}^m l_{ik}\mathbf{e}_i.$$

Thus, letting $\mathbf{a} = \sum_{k=1}^n a_k \mathbf{e}_k$,

$$L(\mathbf{a}) = L\left(\sum_{k=1}^n a_k \mathbf{e}_k\right) = \sum_{k=1}^n a_k L(\mathbf{e}_k) = \sum_{k=1}^n \sum_{i=1}^m a_k l_{ik} \mathbf{e}_i.$$

Consequently, letting $K \geq |l_{ik}|$ for all i, k,

$$\begin{aligned} ||L\mathbf{a}|| &\leq \left(\sum_{i=1}^m \left|\sum_{k=1}^n a_k l_{ik}\right|^2\right)^{1/2} \leq Km^{1/2}n \left(\max\{|a_k|, k = 1, \cdots, n\}\right) \\ &\leq Km^{1/2}n \left(\sum_{k=1}^n a_k^2\right)^{1/2} = Km^{1/2}n \, ||\mathbf{a}||. \end{aligned}$$

This type of thing occurs whenever one is dealing with a linear transformation between finite dimensional normed linear spaces. Thus, in finite dimensions the algebraic condition that an operator is linear is sufficient to imply the topological condition that the operator is continuous. The situation is not so simple in infinite dimensional spaces such as $C(X; \mathbb{R}^n)$. This is why we impose the topological condition of continuity as a criterion for membership in $\mathcal{L}(X, Y)$ in addition to the algebraic condition of linearity.

Theorem 14.8 *If Y is a Banach space, then $\mathcal{L}(X, Y)$ is also a Banach space.*

Proof: Let $\{L_n\}$ be a Cauchy sequence in $\mathcal{L}(X, Y)$ and let $x \in X$.

$$||L_n x - L_m x|| \leq ||x|| \, ||L_n - L_m||.$$

Thus $\{L_n x\}$ is a Cauchy sequence. Let

$$Lx = \lim_{n \to \infty} L_n x.$$

Then, clearly, L is linear. Also L is continuous. To see this, note that $\{||L_n||\}$ is a Cauchy sequence of real numbers because $|||L_n|| - ||L_m||| \leq ||L_n - L_m||$. Hence there exists $K > \sup\{||L_n|| : n \in \mathbb{N}\}$. Thus, if $x \in X$,

$$||Lx|| = \lim_{n \to \infty} ||L_n x|| \leq K||x||.$$

This proves the theorem.

14.2 Uniform boundedness closed graph and open mapping theorems

The next big result is sometimes called the Uniform Boundedness theorem, or the Banach-Steinhaus theorem. This is a very surprising theorem which implies that for a collection of bounded linear operators, if they are bounded pointwise, then they are also bounded uniformly. As an example of a situation in which pointwise bounded does not imply uniformly bounded, consider the functions $f_\alpha(x) \equiv \mathcal{X}_{(\alpha,1)}(x)\, x^{-1}$ for $\alpha \in (0,1)$ and $\mathcal{X}_{(\alpha,1)}(x)$ equals zero if $x \notin (\alpha,1)$ and one if $x \in (\alpha,1)$. Clearly each function is bounded and the collection of functions is bounded at each point of $(0,1)$, but there is no bound for all the functions taken together.

Theorem 14.9 *Let X be a Banach space and let Y be a normed linear space. Let $\{L_\alpha\}_{\alpha\in\Lambda}$ be a collection of elements of $\mathcal{L}(X,Y)$. Then one of the following happens.*

a.) $\sup\{||L_\alpha|| \ : \alpha \in \Lambda\} < \infty$

b.) There exists a dense G_δ set, D, such that for all $x \in D$,

$$\sup\{||L_\alpha x|| \ \alpha \in \Lambda\} = \infty.$$

Proof: For each $n \in \mathbb{N}$, define

$$U_n = \{x \in X : \sup\{||L_\alpha x|| : \alpha \in \Lambda\} > n\}.$$

Then U_n is an open set. Case b.) is obtained from Theorem 14.2 if each U_n is dense. The other case is that for some n, U_n is not dense. If this occurs, there exists x_0 and $r > 0$ such that for all $x \in B(x_0,r)$, $||L_\alpha x|| \leq n$ for all α. Now if $y \in B(0,r)$, $x_0 + y \in B(x_0,r)$. Consequently, for all such y, $||L_\alpha(x_0+y)|| \leq n$. This implies that for such y and all α,

$$||L_\alpha y|| \leq n + ||L_\alpha(x_0)|| \leq 2n.$$

Hence $||L_\alpha|| \leq \frac{2n}{r}$ for all α, and we obtain case a.).

The next theorem is called the Open Mapping theorem. Unlike Theorem 14.9 it requires both X and Y to be Banach spaces.

Theorem 14.10 *Let X and Y be Banach spaces, let $L \in \mathcal{L}(X,Y)$, and suppose L is onto. Then L maps open sets onto open sets.*

To aid in the proof of this important theorem, we give a lemma.

Lemma 14.11 *Let a and b be positive constants and suppose*

$$B(0,a) \subseteq \overline{L(B(0,b))}.$$

Then

$$\overline{L(B(0,b))} \subseteq L(B(0,2b)).$$

Proof of Lemma 14.11: Let $y \in \overline{L(B(0,b))}$. Pick $x_1 \in B(0,b)$ such that $||y - Lx_1|| < \frac{a}{2}$. Now

$$2y - 2Lx_1 \in B(0,a) \subseteq \overline{L(B(0,b))}.$$

Then choose $x_2 \in B(0,b)$ such that $||2y - 2Lx_1 - Lx_2|| < a/2$. Thus $||y - Lx_1 - L\left(\frac{x_2}{2}\right)|| < a/2^2$. Continuing in this way, we pick $x_3, x_4, ...$ in $B(0,b)$ such that

$$||y - \sum_{i=1}^{n} 2^{-(i-1)} L(x_i)|| = ||y - L\sum_{i=1}^{n} 2^{-(i-1)} x_i|| < a2^{-n}. \tag{14.2}$$

Let $x = \sum_{i=1}^{\infty} 2^{-(i-1)} x_i$. The series converges because X is complete and

$$||\sum_{i=m}^{n} 2^{-(i-1)} x_i|| \leq b \sum_{i=m}^{\infty} 2^{-(i-1)} = b\, 2^{-m+2}.$$

Thus the sequence of partial sums is Cauchy. Letting $n \to \infty$ in (14.2) yields $||y - Lx|| = 0$. Now

$$||x|| = \lim_{n\to\infty} ||\sum_{i=1}^{n} 2^{-(i-1)} x_i||$$

$$\leq \lim_{n\to\infty} \sum_{i=1}^{n} 2^{-(i-1)} ||x_i|| < \lim_{n\to\infty} \sum_{i=1}^{n} 2^{-(i-1)} b = 2b.$$

This proves the lemma.

Proof of Theorem 14.10: $Y = \cup_{n=1}^{\infty} \overline{L(B(0,n))}$. By Corollary 14.3, the set, $\overline{L(B(0,n_0))}$ has nonempty interior for some n_0. Thus $B(y,r) \subseteq \overline{L(B(0,n_0))}$ for some y and some $r > 0$. Since L is linear $B(-y,r) \subseteq \overline{L(B(0,n_0))}$ also (why?). Therefore

$$\begin{aligned} B(0,r) &\subseteq B(y,r) + B(-y,r) \\ &\equiv \{x + z : x \in B\,(y,r) \text{ and } z \in B\,(-y,r)\} \\ &\subseteq \overline{L(B(0,2n_0))} \end{aligned}$$

By Lemma 14.11, $\overline{L(B(0,2n_0))} \subseteq L(B(0,4n_0))$. Letting $a = r(4n_0)^{-1}$, it follows, since L is linear, that $B(0,a) \subseteq L(B(0,1))$.

Now let U be open in X and let $x + B(0,r) = B(x,r) \subseteq U$. Then

$$L(U) \supseteq L(x + B(0,r))$$

$$= Lx + L(B(0,r)) \supseteq Lx + B(0,ar) = B(Lx, ar)$$

$(L(B(0,r)) \supseteq B(0,ar)$ because $L(B(0,1)) \supseteq B(0,a)$ and L is linear). Hence

$$Lx \in B(Lx, ar) \subseteq L(U).$$

This shows that every point, $Lx \in LU$, is an interior point of LU and so LU is open. This proves the theorem.

This theorem is surprising because it implies that if $|\cdot|$ and $||\cdot||$ are two norms with respect to which a vector space X is a Banach space such that $|\cdot| \leq K\,||\cdot||$, then there exists a constant k, such that $||\cdot|| \leq k\,|\cdot|$. This can be useful because sometimes it is not clear how to compute k when all that is needed is its existence. To see the open mapping theorem implies this, consider the identity map $ix = x$. Then $i : (X, ||\cdot||) \to (X, |\cdot|)$ is continuous and onto. Hence i is an open map which implies i^{-1} is continuous. This gives the existence of the constant k. Of course there are many other situations where this theorem will be of use.

Definition 14.12 *Let $f : D \to E$. The set of all ordered pairs of the form $\{(x, f(x)) : x \in D\}$ is called the graph of f.*

Definition 14.13 *If X and Y are normed linear spaces, we make $X \times Y$ into a normed linear space by using the norm $||(x,y)|| = ||x|| + ||y||$ along with component-wise addition and scalar multiplication. Thus $a(x,y) + b(z,w) \equiv (ax + bz, ay + bw)$.*

There are other ways to give a norm for $X \times Y$. See Problem 5 for some alternatives.

Lemma 14.14 *The norm defined in Definition 14.13 on $X \times Y$ along with the definition of addition and scalar multiplication given there make $X \times Y$ into a normed linear space. Furthermore, the topology induced by this norm is identical to the product topology defined in Chapter 3.*

Lemma 14.15 *If X and Y are Banach spaces, then $X \times Y$ with the norm and vector space operations defined in Definition 14.13 is also a Banach space.*

Lemma 14.16 *Every closed subspace of a Banach space is a Banach space.*

Definition 14.17 *Let X and Y be Banach spaces and let $D \subseteq X$ be a subspace. A linear map $L : D \to Y$ is said to be closed if its graph is a closed subspace of $X \times Y$. Equivalently, L is closed if $x_n \to x$ and $Lx_n \to y$ implies $x \in D$ and $y = Lx$.*

Note the distinction between closed and continuous. If the operator is closed the assertion that $y = Lx$ only follows if it is known that the sequence $\{Lx_n\}$ converges. In the case of a continuous operator, the convergence of $\{Lx_n\}$ follows from the assumption that $x_n \to x$. It is not always the case that a mapping which is closed is necessarily continuous. Consider the function $f(x) = \tan(x)$ if x is not an odd multiple of $\frac{\pi}{2}$ and $f(x) \equiv 0$ at every odd multiple of $\frac{\pi}{2}$. Then the graph is closed and the function is defined on $\mathbb{R}$ but it clearly fails to be continuous. The next theorem, the closed graph theorem, gives conditions under which closed implies continuous.

Theorem 14.18 *Let X and Y be Banach spaces and suppose $L : X \to Y$ is closed and linear. Then L is continuous.*

Proof: Let G be the graph of L. $G = \{(x, Lx) : x \in X\}$. Define $P : G \to X$ by $P(x, Lx) = x$. P maps the Banach space G onto the Banach space X and is continuous and linear. By the open mapping theorem, P maps open sets onto open sets. Since P is also 1-1, this says that P^{-1} is continuous. Thus $||P^{-1}x|| \leq K||x||$. Hence

$$||x|| + ||Lx|| \leq K||x||$$

and so $||Lx|| \leq (K-1)||x||$. This shows L is continuous and proves the theorem.

14.3 Hahn Banach theorem

The closed graph, open mapping, and uniform boundedness theorems are the three major topological theorems in functional analysis. The other major theorem is the Hahn-Banach theorem which has nothing to do with topology. Before presenting this theorem, we need some preliminaries.

Definition 14.19 *Let $\mathcal{F}$ be a nonempty set. $\mathcal{F}$ is called a partially ordered set if there is a relation, denoted here by $\leq$, such that*

$$x \leq x \text{ for all } x \in \mathcal{F}.$$

$$\text{If } x \leq y \text{ and } y \leq z \text{ then } x \leq z.$$

$\mathcal{C} \subseteq \mathcal{F}$ is said to be a chain if every two elements of $\mathcal{C}$ are related. By this we mean that if $x, y \in \mathcal{C}$, then either $x \leq y$ or $y \leq x$. Sometimes we call a chain a totally ordered set. $\mathcal{C}$ is said to be a maximal chain if whenever $\mathcal{D}$ is a chain containing $\mathcal{C}$, $\mathcal{D} = \mathcal{C}$.

The most common example of a partially ordered set is the power set of a given set with $\subseteq$ being the relation. The following theorem is equivalent to the axiom of choice. For a discussion of this, see the appendix on the subject.

Theorem 14.20 *(Hausdorff Maximal Principle) Let $\mathcal{F}$ be a nonempty partially ordered set. Then there exists a maximal chain.*

Definition 14.21 *Let X be a real vector space $\rho : X \to \mathbb{R}$ is called a gauge function if*

$$\rho(x+y) \le \rho(x) + \rho(y),$$

$$\rho(ax) = a\rho(x) \text{ if } a \ge 0. \tag{14.3}$$

Suppose M is a subspace of X and $z \notin M$. Suppose also that f is a linear real-valued function having the property that $f(x) \le \rho(x)$ for all $x \in M$. We want to consider the problem of extending f to $M \oplus \mathbb{R}z$ such that if F is the extended function, $F(y) \le \rho(y)$ for all $y \in M \oplus \mathbb{R}z$ and F is linear. Since F is to be linear, we see that we only need to determine how to define $F(z)$. Letting $a > 0$, we need to have the following hold for all $x, y \in M$.

$$F(x+az) \le \rho(x+az),\ F(y-az) \le \rho(y-az).$$

Multiplying by a^{-1} using the fact that M is a subspace, and (14.3), we see this is the same as

$$f(x) + F(z) \le \rho(x+z),\ f(y) - \rho(y-z) \le F(z)$$

for all $x, y \in M$. Hence we need to have $F(z)$ such that for all $x, y \in M$

$$f(y) - \rho(y-z) \le F(z) \le \rho(x+z) - f(x). \tag{14.4}$$

Is there any such number between $f(y) - \rho(y-z)$ and $\rho(x+z) - f(x)$ for every pair $x, y \in M$? This is where we use that $f(x) \le \rho(x)$ on M. For $x, y \in M$,

$$\rho(x+z) - f(x) - [f(y) - \rho(y-z)]$$

$$= \rho(x+z) + \rho(y-z) - (f(x) + f(y))$$

$$\ge \rho(x+y) - f(x+y) \ge 0.$$

Therefore there exists a number between

$$\sup\{f(y) - \rho(y-z) : y \in M\}$$

and

$$\inf\{\rho(x+z) - f(x) : x \in M\}$$

We choose $F(z)$ to satisfy (14.4). With this preparation, we state a simple lemma which will be used to prove the Hahn Banach theorem.

Lemma 14.22 *Let M be a subspace of X, a real linear space, and let ρ be a gauge function on X. Suppose $f : M \to \mathbb{R}$ is linear and $z \notin M$, and $f(x) \le \rho(x)$ for all $x \in M$. Then f can be extended to $M \oplus \mathbb{R}z$ such that, if F is the extended function, F is linear and $F(x) \le \rho(x)$ for all $x \in M \oplus \mathbb{R}z$.*

Proof: Let $f(y) - \rho(y-z) \le F(z) \le \rho(x+z) - f(x)$ for all $x, y \in M$ and let $F(x+az) = f(x) + aF(z)$ whenever $x \in M$, $a \in \mathbb{R}$. If $a > 0$

$$\begin{aligned} F(x+az) &= f(x) + aF(z) \\ &\le f(x) + a\left[\rho\left(\frac{x}{a} + z\right) - f\left(\frac{x}{a}\right)\right] \\ &= \rho(x+az). \end{aligned}$$

If $a < 0$,

$$\begin{aligned} F(x + az) &= f(x) + aF(z) \\ &\leq f(x) + a\left[f\left(\frac{-x}{a}\right) - \rho\left(\frac{-x}{a} - z\right)\right] \\ &= f(x) - f(x) + \rho(x + az) = \rho(x + az). \end{aligned}$$

This proves the lemma.

Theorem 14.23 *(Hahn Banach theorem) Let X be a real vector space, let M be a subspace of X, let $f : M \to \mathbb{R}$ be linear, let ρ be a gauge function on X, and suppose $f(x) \leq \rho(x)$ for all $x \in M$. Then there exists a linear function, $F : X \to \mathbb{R}$, such that*
a.) $F(x) = f(x)$ for all $x \in M$
b.) $F(x) \leq \rho(x)$ for all $x \in X$.

Proof: Let $\mathcal{F} = \{(V, g) : V \supseteq M,\ V$ is a subspace of $X,\ g : V \to \mathbb{R}$ is linear, $g(x) = f(x)$ for all $x \in M$, and $g(x) \leq \rho(x)\}$. Then $(M, f) \in \mathcal{F}$ so $\mathcal{F} \neq \emptyset$. Define a partial order by the following rule.

$$(V, g) \leq (W, h)$$

means

$$V \subseteq W \text{ and } h(x) = g(x) \text{ if } x \in V.$$

Let $\mathcal{C}$ be a maximal chain in $\mathcal{F}$ (Hausdorff Maximal theorem). Let $Y = \cup\{V : (V, g) \in \mathcal{C}\}$. Let $h : Y \to \mathbb{R}$ be defined by $h(x) = g(x)$ where $x \in V$ and $(V, g) \in \mathcal{C}$. This is well defined since $\mathcal{C}$ is a chain. Also h is clearly linear and $h(x) \leq \rho(x)$ for all $x \in Y$. We want to argue that $Y = X$. If not, there exists $z \in X \setminus Y$ and we can extend h to $Y \oplus \mathbb{R}z$ using Lemma 14.22. But this will contradict the maximality of $\mathcal{C}$. Indeed, $\mathcal{C} \cup \{\left(Y \oplus \mathbb{R}z,\ \overline{h}\right)\}$ would be a longer chain where $\overline{h}$ is the extended h. This proves the Hahn Banach theorem.

This is the original version of the theorem. There is also a version of this theorem for complex vector spaces which is based on a trick.

Corollary 14.24 *(Hahn Banach) Let M be a subspace of a complex normed linear space, X, and suppose $f : M \to \mathbb{C}$ is linear and satisfies $|f(x)| \leq K||x||$ for all $x \in M$. Then there exists a linear function, F, defined on all of X such that $F(x) = f(x)$ for all $x \in M$ and $|F(x)| \leq K||x||$ for all x.*

Proof: First note $f(x) = \operatorname{Re} f(x) + i \operatorname{Im} f(x)$ and so

$$\operatorname{Re} f(ix) + i\ \operatorname{Im} f(ix) = f(ix) = i\ f(x) = i\ \operatorname{Re} f(x) - \operatorname{Im} f(x).$$

Therefore, $\operatorname{Im} f(x) = -\operatorname{Re} f(ix)$, and we may write

$$f(x) = \operatorname{Re} f(x) - i \operatorname{Re} f(ix).$$

If c is a real scalar

$$\operatorname{Re} f(cx) - i\ \operatorname{Re} f(icx) = cf(x) = c \operatorname{Re} f(x) - i\ c \operatorname{Re} f(ix).$$

Thus $\operatorname{Re} f(cx) = c \operatorname{Re} f(x)$. It is also clear that $\operatorname{Re} f(x+y) = \operatorname{Re} f(x) + \operatorname{Re} f(y)$. Consider X as a real vector space and let $\rho(x) = K||x||$. Then for all $x \in M$,

$$|\operatorname{Re} f(x)| \leq K||x|| = \rho(x).$$

From Theorem 14.23, $\operatorname{Re} f$ may be extended to a function, h which satisfies

$$\begin{aligned} h(ax+by) &= ah(x)+bh(y) \text{ if } a,b\in\mathbb{R} \\ |h(x)| &\leq K||x|| \text{ for all } x\in X. \end{aligned}$$

Let

$$F(x) \equiv h(x) - i\, h(ix).$$

It is routine to show F is linear. Now $wF(x) = |F(x)|$ for some $|w| = 1$. Therefore

$$\begin{aligned} |F(x)| &= wF(x) = h(wx) - i\, h(iwx) = h(wx) \\ &= |h(wx)| \leq K||x||. \end{aligned}$$

This proves the corollary.

Definition 14.25 *Let X be a Banach space. We denote by X' the space $\mathcal{L}(X,\mathbb{C})$. By Theorem 14.8, X' is a Banach space. Remember*

$$||f|| = \sup\{|f(x)| : ||x|| \leq 1\}$$

for $f \in X'$. We call X' the dual space.

Definition 14.26 *Let X and Y be Banach spaces and suppose $L \in \mathcal{L}(X,Y)$. Then we define the adjoint map in $\mathcal{L}(Y',X')$, denoted by L^*, by*

$$L^*y^*(x) \equiv y^*(Lx)$$

for all $y^* \in Y'$.

$$\begin{array}{ccc} & L^* & \\ X' & \leftarrow & Y' \\ X & \rightarrow & Y \\ & L & \end{array}$$

In terms of linear algebra, this adjoint map is algebraically very similar to, and is in fact a generalization of, the transpose of a matrix considered as a map on $\mathbb{R}^n$. Recall that if A is such a matrix, A^T satisfies $A^T\mathbf{x}\cdot\mathbf{y} = \mathbf{x}\cdot A\mathbf{y}$. In the case of $\mathbb{C}^n$ the adjoint is similar to the conjugate transpose of the matrix and it behaves the same with respect to the complex inner product on $\mathbb{C}^n$. What is being done here is to generalize this algebraic concept to arbitrary Banach spaces.

Theorem 14.27 *Let $L \in \mathcal{L}(X,Y)$ where X and Y are Banach spaces. Then*
a.) $L^* \in \mathcal{L}(Y',X')$ as claimed and $||L^*|| \leq ||L||$.
b.) If L is 1-1 onto a closed subspace of Y, then L^* is onto.
c.) If L is onto a dense subset of Y, then L^* is 1-1.

Proof: Clearly L^*y^* is linear and L^* is also a linear map.

$$\begin{aligned} ||L^*|| &= \sup_{||y^*||\leq 1} ||L^*y^*|| = \sup_{||y^*||\leq 1}\sup_{||x||\leq 1} |L^*y^*(x)| \\ &= \sup_{||y^*||\leq 1}\sup_{||x||\leq 1} |y^*(Lx)| \leq \sup_{||x||\leq 1} |(Lx)| = ||L|| \end{aligned}$$

Hence, $||L^*|| \leq ||L||$ and this shows part a.).

If L is 1-1 and onto a closed subset of Y, then we can apply the Open Mapping theorem to conclude that $L^{-1} : L(X) \to X$ is continuous. Hence

$$||x|| = ||L^{-1}Lx|| \leq K||Lx||$$

for some K. Now let $x^* \in X'$ be given. Define $f \in \mathcal{L}(L(X), \mathbb{C})$ by $f(Lx) = x^*(x)$. Since L is 1-1, it follows that f is linear and well defined. Also

$$|f(Lx)| = |x^*(x)| \leq ||x^*|| \, ||x|| \leq K||x^*|| \, ||Lx||.$$

By the Hahn Banach theorem, we can extend f to an element $y^* \in Y'$ such that $||y^*|| \leq K||x^*||$. Then

$$L^*y^*(x) = y^*(Lx) = f(Lx) = x^*(x)$$

so $L^*y^* = x^*$ and we have shown L^* is onto. This shows b.).

Now suppose LX is dense in Y. If $L^*y^* = 0$, then $y^*(Lx) = 0$ for all x. Since LX is dense, this can only happen if $y^* = 0$. Hence L^* is 1-1.

Corollary 14.28 *Suppose X and Y are Banach spaces, $L \in \mathcal{L}(X, Y)$, and L is 1-1 and onto. Then L^* is also 1-1 and onto.*

There exists a natural mapping from a normed linear space, X, to the dual of the dual space.

Definition 14.29 *Define $J : X \to X''$ by $J(x)(x^*) = x^*(x)$. This map is called the James map.*

Theorem 14.30 *The map, J, has the following properties.*

a.) J is 1-1 and linear.

b.) $||Jx|| = ||x||$ and $||J|| = 1$.

c.) $J(X)$ is a closed subspace of X'' if X is complete.

Also if $x^* \in X'$,

$$||x^*|| = \sup\left\{|x^{**}(x^*)| : ||x^{**}|| \leq 1, \ x^{**} \in X''\right\}.$$

Proof: To prove this, we will use a simple but useful lemma which depends on the Hahn Banach theorem.

Lemma 14.31 *Let X be a normed linear space and let $x \in X$. Then there exists $x^* \in X'$ such that $||x^*|| = 1$ and $x^*(x) = ||x||$.*

Proof: Let $f : \mathbb{C}x \to \mathbb{C}$ be defined by $f(\alpha x) = \alpha||x||$. Then for $y \in \mathbb{C}x$, $|f(y)| \leq \|y\|$. By the Hahn Banach theorem, there exists $x^* \in X'$ such that $x^*(\alpha x) = f(\alpha x)$ and $||x^*|| \leq 1$. Since $x^*(x) = ||x||$ it follows that $||x^*|| = 1$. This proves the lemma.

Now we prove the theorem. It is obvious that J is linear. If $Jx = 0$, then let $x^*(x) = ||x||$ with $||x^*|| = 1$.

$$0 = J(x)(x^*) = x^*(x) = ||x||.$$

This shows a.). To show b.), let $x \in X$ and $x^*(x) = ||x||$ with $||x^*|| = 1$. Then

$$\begin{aligned} ||x|| &\geq \sup\{|y^*(x)| : ||y^*|| \leq 1\} = \sup\{|J(x)(y^*)| : ||y^*|| \leq 1\} = ||Jx|| \\ &\geq |J(x)(x^*)| = |x^*(x)| = ||x|| \end{aligned}$$

$$||J|| = \sup\{||Jx|| : ||x|| \leq 1\} = \sup\{||x|| : ||x|| \leq 1\} = 1.$$

This shows b.). To verify c.), use b.). If $Jx_n \to y^{**} \in X''$ then by b.), x_n is a Cauchy sequence converging to some $x \in X$. Then $Jx = \lim_{n\to\infty} Jx_n = y^{**}$.

Finally, to show the assertion about the norm of x^*, use what was just shown applied to the James map from X' to X'''. More specifically,

$$||x^*|| = \sup\{|x^*(x)| : ||x|| \leq 1\} = \sup\{|J(x)(x^*)| : ||Jx|| \leq 1\}$$

$$\leq \sup\{|x^{**}(x^*)| : ||x^{**}|| \leq 1\} = \sup\{|J(x^*)(x^{**})| : ||x^{**}|| \leq 1\}$$

$$\equiv ||Jx^*|| = ||x^*||.$$

This proves the theorem.

Definition 14.32 *When J maps X onto X'', we say that X is Reflexive.*

Later on we will give examples of reflexive spaces. In particular, it will be shown that the space of square integrable and *pth* power integrable functions for $p > 1$ are reflexive.

14.4 Exercises

1. Show that no countable dense subset of $\mathbb{R}$ is a G_δ set. In particular, the rational numbers are not a G_δ set.

2. $\uparrow$ Let $f : \mathbb{R} \to \mathbb{C}$ be a function. Let $\omega_r f(x) = \sup\{|f(z) - f(y)| : y, z \in B(x,r)\}$. Let $\omega f(x) = \lim_{r \to 0} \omega_r f(x)$. Show f is continuous at x if and only if $\omega f(x) = 0$. Then show the set of points where f is continuous is a G_δ set (try $U_n = \{x : \omega f(x) < \frac{1}{n}\}$). Does there exist a function continuous at only the rational numbers? Does there exist a function continuous at every irrational and discontinuous elsewhere? **Hint:** Suppose D is any countable set, $D = \{d_i\}_{i=1}^\infty$, and define the function, $f_n(x)$ to equal zero for every $x \notin \{d_1, \cdots, d_n\}$ and 2^{-n} for x in this finite set. Then consider $g(x) \equiv \sum_{n=1}^\infty f_n(x)$. Show that this series converges uniformly.

3. Let $f \in C([0,1])$ and suppose $f'(x)$ exists. Show there exists a constant, K, such that $|f(x) - f(y)| \leq K|x - y|$ for all $y \in [0,1]$. Let $U_n = \{f \in C([0,1])$ such that for each $x \in [0,1]$ there exists $y \in [0,1]$ such that $|f(x) - f(y)| > n|x - y|\}$. Show that U_n is open and dense in $C([0,1])$ where for $f \in C([0,1])$,

$$||f|| \equiv \sup\{|f(x)| : x \in [0,1]\}.$$

   Show that if $f \in C([0,1])$, there exists $g \in C([0,1])$ such that $||g - f|| < \varepsilon$ but $g'(x)$ does not exist for any $x \in [0,1]$.

4. Let X be a normed linear space and suppose $A \subseteq X$ is "weakly bounded". This means that for each $x^* \in X'$, $\sup\{|x^*(x)| : x \in A\} < \infty$. Show A is bounded. That is, show $\sup\{||x|| : x \in A\} < \infty$.

5. Let X and Y be two Banach spaces. Define the norm

$$|||(x,y)||| \equiv \max(||x||_X, ||y||_Y).$$

   Show this is a norm on $X \times Y$ which is equivalent to the norm given in the chapter for $X \times Y$. Can you do the same for the norm defined by

$$|(x,y)| \equiv \left(||x||_X^2 + ||y||_Y^2\right)^{1/2}?$$

6. Prove Lemmas 14.14 - 14.16.

7. Let $f : \mathbb{R} \to \mathbb{C}$ be continuous and periodic with period 2π. That is $f(x + 2\pi) = f(x)$ for all x. Then it is not unreasonable to try to write $f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}$. Find what a_n should be. **Hint:** Multiply both sides by e^{-imx} and do $\int_{-\pi}^{\pi}$. Pretend there is no problem writing $\int \sum = \sum \int$. Recall the series which results is called a Fourier series.

8. ↑ If you did 7 correctly, you found

$$a_n = \left(\int_{-\pi}^{\pi} f(x) e^{-inx} dx \right) (2\pi)^{-1}.$$

The n^{th} partial sum will be denoted by $S_n f$ and defined by $S_n f(x) = \sum_{k=-n}^{n} a_k e^{ikx}$. Show $S_n f(x) = \int_{-\pi}^{\pi} f(y) D_n(x - y) dy$ where

$$D_n(t) = \frac{\sin((n + \frac{1}{2})t)}{2\pi \sin(\frac{t}{2})}.$$

This is called the Dirichlet kernel. If you have trouble, review the chapter on Fourier series.

9. ↑ Let $Y = \{f$ such that f is continuous, defined on $\mathbb{R}$, and 2π periodic$\}$. Define $||f||_Y = \sup\{|f(x)| : x \in [-\pi, \pi]\}$. Show that $(Y, || \, ||_Y)$ is a Banach space. Let $x \in \mathbb{R}$ and define $L_n(f) = S_n f(x)$. Show $L_n \in Y'$ but $\lim_{n\to\infty} ||L_n|| = \infty$. **Hint:** Let $f(y)$ approximate $sign(D_n(x - y))$.

10. ↑ Show there exists a dense G_δ subset of Y such that for f in this set, $|S_n f(x)|$ is unbounded. Show there is a dense G_δ subset of Y having the property that $|S_n f(x)|$ is unbounded on a dense G_δ subset of $\mathbb{R}$. This shows Fourier series can fail to converge pointwise to continuous periodic functions in a fairly spectacular way. **Hint:** First show there is a dense G_δ subset of Y, G, such that for all $f \in G$, we have $\sup\{|S_n f(x)| : n \in \mathbb{N}\} = \infty$ for all $x \in \mathbb{Q}$. Of course $\mathbb{Q}$ is not a G_δ set but this is still pretty impressive. Next consider $H_k \equiv \{(x, f) \in \mathbb{R} \times Y : \sup_n |S_n f(x)| > k\}$ and argue that H_k is open and dense. Next let $H_k^1 \equiv \{x \in \mathbb{R} :$ for some $f \in Y, \ (x, f) \in H_k\}$and define H_k^2 similarly. Argue that H_k^i is open and dense and then consider $P_i \equiv \cap_{k=1}^{\infty} H_k^i$.

11. Let $\Lambda_n f = \int_0^{\pi} \sin\left(\left(n + \frac{1}{2}\right) y\right) f(y)\, dy$ for $f \in L^1(0, \pi)$. Show that $\sup\{||\Lambda_n|| : n = 1, 2, \cdots\} < \infty$ using the Riemann Lebesgue lemma.

12. Let X be a normed linear space and let M be a convex open set containing 0. Define

$$\rho(x) = \inf\{t > 0 : \frac{x}{t} \in M\}.$$

Show ρ is a gauge function defined on X. This particular example is called a Minkowski functional. Recall a set, M, is convex if $\lambda x + (1 - \lambda) y \in M$ whenever $\lambda \in [0, 1]$ and $x, y \in M$.

13. ↑ This problem explores the use of the Hahn Banach theorem in establishing separation theorems. Let M be an open convex set containing 0. Let $x \notin M$. Show there exists $x^* \in X'$ such that $\operatorname{Re} x^*(x) \geq 1 > \operatorname{Re} x^*(y)$ for all $y \in M$. **Hint:** If $y \in M, \rho(y) < 1$. Show this. If $x \notin M, \ \rho(x) \geq 1$. Try $f(\alpha x) = \alpha \rho(x)$ for $\alpha \in \mathbb{R}$. Then extend f to F, show F is continuous, then fix it so F is the real part of $x^* \in X'$.

14. A Banach space is said to be strictly convex if whenever $||x|| = ||y||$ and $x \neq y$, then

$$\left|\left| \frac{x + y}{2} \right|\right| < ||x||.$$

$F : X \to X'$ is said to be a duality map if it satisfies the following: a.) $||F(x)|| = ||x||$. b.) $F(x)(x) = ||x||^2$. Show that if X' is strictly convex, then such a duality map exists. **Hint:** Let $f(\alpha x) = \alpha ||x||^2$ and use Hahn Banach theorem, then strict convexity.

15. Suppose $D \subseteq X$, a Banach space, and $L : D \to Y$ is a closed operator. D might not be a Banach space with respect to the norm on X. Define a new norm on D by $||x||_D = ||x||_X + ||Lx||_Y$. Show $(D, || \, ||_D)$ is a Banach space.

16. Prove the following theorem which is an improved version of the open mapping theorem, [8]. Let X and Y be Banach spaces and let $A \in \mathcal{L}(X, Y)$. Then the following are equivalent.

$$AX = Y,$$

$$A \text{ is an open map.}$$

There exists a constant M such that for every $y \in Y$, there exists $x \in X$ with $y = Ax$ and

$$||x|| \leq M \, ||y||.$$

17. Here is an example of a closed unbounded operator. Let $X = Y = C([0, 1])$ and let

$$D = \{f \in C^1([0, 1]) : f(0) = 0\}.$$

$L : D \to C([0, 1])$ is defined by $Lf = f'$. Show L is closed.

18. Suppose $D \subseteq X$ and D is dense in X. Suppose $L : D \to Y$ is linear and $||Lx|| \leq K||x||$ for all $x \in D$. Show there is a unique extension of L, $\widetilde{L}$, defined on all of X with $||\widetilde{L}x|| \leq K||x||$ and $\widetilde{L}$ is linear.

19. $\uparrow$ A Banach space is uniformly convex if whenever $||x_n||, \; ||y_n|| \leq 1$ and $||x_n + y_n|| \to 2$, it follows that $||x_n - y_n|| \to 0$. Show uniform convexity implies strict convexity.

20. We say that x_n converges weakly to x if for every $x^* \in X', \; x^*(x_n) \to x^*(x)$. We write $x_n \rightharpoonup x$ to denote weak convergence. Show that if $||x_n - x|| \to 0$, then $x_n \rightharpoonup x$.

21. $\uparrow$ Show that if X is uniformly convex, then $x_n \rightharpoonup x$ and $||x_n|| \to ||x||$ implies $||x_n - x|| \to 0$. **Hint:** Use Lemma 14.31 to obtain $f \in X'$ with $||f|| = 1$ and $f(x) = ||x||$. See Problem 19 for the definition of uniform convexity.

22. Suppose $L \in \mathcal{L}(X, Y)$ and $M \in \mathcal{L}(Y, Z)$. Show $ML \in \mathcal{L}(X, Z)$ and that $(ML)^* = L^* M^*$.

23. In Theorem 14.27, it was shown that $||L^*|| \leq ||L||$. Are these actually equal? **Hint:** You might show that $\sup_{\beta \in B} \sup_{\alpha \in A} a(\alpha, \beta) = \sup_{\alpha \in A} \sup_{\beta \in B} a(\alpha, \beta)$ and then use this in the string of inequalities used to prove $||L^*|| \leq ||L||$ along with the fact that $||Jx|| = ||x||$ which was established in Theorem 14.30.

Hilbert Spaces

15.1 Basic theory

Let X be a vector space. An inner product is a mapping from $X \times X$ to $\mathbb{C}$ if X is complex and from $X \times X$ to $\mathbb{R}$ if X is real, denoted by (x, y) which satisfies the following.

$$(x, x) \geq 0, \ (x, x) = 0 \text{ if and only if } x = 0, \tag{15.1}$$

$$(x, y) = \overline{(y, x)}. \tag{15.2}$$

For $a, b \in \mathbb{C}$ and $x, y, z \in X$,

$$(ax + by, z) = a(x, z) + b(y, z). \tag{15.3}$$

Note that (15.2) and (15.3) imply $(x, ay + bz) = \overline{a}(x, y) + \overline{b}(x, z)$.

We will show that if $(\cdot, \cdot)$ is an inner product, then $(x, x)^{1/2}$ defines a norm and we say that a normed linear space is an inner product space if $||x|| = (x, x)^{1/2}$.

Definition 15.1 *A normed linear space in which the norm comes from an inner product as just described is called an inner product space. A Hilbert space is a complete inner product space.*

Thus a Hilbert space is a Banach space whose norm comes from an inner product as just described. The difference between what we are doing here and the earlier references to Hilbert space is that here we will be making no assumption that the Hilbert space is finite dimensional. Thus we include the finite dimensional material as a special case of that which is presented here.

Example 15.2 *Let $X = \mathbb{C}^n$ with the inner product given by $(\mathbf{x}, \mathbf{y}) \equiv \sum_{i=1}^n x_i \overline{y}_i$. This is a complex Hilbert space.*

Example 15.3 *Let $X = \mathbb{R}^n$, $(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$. This is a real Hilbert space.*

Theorem 15.4 *(Cauchy Schwarz) In any inner product space*

$$|(x, y)| \leq ||x|| \, ||y||.$$

Proof: Let $\omega \in \mathbb{C}$, $|\omega| = 1$, and $\overline{\omega}(x, y) = |(x, y)| = \text{Re}(x, y\omega)$. Let

$$F(t) = (x + ty\omega, x + t\omega y).$$

If $y = 0$ there is nothing to prove because

$$(x, 0) = (x, 0 + 0) = (x, 0) + (x, 0)$$

and so $(x,0)=0$. Thus, we may assume $y \neq 0$. Then from the axioms of the inner product, (15.1),

$$F(t) = ||x||^2 + 2t \operatorname{Re}(x, \omega y) + t^2 ||y||^2 \geq 0.$$

This yields

$$||x||^2 + 2t|(x,y)| + t^2||y||^2 \geq 0.$$

Since this inequality holds for all $t \in \mathbb{R}$, it follows from the quadratic formula that

$$4|(x,y)|^2 - 4||x||^2||y||^2 \leq 0.$$

This yields the conclusion and proves the theorem.

Earlier it was claimed that the inner product defines a norm. In this next proposition this claim is proved.

Proposition 15.5 *For an inner product space, $||x|| \equiv (x,x)^{1/2}$ does specify a norm.*

Proof: All the axioms are obvious except the triangle inequality. To verify this,

$$\begin{aligned} ||x+y||^2 &\equiv (x+y, x+y) \equiv ||x||^2 + ||y||^2 + 2\operatorname{Re}(x,y) \\ &\leq ||x||^2 + ||y||^2 + 2|(x,y)| \\ &\leq ||x||^2 + ||y||^2 + 2||x||\,||y|| = (||x|| + ||y||)^2. \end{aligned}$$

The following lemma is called the parallelogram identity.

Lemma 15.6 *In an inner product space,*

$$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2.$$

The proof, a straightforward application of the inner product axioms, is left to the reader. See Problem 7. Also note that

$$||x|| = \sup_{||y|| \leq 1} |(x,y)| \tag{15.4}$$

because by the Cauchy Schwarz inequality, if $x \neq 0$,

$$||x|| \geq \sup_{||y|| \leq 1} |(x,y)| \geq \left(x, \frac{x}{||x||}\right) = ||x||.$$

It is obvious that (15.4) holds in the case that $x = 0$.

One of the best things about Hilbert space is the theorem about projection onto a closed convex set. Recall that a set, K, is convex if whenever $\lambda \in [0,1]$ and $x, y \in K$, $\lambda x + (1-\lambda) y \in K$.

Theorem 15.7 *Let K be a closed convex nonempty subset of a Hilbert space, H, and let $x \in H$. Then there exists a unique point $Px \in K$ such that $||Px - x|| \leq ||y - x||$ for all $y \in K$.*

Proof: First we show uniqueness. Suppose $||z_i - x|| \leq ||y - x||$ for all $y \in K$. Then using the parallelogram identity and

$$||z_1 - x|| \leq ||y - x||$$

for all $y \in K$,

$$\begin{aligned} ||z_1 - x||^2 &\leq ||\frac{z_1 + z_2}{2} - x||^2 = ||\frac{z_1 - x}{2} + \frac{z_2 - x}{2}||^2 \\ &= 2(||\frac{z_1 - x}{2}||^2 + ||\frac{z_2 - x}{2}||^2) - ||\frac{z_1 - z_2}{2}||^2 \\ &\leq ||z_1 - x||^2 - ||\frac{z_1 - z_2}{2}||^2, \end{aligned}$$

where the last inequality holds because

$$||z_2 - x|| \leq ||z_1 - x||.$$

Hence $z_1 = z_2$ and this shows uniqueness.

Now let $\lambda = \inf\{||x-y|| : y \in K\}$ and let y_n be a minimizing sequence. Thus $\lim_{n\to\infty} ||x-y_n|| = \lambda,\ y_n \in K$. Then by the parallelogram identity,

$$\begin{aligned} ||y_n - y_m||^2 &= 2(||y_n - x||^2 + ||y_m - x||^2) - 4(||\frac{y_n + y_m}{2} - x||^2) \\ &\leq 2(||y_n - x||^2 + ||y_m - x||^2) - 4\lambda^2. \end{aligned}$$

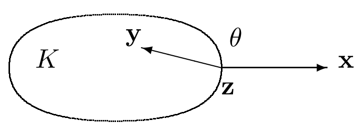
Since $||x - y_n|| \to \lambda$, this shows $\{y_n\}$ is a Cauchy sequence. Since H is complete, $y_n \to y$ for some $y \in H$ which must be in K because K is closed. Therefore $||x - y|| = \lambda$ and we let $Px = y$.

Corollary 15.8 *Let K be a closed, convex, nonempty subset of a Hilbert space, H, and let $x \notin K$. Then for $z \in K,\ z = Px$ if and only if*

$$\text{Re}(x - z, y - z) \leq 0 \tag{15.5}$$

for all $y \in K$.

Before proving this, consider what it says in the case where the Hilbert space is $\mathbb{R}^n$.



Condition (15.5) says the angle, θ, shown in the diagram is always obtuse. Remember, the sign of $\mathbf{x} \cdot \mathbf{y}$ is the same as the sign of the cosine of their included angle.

The inequality (15.5) is an example of a variational inequality and this corollary characterizes the projection of x onto K as the solution of this variational inequality.

Proof of Corollary: Let $z \in K$. Since K is convex, every point of K is in the form $z + t(y - z)$ where $t \in [0,1]$ and $y \in K$. Therefore, $z = Px$ if and only if for all $y \in K$ and $t \in [0,1]$,

$$||x - (z + t(y - z))||^2 = ||(x - z) - t(y - z)||^2 \geq ||x - z||^2$$

for all $t \in [0,1]$ if and only if for all $t \in [0,1]$ and $y \in K$

$$||x - z||^2 + t^2 ||y - z||^2 - 2t \,\text{Re}\,(x - z, y - z) \geq ||x - z||^2$$

which is equivalent to (15.5). This proves the corollary.

Definition 15.9 *Let H be a vector space and let U and V be subspaces. We write $U \oplus V = H$ if every element of H can be written as a sum of an element of U and an element of V in a unique way.*

The case where the closed convex set is a closed subspace is of special importance and in this case the above corollary implies the following.

Corollary 15.10 *Let K be a closed subspace of a Hilbert space, H, and let $x \notin K$. Then for $z \in K,\ z = Px$ if and only if*

$$(x - z, y) = 0$$

for all $y \in K$. Furthermore, $H = K \oplus K^{\perp}$ where

$$K^{\perp} \equiv \{x \in H : (x, k) = 0 \text{ for all } k \in K\}$$

Proof: Since K is a subspace, the condition (15.5) implies

$$\mathrm{Re}(x - z, y) \leq 0$$

for all $y \in K$. But this implies this inequality holds with $\leq$ replaced with $=$. To see this, replace y with $-y$. Now let $|\alpha| = 1$ and

$$\alpha (x - z, y) = |(x - z, y)|.$$

Since $\overline{\alpha} y \in K$ for all $y \in K$,

$$0 = \mathrm{Re}(x - z, \overline{\alpha} y) = (x - z, \overline{\alpha} y) = \alpha (x - z, y) = |(x - z, y)|.$$

Now let $x \in H$. Then $x = x - Px + Px$ and from what was just shown, $x - Px \in K^{\perp}$ and $Px \in K$. This shows that $K^{\perp} + K = H$. We need to verify that $K \cap K^{\perp} = \{0\}$ because this implies that there is at most one way to write an element of H as a sum of one from K and one from $K^{\perp}$. Suppose then that $z \in K \cap K^{\perp}$. Then from what was just shown, $(z, z) = 0$ and so $z = 0$. This proves the corollary.

The following theorem is called the Riesz representation theorem for the dual of a Hilbert space. If $z \in H$ then we may define an element $f \in H'$ by the rule

$$(x, z) \equiv f(x).$$

It follows from the Cauchy Schwartz inequality and the properties of the inner product that $f \in H'$. The Riesz representation theorem says that all elements of H' are of this form.

Theorem 15.11 *Let H be a Hilbert space and let $f \in H'$. Then there exists a unique $z \in H$ such that $f(x) = (x, z)$ for all $x \in H$.*

Proof: If $f = 0$, there is nothing to prove so assume without loss of generality that $f \neq 0$. Let $M = \{x \in H : f(x) = 0\}$. Thus M is a closed proper subspace of H. Let $y \notin M$. Then $y - Py \equiv w$ has the property that $(x, w) = 0$ for all $x \in M$ by Corollary 15.10. Let $x \in H$ be arbitrary. Then

$$xf(w) - f(x)w \in M$$

so

$$0 = (f(w)x - f(x)w, w) = f(w)(x, w) - f(x)||w||^2.$$

Thus

$$f(x) = (x, \frac{\overline{f(w)}w}{||w||^2})$$

and so we let

$$z = \frac{\overline{f(w)}w}{||w||^2}.$$

This proves the existence of z. If $f(x) = (x, z_i)$ $i = 1, 2$, for all $x \in H$, then for all $x \in H$,

$$(x, z_1 - z_2) = 0.$$

Let $x = z_1 - z_2$ to conclude uniqueness. This proves the theorem.

15.2 Orthonormal sets

The concept of an orthonormal set of vectors is a generalization of the notion of the standard basis vectors of $\mathbb{R}^n$ or $\mathbb{C}^n$.

Definition 15.12 *Let H be a Hilbert space. $S \subseteq H$ is called an orthonormal set if $||x|| = 1$ for all $x \in S$ and $(x, y) = 0$ if $x, y \in S$ and $x \neq y$. For any set, D, we define $D^{\perp} \equiv \{x \in H : (x, d) = 0 \text{ for all } d \in D\}$. If S is a set, we denote by span (S) the set of all finite linear combinations of vectors from S.*

We leave it as an exercise to verify that $D^{\perp}$ is always a closed subspace of H.

Theorem 15.13 *In any separable Hilbert space, H, there exists a countable orthonormal set, $S = \{x_i\}$ such that the span of these vectors is dense in H. Furthermore, if $x \in H$, then*

$$x = \sum_{i=1}^{\infty} (x, x_i)\, x_i \equiv \lim_{n \to \infty} \sum_{i=1}^{n} (x, x_i)\, x_i. \tag{15.6}$$

Proof: Let $\mathcal{F}$ denote the collection of all orthonormal subsets of H. We note $\mathcal{F}$ is nonempty because $\{x\} \in \mathcal{F}$ where $||x|| = 1$. The set, $\mathcal{F}$ is a partially ordered set if we let the order be given by set inclusion. By the Hausdorff maximal theorem, there exists a maximal chain, $\mathfrak{C}$ in $\mathcal{F}$. Then we let $S \equiv \cup \mathfrak{C}$. It follows S must be a maximal orthonormal set of vectors. It remains to verify that S is countable *span* (S) is dense, and the condition, (15.6) holds. To see S is countable note that if $x, y \in S$, then

$$||x - y||^2 = ||x||^2 + ||y||^2 - 2\operatorname{Re}(x, y) = ||x||^2 + ||y||^2 = 2.$$

Therefore, the open sets, $B\left(x, \frac{1}{2}\right)$ for $x \in S$ are disjoint and cover S. Since H is assumed to be separable, there exists a point from a countable dense set in each of these disjoint balls showing there can only be countably many of the balls and that consequently, S is countable as claimed.

It remains to verify (15.6) and that *span* (S) is dense. If *span* (S) is not dense, then $\overline{span\,(S)}$ is a closed proper subspace of H and letting $y \notin \overline{span\,(S)}$ we see that $z \equiv y - Py \in span\,(S)^{\perp}$. But then $S \cup \{z\}$ would be a larger orthonormal set of vectors contradicting the maximality of S.

It remains to verify (15.6). Let $S = \{x_i\}_{i=1}^{\infty}$ and consider the problem of choosing the constants, c_k in such a way as to minimize the expression

$$\left|\left| x - \sum_{k=1}^{n} c_k x_k \right|\right|^2 =$$

$$||x||^2 + \sum_{k=1}^{n} |c_k|^2 - \sum_{k=1}^{n} \overline{c_k}\,(x, x_k) - \sum_{k=1}^{n} c_k \overline{(x, x_k)}.$$

We see this equals

$$||x||^2 + \sum_{k=1}^{n} |c_k - (x, x_k)|^2 - \sum_{k=1}^{n} |(x, x_k)|^2$$

and therefore, this minimum is achieved when $c_k = (x, x_k)$. Now since *span* (S) is dense, there exists n large enough that for some choice of constants, c_k,

$$\left|\left| x - \sum_{k=1}^{n} c_k x_k \right|\right|^2 < \varepsilon.$$

However, from what was just shown,

$$\left\|x-\sum_{i=1}^{n}(x,x_i)\,x_i\right\|^2 \leq \left\|x-\sum_{k=1}^{n}c_k x_k\right\|^2 < \varepsilon$$

showing that $\lim_{n\to\infty}\sum_{i=1}^{n}(x,x_i)\,x_i = x$ as claimed. This proves the theorem.

In the proof of this theorem, we established the following corollary.

Corollary 15.14 *Let S be any orthonormal set of vectors and let $\{x_1,\cdot\cdot\cdot,x_n\}\subseteq S$. Then if $x\in H$*

$$\left\|x-\sum_{k=1}^{n}c_k x_k\right\|^2 \geq \left\|x-\sum_{i=1}^{n}(x,x_i)\,x_i\right\|^2$$

for all choices of constants, c_k. In addition to this, we have Bessel's inequality

$$\|x\|^2 \geq \sum_{k=1}^{n}|(x,x_k)|^2 .$$

If S is countable and $span\,(S)$ is dense, then letting $\{x_i\}_{i=1}^{\infty}=S$, we obtain (15.6).

Definition 15.15 *Let $A\in\mathcal{L}(H,H)$ where H is a Hilbert space. Then $|(Ax,y)|\leq\|A\|\,\|x\|\,\|y\|$ and so the map, $x\to(Ax,y)$ is continuous and linear. By the Riesz representation theorem, there exists a unique element of H, denoted by A^*y such that*

$$(Ax,y)=(x,A^*y)\,.$$

It is clear $y\to A^*y$ is linear and continuous. We call A^* the adjoint of A. We say A is a self adjoint operator if $A=A^*$. Thus for all $x,y\in H$, $(Ax,y)=(x,Ay)$. We say A is a compact operator if whenever $\{x_k\}$ is a bounded sequence, there exists a convergent subsequence of $\{Ax_k\}$.

The big result in this subject is sometimes called the Hilbert Schmidt theorem.

Theorem 15.16 *Let A be a compact self adjoint operator defined on a Hilbert space, H. Then there exists a countable set of eigenvalues, $\{\lambda_i\}$ and an orthonormal set of eigenvectors, u_i satisfying*

$$\lambda_i \text{ is real, } |\lambda_n|\geq|\lambda_{n+1}|\,,\ Au_i=\lambda_i u_i, \tag{15.7}$$

and either

$$\lim_{n\to\infty}\lambda_n=0, \tag{15.8}$$

or for some n,

$$span\,(u_1,\cdot\cdot\cdot,u_n)=H. \tag{15.9}$$

In any case,

$$span\,(\{u_i\}_{i=1}^{\infty}) \text{ is dense in } A\,(H)\,. \tag{15.10}$$

and for all $x\in H$,

$$Ax=\sum_{k=1}^{\infty}\lambda_k\,(x,u_k)\,u_k. \tag{15.11}$$

This sequence of eigenvectors and eigenvalues also satisfies

$$|\lambda_n| = ||A_n||, \tag{15.12}$$

and

$$A_n : H_n \to H_n. \tag{15.13}$$

where $H \equiv H_1$ *and* $H_n \equiv \{u_1, \cdots, u_{n-1}\}^\perp$ *and* A_n *is the restriction of* A *to* H_n.

Proof: If $A = 0$ then we may pick $u \in H$ with $||u|| = 1$ and let $\lambda_1 = 0$. Since $A(H) = 0$ it follows the span of u is dense in $A(H)$ and we have proved the theorem in this case.

Thus, we may assume $A \neq 0$. Let λ_1 be real and $\lambda_1^2 \equiv ||A||^2$. We know from the definition of $||A||$ there exists $x_n, ||x_n|| = 1$, and $||Ax_n|| \to ||A|| = |\lambda_1|$. Now it is clear that A^2 is also a compact self adjoint operator. We consider

$$\left(\left(\lambda_1^2 - A^2\right) x_n, x_n\right) = \lambda_1^2 - ||Ax_n||^2 \to 0.$$

Since A is compact, we may replace $\{x_n\}$ by a subsequence, still denoted by $\{x_n\}$ such that Ax_n converges to some element of H. Thus since $\lambda_1^2 - A^2$ satisfies

$$\left(\left(\lambda_1^2 - A^2\right) y, y\right) \geq 0$$

in addition to being self adjoint, it follows $x, y \to \left(\left(\lambda_1^2 - A^2\right) x, y\right)$ satisfies all the axioms for an inner product except for the one which says that $(z, z) = 0$ only if $z = 0$. Therefore, the Cauchy Schwartz inequality (see Problem 6) may be used to write

$$\begin{aligned} \left|\left(\left(\lambda_1^2 - A^2\right) x_n, y\right)\right| &\leq \left(\left(\lambda_1^2 - A^2\right) y, y\right)^{1/2} \left(\left(\lambda_1^2 - A^2\right) x_n, x_n\right)^{1/2} \\ &\leq e_n ||y||. \end{aligned}$$

where $e_n \to 0$ as $n \to \infty$. Therefore, taking the sup over all $||y|| \leq 1$, we see

$$\lim_{n\to\infty} \left|\left|\left(\lambda_1^2 - A^2\right) x_n\right|\right| = 0.$$

Since $A^2 x_n$ converges, it follows since $\lambda_1 \neq 0$ that $\{x_n\}$ is a Cauchy sequence converging to x with $||x|| = 1$. Therefore, $A^2 x_n \to A^2 x$ and so

$$\left|\left|\left(\lambda_1^2 - A^2\right) x\right|\right| = 0.$$

Now

$$(\lambda_1 I - A)(\lambda_1 I + A) x = (\lambda_1 I + A)(\lambda_1 I - A) x = 0.$$

If $(\lambda_1 I - A) x = 0$, we let $u_1 \equiv \frac{x}{||x||}$. If $(\lambda_1 I - A) x = y \neq 0$, we let $u_1 \equiv \frac{y}{||y||}$.

Suppose we have found $\{u_1, \cdots, u_n\}$ such that $Au_k = \lambda_k u_k$ and $|\lambda_k| \geq |\lambda_{k+1}|$, $|\lambda_k| = ||A_k||$ and $A_k : H_k \to H_k$ for $k \leq n$. If

$$span(u_1, \cdots, u_n) = H$$

we have obtained the conclusion of the theorem and we are in the situation of (15.9). Therefore, we assume the span of these vectors is always a proper subspace of H. We show that $A_{n+1} : H_{n+1} \to H_{n+1}$. Let

$$y \in H_{n+1} \equiv \{u_1, \cdots, u_n\}^\perp$$

Then for $k \leq n$

$$(Ay, u_k) = (y, Au_k) = \lambda_k (y, u_k) = 0,$$

showing $A_{n+1} : H_{n+1} \to H_{n+1}$ as claimed. We have two cases to consider. Either $\lambda_n = 0$ or it is not. In the case where $\lambda_n = 0$ we see $A_n = 0$. Then every element of H is the sum of one in $span(u_1, \cdots, u_n)$ and one in $span(u_1, \cdots, u_n)^{\perp}$. (note $span(u_1, \cdots, u_n)$ is a closed subspace. See Problem 11.) Thus, if $x \in H$, we can write $x = y + z$ where $y \in span(u_1, \cdots, u_n)$ and $z \in span(u_1, \cdots, u_n)^{\perp}$ and $Az = 0$. Therefore, $y = \sum_{j=1}^{n} c_j u_j$ and so

$$\begin{aligned} Ax &= Ay = \sum_{j=1}^{n} c_j A u_j \\ &= \sum_{j=1}^{n} c_j \lambda_j u_j \in span(u_1, \cdots, u_n). \end{aligned}$$

It follows that we have the conclusion of the theorem in this case because the above equation holds if we let $c_i = (x, u_i)$.

Now consider the case where $\lambda_n \neq 0$. In this case we repeat the above argument used to find u_1 and λ_1 for the operator, A_{n+1}. This yields $u_{n+1} \in H_{n+1} \equiv \{u_1, \cdots, u_n\}^{\perp}$ such that

$$||u_{n+1}|| = 1, ||Au_{n+1}|| = |\lambda_{n+1}| = ||A_{n+1}|| \leq ||A_n|| = |\lambda_n|$$

and if it is ever the case that $\lambda_n = 0$, it follows from the above argument that the conclusion of the theorem is obtained.

Now we claim $\lim_{n\to\infty} \lambda_n = 0$. If this were not so, we would have $0 < \varepsilon = \lim_{n\to\infty} |\lambda_n|$ but then

$$\begin{aligned} ||Au_n - Au_m||^2 &= ||\lambda_n u_n - \lambda_m u_m||^2 \\ &= |\lambda_n|^2 + |\lambda_m|^2 \geq 2\varepsilon^2 \end{aligned}$$

and so there would not exist a convergent subsequence of $\{Au_k\}_{k=1}^{\infty}$ contrary to the assumption that A is compact. Thus we have verified the claim that $\lim_{n\to\infty} \lambda_n = 0$. It remains to verify that $span(\{u_i\})$ is dense in $A(H)$. If $w \in span(\{u_i\})^{\perp}$ then $w \in H_n$ for all n and so for all n, we have

$$||Aw|| \leq ||A_n|| \, ||w|| \leq |\lambda_n| \, ||w|| .$$

Therefore, $Aw = 0$. Now every vector from H can be written as a sum of one from

$$span(\{u_i\})^{\perp} = \overline{span(\{u_i\})}^{\perp}$$

and one from $\overline{span(\{u_i\})}$. Therefore, if $x \in H$, we can write $x = y + w$ where $y \in \overline{span(\{u_i\})}$ and $w \in \overline{span(\{u_i\})}^{\perp}$. From what we just showed, we see $Aw = 0$. Also, since $y \in \overline{span(\{u_i\})}$, there exist constants, c_k and n such that

$$\left|\left| y - \sum_{k=1}^{n} c_k u_k \right|\right| < \varepsilon.$$

Therefore, from Corollary 15.14,

$$\left|\left| y - \sum_{k=1}^{n} (y, u_k) u_k \right|\right| = \left|\left| y - \sum_{k=1}^{n} (x, u_k) u_k \right|\right| < \varepsilon.$$

Therefore,

$$\begin{aligned}||A||\,\varepsilon &> \left|\left|A\left(y-\sum_{k=1}^{n}(x,u_k)\,u_k\right)\right|\right| \\ &= \left|\left|Ax-\sum_{k=1}^{n}(x,u_k)\,\lambda_k u_k\right|\right|.\end{aligned}$$

Since ε is arbitrary, this shows $span\left(\{u_i\}\right)$ is dense in $A\left(H\right)$ and also implies (15.11). This proves the theorem.

Note that if we define $v\otimes u\in\mathcal{L}\left(H,H\right)$ by

$$v\otimes u\left(x\right)=\left(x,u\right)v,$$

then we can write (15.11) in the form

$$A=\sum_{k=1}^{\infty}\lambda_k u_k\otimes u_k$$

We give the following useful corollary.

Corollary 15.17 *Let A be a compact self adjoint operator defined on a separable Hilbert space, H. Then there exists a countable set of eigenvalues, $\{\lambda_i\}$ and an orthonormal set of eigenvectors, v_i satisfying*

$$Av_i=\lambda_i v_i,\,||v_i||=1, \tag{15.14}$$

$$span\left(\{v_i\}_{i=1}^{\infty}\right)\ \text{is dense in } H. \tag{15.15}$$

Furthermore, if $\lambda_i\neq 0$, the space, $V_{\lambda_i}\equiv\{x\in H:Ax=\lambda_i x\}$ is finite dimensional.

Proof: In the proof of the above theorem, let $W\equiv\overline{span\left(\{u_i\}\right)}^{\perp}$. By Theorem 15.13, there is an orthonormal set of vectors, $\{w_i\}_{i=1}^{\infty}$ whose span is dense in W. As shown in the proof of the above theorem, $Aw=0$ for all $w\in W$. Let $\{v_i\}_{i=1}^{\infty}=\{u_i\}_{i=1}^{\infty}\cup\{w_i\}_{i=1}^{\infty}$.

It remains to verify the space, V_{λ_i}, is finite dimensional. First we observe that $A:V_{\lambda_i}\to V_{\lambda_i}$. Since A is continuous, it follows that $A:\overline{V_{\lambda_i}}\to\overline{V_{\lambda_i}}$. Thus A is a compact self adjoint operator on $\overline{V_{\lambda_i}}$ and by Theorem 15.16, we must be in the situation of (15.9) because the only eigenvalue is λ_i. This proves the corollary.

Note the last claim of this corollary holds independent of the separability of H.

Suppose $\lambda\notin\{\lambda_n\}$ and $\lambda\neq 0$. Then we can use the above formula for A, (15.11), to give a formula for $\left(A-\lambda I\right)^{-1}$. We note first that since $\lim_{n\to\infty}\lambda_n=0$, it follows that $\lambda_n^2/\left(\lambda_n-\lambda\right)^2$ must be bounded, say by a positive constant, M.

Corollary 15.18 *Let A be a compact self adjoint operator and let $\lambda\notin\{\lambda_n\}_{n=1}^{\infty}$ and $\lambda\neq 0$ where the λ_n are the eigenvalues of A. Then*

$$\left(A-\lambda I\right)^{-1}x=-\frac{1}{\lambda}x+\frac{1}{\lambda}\sum_{k=1}^{\infty}\frac{\lambda_k}{\lambda_k-\lambda}\left(x,u_k\right)u_k. \tag{15.16}$$

Proof: Let $m<n$. Then since the $\{u_k\}$ form an orthonormal set,

$$\begin{aligned}\left|\sum_{k=m}^{n}\frac{\lambda_k}{\lambda_k-\lambda}\left(x,u_k\right)u_k\right| &= \left(\sum_{k=m}^{n}\left(\frac{\lambda_k}{\lambda_k-\lambda}\right)^2\left|\left(x,u_k\right)\right|^2\right)^{1/2} \\ &\leq M\left(\sum_{k=m}^{n}\left|\left(x,u_k\right)\right|^2\right)^{1/2}.\end{aligned} \tag{15.17}$$

But we have from Bessel's inequality,

$$\sum_{k=1}^{\infty} |(x, u_k)|^2 \leq ||x||^2$$

and so for m large enough, the first term in (15.17) is smaller than ε. This shows the infinite series in (15.16) converges. It is now routine to verify that the formula in (15.16) is the inverse.

15.3 The Fredholm alternative

Recall that if A is an $n \times n$ matrix and if the only solution to the system, $A\mathbf{x} = 0$ is $\mathbf{x} = 0$ then for any $\mathbf{y} \in \mathbb{R}^n$ it follows that there exists a unique solution to the system $A\mathbf{x} = \mathbf{y}$. This holds because the first condition implies A is one to one and therefore, A^{-1} exists. Of course things are much harder in a Hilbert space. Here is a simple example.

Example 15.19 *Let $L^2(\mathbb{N}; \mu) = H$ where μ is counting measure. Thus an element of H is a sequence, $\mathbf{a} = \{a_i\}_{i=1}^{\infty}$ having the property that*

$$||\mathbf{a}||_H \equiv \left(\sum_{k=1}^{\infty} |a_k|^2 \right)^{1/2} < \infty.$$

We define $A : H \to H$ by

$$A\mathbf{a} \equiv \mathbf{b} \equiv \{0, a_1, a_2, \cdots\}.$$

Thus A slides the sequence to the right and puts a zero in the first slot. Clearly A is one to one and linear but it cannot be onto because it fails to yield $\mathbf{e}_1 \equiv \{1, 0, 0, \cdots\}$.

Notwithstanding the above example, there are theorems which are like the linear algebra theorem mentioned above which hold in an arbitrary Hilbert space in the case where some operator is compact. To begin with we give a simple lemma which is interesting for its own sake.

Lemma 15.20 *Suppose A is a compact operator defined on a Hilbert space, H. Then $(I - A)(H)$ is a closed subspace of H.*

Proof: Suppose $(I - A) x_n \to y$. Let

$$\alpha_n \equiv dist\,(x_n, \ker (I - A))$$

and let $z_n \in \ker (I - A)$ be such that

$$\alpha_n \leq ||x_n - z_n|| \leq \left(1 + \frac{1}{n}\right) \alpha_n.$$

Thus $(I - A)(x_n - z_n) \to y$.

Case 1: $\{x_n - z_n\}$ has a bounded subsequence.

If this is so, the compactness of A implies there exists a subsequence, still denoted by n such that $\{A(x_n - z_n)\}_{n=1}^{\infty}$ is a Cauchy sequence. Since $(I - A)(x_n - z_n) \to y$, this implies $\{(x_n - z_n)\}$ is also a Cauchy sequence converging to a point, $x \in H$. Then, taking the limit as $n \to \infty$, we see $(I - A) x = y$ and so $y \in (I - A)(H)$.

Case 2: $\lim_{n\to\infty} ||x_n - z_n|| = \infty$. We will show this case cannot occur.

In this case, we let $w_n \equiv \frac{x_n - z_n}{||x_n - z_n||}$. Thus $(I - A) w_n \to 0$ and w_n is bounded. Therefore, we can take a subsequence, still denoted by n such that $\{Aw_n\}$ is a Cauchy sequence. This implies $\{w_n\}$ is a Cauchy sequence which must converge to some $w_\infty \in H$. Therefore, $(I-A) w_\infty = 0$ and so $w_\infty \in \ker(I-A)$. However, this is impossible because of the following argument. If $z \in \ker(I-A)$,

$$\begin{aligned}||w_n - z|| &= \frac{1}{||x_n - z_n||} ||x_n - z_n - ||x_n - z_n|| z|| \\ &\geq \frac{1}{||x_n - z_n||}\alpha_n \geq \frac{\alpha_n}{\left(1+\frac{1}{n}\right)\alpha_n} = \frac{n}{n+1}.\end{aligned}$$

Taking the limit, we see $||w_\infty - z|| \geq 1$. Since $z \in \ker(I-A)$ is arbitrary, this shows $dist(w_\infty, \ker(I-A)) \geq 1$.

Since Case 2 does not occur, this proves the lemma.

Theorem 15.21 *Let A be a compact operator defined on a Hilbert space, H and let $f \in H$. Then there exists a solution, x, to*

$$x - Ax = f \tag{15.18}$$

if and only if

$$(f, y) = 0 \tag{15.19}$$

for all y solving

$$y - A^* y = 0. \tag{15.20}$$

Proof: Suppose x is a solution to (15.18) and let y be a solution to (15.20). Then

$$\begin{aligned}(f,y) &= (x - Ax, y) = (x,y) - (Ax, y) \\ &= (x,y) - (x, A^*y) = (x, y - A^*y) = 0.\end{aligned}$$

Next suppose $(f,y) = 0$ for all y solving (15.20). We want to show there exists x solving (15.18). By Lemma 15.20, $(I-A)(H)$ is a closed subspace of H. Therefore, there exists a point, $(I-A)x$, in this subspace which is the closest point to f. By Corollary 15.10, we must have

$$(f - (I-A)x, (I-A)y - (I-A)x) = 0$$

for all $y \in H$. Therefore,

$$((I-A^*)[f - (I-A)x], y - x) = 0$$

for all $y \in H$. This implies x is a solution to

$$(I-A^*)(I-A)x = (I-A^*)f$$

and so

$$(I-A^*)[(I-A)x - f] = 0.$$

Therefore $(f, f - (I-A)x) = 0$. Since $(I-A)x \in (I-A)(H)$, we also have

$$((I-A)x, f - (I-A)x) = 0$$

and so

$$(f - (I-A)x, f - (I-A)x) = 0,$$

showing that $f = (I-A)x$. This proves the theorem.

The following corollary is called the Fredholm alternative.

Corollary 15.22 *Let A be a compact operator defined on a Hilbert space, H. Then there exists a solution to the equation*

$$x - Ax = f \tag{15.21}$$

for all $f \in H$ if and only if $(I - A^*)$ is one to one.

Proof: Suppose $(I - A)$ is one to one first. Then if $y - A^*y = 0$ it follows $y = 0$ and so for any $f \in H$, $(f, y) = (f, 0) = 0$. Therefore, by Theorem 15.21 there exists a solution to $(I - A)\,x = f$.

Now suppose there exists a solution, x, to $(I - A)\,x = f$ for every $f \in H$. If $(I - A^*)\,y = 0$, we can let $(I - A)\,x = y$ and so

$$||y||^2 = (y, y) = ((I - A)\,x, y) = (x, (I - A^*)\,y) = 0.$$

Therefore, $(I - A^*)$ is one to one.

15.4 Sturm Liouville problems

A Sturm Liouville problem involves the differential equation,

$$(p\,(x)\,y')' + (\lambda q\,(x) + r\,(x))\,y = 0,\ x \in [a, b] \tag{15.22}$$

where we assume that $q\,(t) \neq 0$ for any t and some boundary conditions,

$$\begin{array}{l} \text{boundary condition at } a \\ \text{boundary condition at } b \end{array} \tag{15.23}$$

For example we typically have boundary conditions of the form

$$\begin{aligned} C_1 y\,(a) + C_2 y'\,(a) &= 0 \\ C_3 y\,(b) + C_4 y'\,(b) &= 0 \end{aligned} \tag{15.24}$$

where

$$C_1^2 + C_2^2 > 0,\ \text{and } C_3^2 + C_4^2 > 0. \tag{15.25}$$

We will assume all the functions involved are continuous although this could certainly be weakened. Also we assume here that a and b are finite numbers. In the example the constants, C_i are given and λ is called the eigenvalue while a solution of the differential equation and given boundary conditions corresponding to λ is called an eigenfunction.

There is a simple but important identity related to solutions of the above differential equation. Suppose λ_i and y_i for $i = 1, 2$ are two solutions of (15.22). Thus from the equation, we obtain the following two equations.

$$(p\,(x)\,y_1')'\,y_2 + (\lambda_1 q\,(x) + r\,(x))\,y_1 y_2 = 0,$$

$$(p\,(x)\,y_2')'\,y_1 + (\lambda_2 q\,(x) + r\,(x))\,y_1 y_2 = 0.$$

Subtracting the second from the first yields

$$(p\,(x)\,y_1')'\,y_2 - (p\,(x)\,y_2')'\,y_1 + (\lambda_1 - \lambda_2)\,q\,(x)\,y_1 y_2 = 0. \tag{15.26}$$

Now we note that

$$\left(p\left(x\right)y_1'\right)'y_2-\left(p\left(x\right)y_2'\right)'y_1=\frac{d}{dx}\left(\left(p\left(x\right)y_1'\right)y_2-\left(p\left(x\right)y_2'\right)y_1\right)$$

and so integrating (15.26) from a to b, we obtain

$$\left(\left(p\left(x\right)y_1'\right)y_2-\left(p\left(x\right)y_2'\right)y_1\right)|_a^b+\left(\lambda_1-\lambda_2\right)\int_a^b q\left(x\right)y_1\left(x\right)y_2\left(x\right)dx=0 \tag{15.27}$$

We have been purposely vague about the nature of the boundary conditions because of a desire to not lose generality. However, we will always assume the boundary conditions are such that whenever y_1 and y_2 are two eigenfunctions, it follows that

$$\left(\left(p\left(x\right)y_1'\right)y_2-\left(p\left(x\right)y_2'\right)y_1\right)|_a^b=0 \tag{15.28}$$

In the case where the boundary conditions are given by (15.24), and (15.25), we obtain (15.28). To see why this is so, consider the top limit. This yields

$$p\left(b\right)\left[y_1'\left(b\right)y_2\left(b\right)-y_2'\left(b\right)y_1\left(b\right)\right]$$

However we know from the boundary conditions that

$$\begin{aligned}C_3y_1\left(b\right)+C_4y_1'\left(b\right)&=0\\C_3y_2\left(b\right)+C_4y_2'\left(b\right)&=0\end{aligned}$$

and that from (15.25) that not both C_3 and C_4 equal zero. Therefore the determinant of the matrix of coefficients must equal zero. But this is implies $\left[y_1'\left(b\right)y_2\left(b\right)-y_2'\left(b\right)y_1\left(b\right)\right]=0$ which yields the top limit is equal to zero. A similar argument holds for the lower limit.

With the identity (15.27) we can give a result on orthogonality of the eigenfunctions.

Proposition 15.23 *Suppose y_i solves the problem (15.22) , (15.23), and (15.28) for $\lambda=\lambda_i$ where $\lambda_1\neq\lambda_2$. Then we have the orthogonality relation*

$$\int_a^b q\left(x\right)y_1\left(x\right)y_2\left(x\right)dx=0. \tag{15.29}$$

In addition to this, if u,v are two solutions to the differential equation corresponding to λ, (15.22), not necessarily the boundary conditions, then there exists a constant, C such that

$$W\left(u,v\right)\left(x\right)p\left(x\right)=C \tag{15.30}$$

for all $x\in[a,b]$. In this formula, $W\left(u,v\right)$ denotes the Wronskian given by

$$\det\begin{pmatrix}u\left(x\right)&v\left(x\right)\\u'\left(x\right)&v'\left(x\right)\end{pmatrix}. \tag{15.31}$$

Proof: The orthogonality relation, (15.29) follows from the fundamental assumption, (15.28) and (15.27). It remains to verify (15.30). We have

$$\left(p\left(x\right)u'\right)'v-\left(p\left(x\right)v'\right)'u+\left(\lambda-\lambda\right)q\left(x\right)uv=0.$$

Now the first term equals

$$\frac{d}{dx}\left(p\left(x\right)u'v-p\left(x\right)v'u\right)=\frac{d}{dx}\left(p\left(x\right)W\left(v,u\right)\left(x\right)\right)$$

and so $p\left(x\right)W\left(u,v\right)\left(x\right)=-p\left(x\right)W\left(v,u\right)\left(x\right)=C$ as claimed.

Now consider the differential equation,

$$\left(p\left(x\right)y'\right)'+r\left(x\right)y=0. \tag{15.32}$$

This is obtained from the one of interest by letting $\lambda=0$.

Criterion 15.24 *Suppose we are able to find functions, u and v such that they solve the differential equation, (15.32) and u solves the boundary condition at $x = a$ while v solves the boundary condition at $x = b$. Suppose also that $W(u, v) \neq 0$.*

If $p(x) > 0$ on $[a, b]$ it is clear from the fundamental existence and uniqueness theorems for ordinary differential equations that such functions, u and v exist. (See any good differential equations book or Problem 10 of Chapter 10.) The following lemma gives an easy to check sufficient condition for Criterion 15.24 to occur in the case where the boundary conditions are given in (15.24), (15.25).

Lemma 15.25 *Suppose $p(x) \neq 0$ for all $x \in [a, b]$. Then for C_1 and C_2 given in (15.24) and u a nonzero solution of (15.32), if*

$$C_3 u(b) + C_4 u'(b) \neq 0,$$

Then if v is any nonzero solution of the equation (15.32) which satisfies the boundary condition at $x = b$, it follows $W(u, v) \neq 0$.

Proof: Thanks to Proposition 15.23 $W(u, v)(x)$ is either equal to zero for all $x \in [a, b]$ or it is never equal to zero on this interval. If the conclusion of the lemma is not so, then $\frac{u}{v}$ equals a constant. This is easy to see from using the quotient rule in which the Wronskian occurs in the numerator. Therefore, $v(x) = u(x)c$ for some nonzero constant, c But then

$$C_3 v(b) + C_4 v'(b) = c(C_3 u(b) + C_4 u'(b)) \neq 0,$$

contrary to the assumption that v satisfies the boundary condition at $x = b$. This proves the lemma.

Lemma 15.26 *Assume Criterion 15.24. In the case where the boundary conditions are given by (15.24) and (15.25), a function, y is a solution to the boundary conditions, (15.24) and (15.25) along with the equation,*

$$(p(x) y')' + r(x) y = g \tag{15.33}$$

if and only if

$$y(x) = \int_a^b G(t, x) g(t) \, dt \tag{15.34}$$

where

$$G_1(t, x) = \begin{cases} c^{-1}(p(t) v(x) u(t)) & \text{if } t < x \\ c^{-1}(p(t) v(t) u(x)) & \text{if } t > x \end{cases}. \tag{15.35}$$

where c is the constant of Proposition 15.23 which satisfies $p(x) W(u, v)(x) = c$.

Proof: We can verify that if

$$y_p(x) = \frac{1}{c} \int_a^x g(t) p(t) u(t) v(x) \, dt + \frac{1}{c} \int_x^b g(t) p(t) v(t) u(x) \, dt,$$

then y_p is a particular solution of the equation (15.33) which satisfies the boundary conditions, (15.24) and (15.25). Therefore, every solution of the equation must be of the form

$$y(x) = \alpha u(x) + \beta v(x) + y_p(x).$$

Substituting in to the given boundary conditions, (15.24), we obtain

$$\beta(C_1 v(a) + C_2 v'(a)) = 0.$$

If $\beta \neq 0$, then we have

$$\begin{aligned} C_1 v(a) + C_2 v'(a) &= 0 \\ C_1 u(a) + C_2 u'(a) &= 0. \end{aligned}$$

Since $C_1^2 + C_2^2 \neq 0$, we must have

$$(v(a) u'(a) - u(a) v'(a)) = W(v, u)(a) = 0$$

which contradicts the assumption in Criterion 15.24 about the Wronskian. Thus $\beta = 0$. Similar reasoning applies to show $\alpha = 0$ also. This proves the lemma.

Now in the case of Criterion 15.24, y is a solution to the Sturm Liouville eigenvalue problem, (15.22), (15.24), and (15.25) if and only if y solves the boundary conditions, (15.24) and the equation,

$$(p(x) y')' + r(x) y(x) = -\lambda q(x) y(x).$$

This happens if and only if

$$y(x) = \frac{-\lambda}{c} \int_a^x q(t) y(t) p(t) u(t) v(x) \, dt + \frac{-\lambda}{c} \int_x^b q(t) y(t) p(t) v(t) u(x) \, dt, \tag{15.36}$$

Letting $\mu = \frac{1}{\lambda}$, this if of the form

$$\mu y(x) = \int_a^b G(t, x) y(t) \, dt \tag{15.37}$$

where

$$G(t, x) = \begin{cases} -c^{-1} (q(t) p(t) v(x) u(t)) & \text{if } t < x \\ -c^{-1} (q(t) p(t) v(t) u(x)) & \text{if } t > x \end{cases}. \tag{15.38}$$

Could $\mu = 0$? If this happened, then from Lemma 15.26, we would have that 0 is the solution of (15.33) where the right side is $-q(t) y(t)$ which would imply that $q(t) y(t) = 0$ for all t which implies $y(t) = 0$ for all t. It follows from (15.38) that $G : [a, b] \times [a, b] \to \mathbb{R}$ is continuous and symmetric, $G(t, x) = G(x, t)$.

Also we see that for $f \in C([a, b])$, and

$$w(x) \equiv \int_a^b G(t, x) f(t) \, dt,$$

Lemma 15.26 implies w is the solution to the boundary conditions (15.24), (15.25) and the equation

$$(p(x) y')' + r(x) y = -q(x) f(x) \tag{15.39}$$

Theorem 15.27 *Suppose u, v are given in Criterion 15.24. Then there exists a sequence of functions, $\{y_n\}_{n=1}^{\infty}$ and real numbers, λ_n such that*

$$(p(x) y_n')' + (\lambda_n q(x) + r(x)) y_n = 0, \; x \in [a, b], \tag{15.40}$$

$$\begin{aligned} C_1 y_n(a) + C_2 y_n'(a) &= 0, \\ C_3 y_n(b) + C_4 y_n'(b) &= 0. \end{aligned} \tag{15.41}$$

and

$$\lim_{n \to \infty} |\lambda_n| = \infty \tag{15.42}$$

such that for all $f \in C([a,b])$, whenever w satisfies (15.39) and the boundary conditions, (15.24),

$$w(x) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} (f, y_n) y_n. \tag{15.43}$$

Also the functions, $\{y_n\}$ form a dense set in $L^2(a,b)$ which satisfy the orthogonality condition, (15.29).

Proof: Let $Ay(x) \equiv \int_a^b G(t,x) y(t)\, dt$ where G is defined above in (15.35). Then $A : L^2(a,b) \to C([a,b]) \subseteq L^2(a,b)$. Also, for $y \in L^2(a,b)$ we may use Fubini's theorem and obtain

$$\begin{aligned}(Ay, z)_{L^2(a,b)} &= \int_a^b \int_a^b G(t,x) y(t) z(x)\, dtdx \\ &= \int_a^b \int_a^b G(t,x) y(t) z(x)\, dxdt \\ &= (y, Az)_{L^2(a,b)}\end{aligned}$$

showing that A is self adjoint.

Now suppose $D \subseteq L^2(a,b)$ is a bounded set and pick $y \in D$. Then

$$\begin{aligned}|Ay(x)| &\equiv \left| \int_a^b G(t,x) y(t)\, dt \right| \\ &\leq \int_a^b |G(t,x)| |y(t)|\, dt \\ &\leq C_G ||y||_{L^2(a,b)} \leq C\end{aligned}$$

where C is a constant which depends on G and the uniform bound on functions from D. Therefore, the functions, $\{Ay : y \in D\}$ are uniformly bounded. Now for $y \in D$, we use the uniform continuity of G on $[a,b] \times [a,b]$ to conclude that when $|x - z|$ is sufficiently small, $|G(t,x) - G(t,z)| < \varepsilon$ and that therefore,

$$\begin{aligned}|Ay(x) - Ay(z)| &= \left| \int_a^b (G(t,x) - G(t,z)) y(t)\, dt \right| \\ &\leq \int_a^b \varepsilon |y(t)| \leq \varepsilon \sqrt{b-a}\, ||y||_{L^2(a,b)}\end{aligned}$$

Thus $\{Ay : y \in D\}$ is uniformly equicontinuous and so by the Ascoli Arzela theorem, Theorem 4.4, this set of functions is precompact in $C([a,b])$ which means there exists a uniformly convergent subsequence, $\{Ay_n\}$. However this sequence must then be uniformly Cauchy in the norm of the space, $L^2(a,b)$ and so A is a compact self adjoint operator defined on the Hilbert space, $L^2(a,b)$. Therefore, by Theorem 15.16, there exist functions y_n and real constants, μ_n such that $||y_n||_{L^2} = 1$ and $Ay_n = \mu_n y_n$ and

$$|\mu_n| \geq |\mu_{n+1}|,\ Au_i = \mu_i u_i, \tag{15.44}$$

and either

$$\lim_{n\to\infty} \mu_n = 0, \tag{15.45}$$

or for some n,

$$span(y_1, \cdots, y_n) = H \equiv L^2(a,b). \tag{15.46}$$

Since (15.46) does not occur, we must have (15.45). Also from Theorem 15.16,

$$span\left(\{y_i\}_{i=1}^{\infty}\right) \text{ is dense in } A\left(H\right). \tag{15.47}$$

and so for all $f \in C\left([a,b]\right)$,

$$Af = \sum_{k=1}^{\infty} \mu_k \left(f, y_k\right) y_k. \tag{15.48}$$

Thus for w a solution of (15.39) and the boundary conditions (15.24),

$$w = Af = \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left(f, y_k\right) y_k.$$

The last claim follows from Corollary 15.17 and the observation above that μ is never equal to zero. This proves the theorem.

Note that if since $q\left(x\right) \neq 0$ we can say that for a given $g \in C\left([a,b]\right)$ we can define f by $g\left(x\right) = -q\left(x\right) f\left(x\right)$ and so if w is a solution to the boundary conditions, (15.24) and the equation

$$\left(p\left(x\right) w'\left(x\right)\right)' + r\left(x\right) w\left(x\right) = g\left(x\right) = -q\left(x\right) f\left(x\right),$$

we may write the formula

$$\begin{aligned} w\left(x\right) &= \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left(f, y_k\right) y_k \\ &= \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left(\frac{-g}{q}, y_k\right) y_k. \end{aligned}$$

Theorem 15.28 *Suppose $f \in L^2\left(a,b\right)$. Then*

$$f = \sum_{k=1}^{\infty} a_k y_k \tag{15.49}$$

where

$$a_k = \frac{\int_a^b f\left(x\right) y_k\left(x\right) q\left(x\right) dx}{\int_a^b y_k^2\left(x\right) q\left(x\right) dx} \tag{15.50}$$

and the convergence of the partial sums takes place in $L^2\left(a,b\right)$.

Proof: By Theorem 15.27 there exist b_k and n such that

$$\left|\left|f - \sum_{k=1}^{n} b_k y_k\right|\right|_{L^2} < \varepsilon.$$

Now we can define an equivalent norm on $L^2\left(a,b\right)$ by

$$|||f|||_{L^2(a,b)} \equiv \left(\int_a^b \left|f\left(x\right)\right|^2 \left|q\left(x\right)\right| dx\right)^{1/2}$$

Thus there exist constants δ and Δ independent of g such that

$$\delta \|g\| \leq |||g||| \leq \Delta \|g\|$$

Therefore,

$$\left|\left|\left| f - \sum_{k=1}^{n} b_k y_k \right|\right|\right|_{L^2} < \Delta\varepsilon.$$

This new norm also comes from the inner product $((f,g)) \equiv \int_a^b f(x) \overline{g(x)} |q(x)| \, dx$. Then as in Theorem 9.21 a completing the square argument shows if we let b_k be given as in (15.50) then the distance between f and the linear combination of the first n of the y_k measured in the norm, $|||\cdot|||$, is minimized Thus letting a_k be given by (15.50), we see that

$$\delta \left|\left| f - \sum_{k=1}^{n} a_k y_k \right|\right|_{L^2} \leq \left|\left|\left| f - \sum_{k=1}^{n} a_k y_k \right|\right|\right|_{L^2} \leq \left|\left|\left| f - \sum_{k=1}^{n} b_k y_k \right|\right|\right|_{L^2} < \Delta\varepsilon$$

Since ε is arbitrary, this proves the theorem.

More can be said about convergence of these series based on the eigenfunctions of a Sturm Liouville problem. In particular, it can be shown that for reasonable functions the pointwise convergence properties are like those of Fourier series and that the series converges to the midpoint of the jump. For more on these topics see the old book by Ince, written in Egypt in the 1920's, [17].

15.5 Exercises

1. Prove Examples 2.14 and 2.15 are Hilbert spaces. For $f, g \in C([0,1])$ let $(f,g) = \int_0^1 f(x) \overline{g(x)} dx$. Show this is an inner product. What does the Cauchy Schwarz inequality say in this context?

2. Generalize the Fredholm theory presented above to the case where $A : X \to Y$ for X, Y Banach spaces. In this context, $A^* : Y' \to X'$ is given by $A^* y^*(x) \equiv y^*(Ax)$. We say A is compact if $A(\text{bounded set}) = \text{precompact set}$, exactly as in the Hilbert space case.

3. Let S denote the unit sphere in a Banach space, X, $S \equiv \{x \in X : \|x\| = 1\}$. Show that if Y is a Banach space, then $A \in \mathcal{L}(X,Y)$ is compact if and only if $A(S)$ is precompact.

4. $\uparrow$ Show that $A \in \mathcal{L}(X,Y)$ is compact if and only if A^* is compact. **Hint:** Use the result of 3 and the Ascoli Arzela theorem to argue that for S^* the unit ball in X', there is a subsequence, $\{y_n^*\} \subseteq S^*$ such that y_n^* converges uniformly on the compact set, $\overline{A(S)}$. Thus $\{A^* y_n^*\}$ is a Cauchy sequence in X'. To get the other implication, apply the result just obtained for the operators A^* and A^{**}. Then use results about the embedding of a Banach space into its double dual space.

5. Prove a version of Problem 4 for Hilbert spaces.

6. Suppose, in the definition of inner product, Condition (15.1) is weakened to read only $(x,x) \geq 0$. Thus the requirement that $(x,x) = 0$ if and only if $x = 0$ has been dropped. Show that then $|(x,y)| \leq |(x,x)|^{1/2} |(y,y)|^{1/2}$. This generalizes the usual Cauchy Schwarz inequality.

7. Prove the parallelogram identity. Next suppose $(X, \| \, \|)$ is a real normed linear space. Show that if the parallelogram identity holds, then $(X, \| \, \|)$ is actually an inner product space. That is, there exists an inner product $(\cdot,\cdot)$ such that $\|x\| = (x,x)^{1/2}$.

8. Let K be a closed convex subset of a Hilbert space, H, and let P be the projection map of the chapter. Thus, $\|Px - x\| \leq \|y - x\|$ for all $y \in K$. Show that $\|Px - Py\| \leq \|x - y\|$.

9. Show that every inner product space is uniformly convex. This means that if x_n, y_n are vectors whose norms are no larger than 1 and if $||x_n + y_n|| \to 2$, then $||x_n - y_n|| \to 0$.

10. Let H be separable and let S be an orthonormal set. Show S is countable.

11. Suppose $\{x_1, \cdots, x_m\}$ is a linearly independent set of vectors in a normed linear space. Show $span\,(x_1, \cdots, x_m)$ is a closed subspace. Also show that if $\{x_1, \cdots, x_m\}$ is an orthonormal set then $span\,(x_1, \cdots, x_m)$ is a closed subspace.

12. Show every Hilbert space, separable or not, has a maximal orthonormal set of vectors.

13. ↑ Prove Bessel's inequality, which says that if $\{x_n\}_{n=1}^{\infty}$ is an orthonormal set in H, then for all $x \in H$, $||x||^2 \geq \sum_{k=1}^{\infty} |(x, x_k)|^2$. **Hint:** Show that if $M = span(x_1, \cdots, x_n)$, then $Px = \sum_{k=1}^{n} x_k(x, x_k)$. Then observe $||x||^2 = ||x - Px||^2 + ||Px||^2$.

14. ↑ Show S is a maximal orthonormal set if and only if

$$span(S) \equiv \{\text{all finite linear combinations of elements of } S\}$$

is dense in H.

15. ↑ Suppose $\{x_n\}_{n=1}^{\infty}$ is a maximal orthonormal set. Show that

$$x = \sum_{n=1}^{\infty} (x, x_n) x_n \equiv \lim_{N \to \infty} \sum_{n=1}^{N} (x, x_n) x_n$$

and $||x||^2 = \sum_{i=1}^{\infty} |(x, x_i)|^2$. Also show $(x, y) = \sum_{n=1}^{\infty} (x, x_n)\overline{(y, x_n)}$.

16. Let $S = \{e^{inx}(2\pi)^{-\frac{1}{2}}\}_{n=-\infty}^{\infty}$. Show S is an orthonormal set if the inner product is given by $(f, g) = \int_{-\pi}^{\pi} f\,(x)\,\overline{g\,(x)}dx$.

17. ↑ Show that if Bessel's equation,

$$||y||^2 = \sum_{n=1}^{\infty} |(y, \phi_n)|^2,$$

holds for all $y \in H$ where $\{\phi_n\}_{n=1}^{\infty}$ is an orthonormal set, then $\{\phi_n\}_{n=1}^{\infty}$ is a maximal orthonormal set and

$$\lim_{N \to \infty} \left|\left| y - \sum_{n=1}^{N} (y, \phi_n)\, \phi_n \right|\right| = 0.$$

18. Suppose X is an infinite dimensional Banach space and suppose

$$\{x_1 \cdots x_n\}$$

are linearly independent with $||x_i|| = 1$. Show $span\,(x_1 \cdots x_n) \equiv X_n$ is a closed linear subspace of X. Now let $z \notin X_n$ and pick $y \in X_n$ such that $||z - y|| \leq 2 \text{ dist}\,(z, X_n)$ and let

$$x_{n+1} = \frac{z - y}{||z - y||}.$$

Show the sequence $\{x_k\}$ satisfies $||x_n - x_k|| \geq 1/2$ whenever $k < n$. **Hint:**

$$\left|\left| \frac{z - y}{||z - y||} - x_k \right|\right| = \left|\left| \frac{z - y - x_k\, ||z - y||}{||z - y||} \right|\right|.$$

Now show the unit ball $\{x \in X : ||x|| \leq 1\}$ is compact if and only if X is finite dimensional.

19. Show that if A is a self adjoint operator and $Ay = \lambda y$ for λ a complex number and $y \neq 0$, then λ must be real. Also verify that if A is self adjoint and $Ax = \mu x$ while $Ay = \lambda y$, then if $\mu \neq \lambda$, it must be the case that $(x, y) = 0$.

20. If Y is a closed subspace of a reflexive Banach space X, show Y is reflexive.

21. Show $H^k(\mathbb{R}^n)$ is a Hilbert space. See Problem 15 of Chapter 13 for a definition of this space.

Brouwer Degree

This chapter is on the Brouwer degree, a very useful concept with numerous and important applications. The degree can be used to prove some difficult theorems in topology such as the Brouwer fixed point theorem, the Jordan separation theorem, and the invariance of domain theorem. It also is used in bifurcation theory and many other areas in which it is an essential tool. Our emphasis in this chapter will be on the Brouwer degree for $\mathbb{R}^n$. When this is understood, it is not too difficult to extend to versions of the degree which hold in Banach space. There is more on degree theory in the book by Deimling [7] and much of the presentation here follows this reference.

16.1 Preliminary results

Definition 16.1 *For Ω a bounded open set, we denote by $C\left(\overline{\Omega}\right)$ the set of functions which are continuous on $\overline{\Omega}$ and by $C^m\left(\overline{\Omega}\right), m \leq \infty$ the space of restrictions of functions in $C_c^\infty\left(\mathbb{R}^n\right)$ to $\overline{\Omega}$. The norm in $C\left(\overline{\Omega}\right)$ is defined as follows.*

$$||f||_\infty = ||f||_{C\left(\overline{\Omega}\right)} \equiv \sup\left\{|f(\mathbf{x})| : \mathbf{x} \in \overline{\Omega}\right\}.$$

If the functions take values in $\mathbb{R}^n$ we will write $C^m\left(\overline{\Omega};\mathbb{R}^n\right)$ or $C\left(\overline{\Omega};\mathbb{R}^n\right)$ if there is no differentiability assumed. The norm on $C\left(\overline{\Omega};\mathbb{R}^n\right)$ is defined in the same way as above,

$$||\mathbf{f}||_\infty = ||\mathbf{f}||_{C\left(\overline{\Omega};\mathbb{R}^n\right)} \equiv \sup\left\{|\mathbf{f}(\mathbf{x})| : \mathbf{x} \in \overline{\Omega}\right\}.$$

Also, we will denote by $C\left(\Omega;\mathbb{R}^n\right)$ functions which are continuous on Ω that have values in $\mathbb{R}^n$ and by $C^m\left(\Omega;\mathbb{R}^n\right)$ functions which have m continuous derivatives defined on Ω.

Theorem 16.2 *Let Ω be a bounded open set in $\mathbb{R}^n$ and let $f \in C\left(\overline{\Omega}\right)$. Then there exists $g \in C^\infty\left(\overline{\Omega}\right)$ with $||g - f||_{C\left(\overline{\Omega}\right)} \leq \varepsilon$.*

Proof: This follows immediately from the Stone Weierstrass theorem. Let $\pi_i(\mathbf{x}) \equiv x_i$. Then the functions π_i and the constant function, $\pi_0(\mathbf{x}) \equiv 1$ separate the points of $\overline{\Omega}$ and annihilate no point. Therefore, the algebra generated by these functions, the polynomials, is dense in $C\left(\overline{\Omega}\right)$. Thus we may take g to be a polynomial.

Applying this result to the components of a vector valued function yields the following corollary.

Corollary 16.3 *If $\mathbf{f} \in C\left(\overline{\Omega};\mathbb{R}^n\right)$ for Ω a bounded subset of $\mathbb{R}^n$, then for all $\varepsilon > 0$, there exists $\mathbf{g} \in C^\infty\left(\overline{\Omega};\mathbb{R}^n\right)$ such that*

$$||\mathbf{g} - \mathbf{f}||_\infty < \varepsilon.$$

We make essential use of the following lemma a little later in establishing the definition of the degree, given below, is well defined.

Lemma 16.4 *Let $\mathbf{g} : U \to V$ be C^2 where U and V are open subsets of $\mathbb{R}^n$. Then*

$$\sum_{j=1}^{n} \left(cof(D\mathbf{g})\right)_{ij,j} = 0,$$

where here $(D\mathbf{g})_{ij} \equiv g_{i,j} \equiv \frac{\partial g_i}{\partial x_j}$.

Proof: $\det(D\mathbf{g}) = \sum_{i=1}^{n} g_{i,j} \text{cof}(D\mathbf{g})_{ij}$ and so

$$\frac{\partial \det(D\mathbf{g})}{\partial g_{i,j}} = \text{cof}(D\mathbf{g})_{ij}. \tag{16.1}$$

Also

$$\delta_{kj} \det(D\mathbf{g}) = \sum_i g_{i,k} \left(\text{cof}(D\mathbf{g})\right)_{ij}. \tag{16.2}$$

The reason for this is that if $k \neq j$ this is just the expansion of a determinant of a matrix in which the k *th* and j *th* columns are equal. Differentiate (16.2) with respect to x_j and sum on j. This yields

$$\sum_{r,s,j} \delta_{kj} \frac{\partial(\det D\mathbf{g})}{\partial g_{r,s}} g_{r,sj} = \sum_{ij} g_{i,kj} \left(\text{cof}(D\mathbf{g})\right)_{ij} + \sum_{ij} g_{i,k} \text{cof}(D\mathbf{g})_{ij,j}.$$

Hence, using $\delta_{kj} = 0$ if $j \neq k$ and (16.1),

$$\sum_{rs} \left(\text{cof}(D\mathbf{g})\right)_{rs} g_{r,sk} = \sum_{rs} g_{r,ks} \left(\text{cof}(D\mathbf{g})\right)_{rs} + \sum_{ij} g_{i,k} \text{cof}(D\mathbf{g})_{ij,j}.$$

Subtracting the first sum on the right from both sides and using the equality of mixed partials,

$$\sum_i g_{i,k} \left(\sum_j \left(\text{cof}(D\mathbf{g})\right)_{ij,j} \right) = 0.$$

If $\det(g_{i,k}) \neq 0$ so that $(g_{i,k})$ is invertible, this shows $\sum_j \left(\text{cof}(D\mathbf{g})\right)_{ij,j} = 0$. If $\det(D\mathbf{g}) = 0$, let

$$g_k = g + \epsilon_k I$$

where $\epsilon_k \to 0$ and $\det(D\mathbf{g} + \epsilon_k I) \equiv \det(D\mathbf{g}_k) \neq 0$. Then

$$\sum_j \left(\text{cof}(D\mathbf{g})\right)_{ij,j} = \lim_{k\to\infty} \sum_j \left(\text{cof}(D\mathbf{g}_k)\right)_{ij,j} = 0$$

and this proves the lemma.

16.2 Definitions and elementary properties

In this section, $\mathbf{f} : \overline{\Omega} \to \mathbb{R}^n$ will be a continuous map. We make the following definitions.

Definition 16.5 $\mathcal{U}_{\mathbf{y}} \equiv \left\{\mathbf{f} \in C\left(\overline{\Omega};\mathbb{R}^n\right) : \mathbf{y} \notin \mathbf{f}\left(\partial\Omega\right)\right\}$. *For two functions,*

$$\mathbf{f},\mathbf{g} \in \mathcal{U}_{\mathbf{y}},$$

we will say $\mathbf{f} \sim \mathbf{g}$ *if there exists a continuous function,*

$$\mathbf{h} : \overline{\Omega} \times [0,1] \to \mathbb{R}^n$$

such that $\mathbf{h}(\mathbf{x},1) = \mathbf{g}(\mathbf{x})$ *and* $\mathbf{h}(\mathbf{x},0) = \mathbf{f}(\mathbf{x})$ *and* $\mathbf{x} \to \mathbf{h}(\mathbf{x},t) \in \mathcal{U}_{\mathbf{y}}$. *This function,* $\mathbf{h}$, *is called a homotopy and we say that* $\mathbf{f}$ *and* $\mathbf{g}$ *are homotopic.*

Definition 16.6 *For* W *an open set in* $\mathbb{R}^n$ *and* $\mathbf{g} \in C^1(W;\mathbb{R}^n)$ *we say* $\mathbf{y}$ *is a regular value of* $\mathbf{g}$ *if whenever* $\mathbf{x} \in \mathbf{g}^{-1}(\mathbf{y})$, $\det(D\mathbf{g}(\mathbf{x})) \neq 0$. *Note that if* $\mathbf{g}^{-1}(\mathbf{y}) = \emptyset$, *it follows that* $\mathbf{y}$ *is a regular value from this definition.*

Lemma 16.7 *The relation,* $\sim$, *is an equivalence relation and, denoting by* $[\mathbf{f}]$ *the equivalence class determined by* $\mathbf{f}$, *it follows that* $[\mathbf{f}]$ *is an open subset of* $\mathcal{U}_{\mathbf{y}}$. *Furthermore,* $\mathcal{U}_{\mathbf{y}}$ *is an open set in* $C\left(\overline{\Omega};\mathbb{R}^n\right)$. *In addition to this, if* $\mathbf{f} \in \mathcal{U}_{\mathbf{y}}$, *there exists* $\mathbf{g} \in [\mathbf{f}] \cap C^2\left(\overline{\Omega};\mathbb{R}^n\right)$ *for which* $\mathbf{y}$ *is a regular value.*

Proof: In showing that $\sim$ is an equivalence relation, it is easy to verify that $\mathbf{f} \sim \mathbf{f}$ and that if $\mathbf{f} \sim \mathbf{g}$, then $\mathbf{g} \sim \mathbf{f}$. To verify the transitive property for an equivalence relation, suppose $\mathbf{f} \sim \mathbf{g}$ and $\mathbf{g} \sim \mathbf{k}$, with the homotopy for $\mathbf{f}$ and $\mathbf{g}$, the function, $\mathbf{h}_1$ and the homotopy for $\mathbf{g}$ and $\mathbf{k}$, the function $\mathbf{h}_2$. Thus $\mathbf{h}_1(\mathbf{x},0) = \mathbf{f}(\mathbf{x})$, $\mathbf{h}_1(\mathbf{x},1) = \mathbf{g}(\mathbf{x})$ and $\mathbf{h}_2(\mathbf{x},0) = \mathbf{g}(\mathbf{x})$, $\mathbf{h}_2(\mathbf{x},1) = \mathbf{k}(\mathbf{x})$. Then we define a homotopy of $\mathbf{f}$ and $\mathbf{k}$ as follows.

$$\mathbf{h}(\mathbf{x},t) \equiv \begin{cases} \mathbf{h}_1(\mathbf{x},2t) \text{ if } t \in \left[0,\frac{1}{2}\right] \\ \mathbf{h}_2(\mathbf{x},2t-1) \text{ if } t \in \left[\frac{1}{2},1\right] \end{cases}.$$

It is obvious that $\mathcal{U}_{\mathbf{y}}$ is an open subset of $C\left(\overline{\Omega};\mathbb{R}^n\right)$. We need to argue that $[\mathbf{f}]$ is also an open set. However, if $\mathbf{f} \in \mathcal{U}_{\mathbf{y}}$, There exists $\delta > 0$ such that $B(\mathbf{y},2\delta) \cap \mathbf{f}(\partial\Omega) = \emptyset$. Let $\mathbf{f}_1 \in C\left(\overline{\Omega};\mathbb{R}^n\right)$ with $||\mathbf{f}_1 - \mathbf{f}||_\infty < \delta$. Then if $t \in [0,1]$, and $\mathbf{x} \in \partial\Omega$

$$|\mathbf{f}(\mathbf{x}) + t(\mathbf{f}_1(\mathbf{x}) - \mathbf{f}(\mathbf{x})) - \mathbf{y}| \geq |\mathbf{f}(\mathbf{x}) - \mathbf{y}| - t\,||\mathbf{f} - \mathbf{f}_1||_\infty > 2\delta - t\delta > 0.$$

Therefore, $B(\mathbf{f},\delta) \subseteq [\mathbf{f}]$ because if $\mathbf{f}_1 \in B(\mathbf{f},\delta)$, this shows that, letting $\mathbf{h}(\mathbf{x},t) \equiv \mathbf{f}(\mathbf{x}) + t(\mathbf{f}_1(\mathbf{x}) - \mathbf{f}(\mathbf{x}))$, $\mathbf{f}_1 \sim \mathbf{f}$.

It remains to verify the last assertion of the lemma. Since $[\mathbf{f}]$ is an open set, there exists $\mathbf{g} \in [\mathbf{f}] \cap C^2\left(\overline{\Omega};\mathbb{R}^n\right)$. If $\mathbf{y}$ is a regular value of $\mathbf{g}$, leave $\mathbf{g}$ unchanged. Otherwise, let

$$S \equiv \left\{\mathbf{x} \in \overline{\Omega} : \det D\mathbf{g}(\mathbf{x}) = 0\right\}$$

and pick $\delta > 0$ small enough that $B(\mathbf{y},2\delta) \cap \mathbf{g}(\partial\Omega) = \emptyset$. By Sard's lemma, $\mathbf{g}(S)$ is a set of measure zero and so there exists $\widetilde{\mathbf{y}} \in B(\mathbf{y},\delta) \setminus \mathbf{g}(S)$. Thus $\widetilde{\mathbf{y}}$ is a regular value of $\mathbf{g}$. Now define $\mathbf{g}_1(\mathbf{x}) \equiv \mathbf{g}(\mathbf{x}) + \mathbf{y} - \widetilde{\mathbf{y}}$. It follows that $\mathbf{g}_1(\mathbf{x}) = \mathbf{y}$ if and only if $\mathbf{g}(\mathbf{x}) = \widetilde{\mathbf{y}}$ and so, since $D\mathbf{g}(\mathbf{x}) = D\mathbf{g}_1(\mathbf{x})$, $\mathbf{y}$ is a regular value of $\mathbf{g}_1$. Then for $t \in [0,1]$ and $\mathbf{x} \in \partial\Omega$,

$$|\mathbf{g}(\mathbf{x}) + t(\mathbf{g}_1(\mathbf{x}) - \mathbf{g}(\mathbf{x})) - \mathbf{y}| \geq |\mathbf{g}(\mathbf{x}) - \mathbf{y}| - t\,|\mathbf{y} - \widetilde{\mathbf{y}}| > 2\delta - t\delta \geq \delta > 0.$$

It follows $\mathbf{g}_1 \sim \mathbf{g}$ and so $\mathbf{g}_1 \sim \mathbf{f}$. This proves the lemma since $\mathbf{y}$ is a regular value of $\mathbf{g}_1$.

Definition 16.8 *Let* $\mathbf{g} \in \mathcal{U}_{\mathbf{y}} \cap C^2\left(\overline{\Omega};\mathbb{R}^n\right)$ *and let* $\mathbf{y}$ *be a regular value of* $\mathbf{g}$. *Then*

$$d(\mathbf{g},\Omega,\mathbf{y}) \equiv \sum \left\{ \mathit{sign}\left(\det D\mathbf{g}(\mathbf{x})\right) : \mathbf{x} \in \mathbf{g}^{-1}(\mathbf{y}) \right\}, \tag{16.3}$$

$$d(\mathbf{g},\Omega,\mathbf{y}) = 0 \text{ } \mathit{if}\ \mathbf{g}^{-1}(\mathbf{y}) = \emptyset, \tag{16.4}$$

and if $\mathbf{f} \in \mathcal{U}_{\mathbf{y}}$,

$$d(\mathbf{f}, \Omega, \mathbf{y}) = d(\mathbf{g}, \Omega, \mathbf{y}) \tag{16.5}$$

where $\mathbf{g} \in [\mathbf{f}] \cap C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$, *and* $\mathbf{y}$ *is a regular value of* $\mathbf{g}$. *This number, denoted by* $d(\mathbf{f}, \Omega, \mathbf{y})$ *is called the degree or Brouwer degree.*

We need to verify this is well defined. We begin with the definition, (16.3). We need to show that the sum is finite.

Lemma 16.9 *When* $\mathbf{y}$ *is a regular value, the sum in (16.3) is finite.*

Proof: This follows from the inverse function theorem because $\mathbf{g}^{-1}(\mathbf{y})$ is a closed, hence compact subset of Ω due to the assumption that $\mathbf{y} \notin \mathbf{g}(\partial\Omega)$. Since $\mathbf{y}$ is a regular value, it follows that $\det(D\mathbf{g}(\mathbf{x})) \neq 0$ for each $\mathbf{x} \in \mathbf{g}^{-1}(\mathbf{y})$. By the inverse function theorem, there is an open set, $U_{\mathbf{x}}$, containing $\mathbf{x}$ such that $\mathbf{g}$ is one to one on this open set. Since $\mathbf{g}^{-1}(\mathbf{y})$ is compact, this means there are finitely many sets, $U_{\mathbf{x}}$ which cover $\mathbf{g}^{-1}(\mathbf{y})$, each containing only one point of $\mathbf{g}^{-1}(\mathbf{y})$. Therefore, this set is finite and so the sum is well defined.

A much more difficult problem is to show there are no contradictions in the two ways $d(\mathbf{f}, \Omega, \mathbf{y})$ is defined in the case when $\mathbf{f} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ and $\mathbf{y}$ is a regular value of $\mathbf{f}$. We need to verify that if $\mathbf{g}_0 \sim \mathbf{g}_1$ for $\mathbf{g}_i \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ and $\mathbf{y}$ a regular value of $\mathbf{g}_i$, it follows that $d(\mathbf{g}_1, \Omega, \mathbf{y}) = d(\mathbf{g}_2, \Omega, \mathbf{y})$ under the conditions of (16.3) and (16.4). To aid in this, we give the following lemma.

Lemma 16.10 *Suppose* $\mathbf{k} \sim \mathbf{l}$. *Then there exists a sequence of functions of* $\mathcal{U}_{\mathbf{y}}$,

$$\{\mathbf{g}_i\}_{i=1}^m,$$

such that $\mathbf{g}_i \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$, $\mathbf{y}$ *is a regular value for* $\mathbf{g}_i$, *and defining* $\mathbf{g}_0 \equiv \mathbf{k}$ *and* $\mathbf{g}_{m+1} \equiv \mathbf{l}$, *there exists* $\delta > 0$ *such that for* $i = 1, \cdots, m+1$,

$$B(\mathbf{y}, \boldsymbol{\delta}) \cap (t\mathbf{g}_i + (1-t)\mathbf{g}_{i-1})(\partial\Omega) = \emptyset, \text{ for all } t \in [0,1]. \tag{16.6}$$

Proof: Let $\mathbf{h} : \overline{\Omega} \times [0,1] \to \mathbb{R}^n$ be a function which shows $\mathbf{k}$ and $\mathbf{l}$ are equivalent. Now let $0 = t_0 < t_1 < \cdots < t_m = 1$ be such that

$$||\mathbf{h}(\cdot, t_i) - \mathbf{h}(\cdot, t_{i-1})||_\infty < \delta \tag{16.7}$$

where $\delta > 0$ is small enough that

$$B(\mathbf{y}, 8\delta) \cap \mathbf{h}(\partial\Omega \times [0,1]) = \emptyset. \tag{16.8}$$

Now for $i \in \{1, \cdots, m\}$, let $\mathbf{g}_i \in \mathcal{U}_{\mathbf{y}} \cap C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ be such that

$$||\mathbf{g}_i - \mathbf{h}(\cdot, t_i)||_\infty < \delta. \tag{16.9}$$

This is possible because $C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ is dense in $C\left(\overline{\Omega}; \mathbb{R}^n\right)$ from Corollary 16.3. If $\mathbf{y}$ is a regular value for $\mathbf{g}_i$, leave $\mathbf{g}_i$ unchanged. Otherwise, using Sard's lemma, let $\widetilde{\mathbf{y}}$ be a regular value of $\mathbf{g}_i$ close enough to $\mathbf{y}$ that the function, $\widetilde{\mathbf{g}_i} \equiv \mathbf{g}_i + \mathbf{y} - \widetilde{\mathbf{y}}$ also satisfies (16.9). Then $\widetilde{\mathbf{g}_i}(\mathbf{x}) = \mathbf{y}$ if and only if $\mathbf{g}_i(\mathbf{x}) = \widetilde{\mathbf{y}}$. Thus $\mathbf{y}$ is a regular value for $\widetilde{\mathbf{g}_i}$ and we may replace $\mathbf{g}_i$ with $\widetilde{\mathbf{g}_i}$ in (16.9). Therefore, we can assume that $\mathbf{y}$ is a regular value for $\mathbf{g}_i$ in (16.9). Now from this construction,

$$||\mathbf{g}_i - \mathbf{g}_{i-1}||_\infty \leq ||\mathbf{g}_i - \mathbf{h}(\cdot, t_i)||$$

$$+ ||\mathbf{h}(\cdot, t_i) - \mathbf{h}(\cdot, t_{i-1})|| + ||\mathbf{g}_{i-1} - \mathbf{h}(\cdot, t_{i-1})|| < 3\delta.$$

Now we verify (16.6). We just showed that for each $\mathbf{x} \in \partial\Omega$,

$$\mathbf{g}_i(\mathbf{x}) \in B(\mathbf{g}_{i-1}(\mathbf{x}), 3\delta)$$

and we also know from (16.8) and (16.9) that for any j,

$$\begin{aligned} |\mathbf{g}_j(\mathbf{x}) - \mathbf{y}| &\geq -|\mathbf{g}_j(\mathbf{x}) - \mathbf{h}(\mathbf{x}, t_j)| + |\mathbf{h}(\mathbf{x}, t_j) - \mathbf{y}| \\ &\geq 8\delta - \delta = 7\delta. \end{aligned}$$

Therefore, for $\mathbf{x} \in \partial\Omega$,

$$|t\mathbf{g}_i(\mathbf{x}) + (1-t)\mathbf{g}_{i-1}(\mathbf{x}) - \mathbf{y}| = |\mathbf{g}_{i-1}(\mathbf{x}) + t(\mathbf{g}_i(\mathbf{x}) - \mathbf{g}_{i-1}(\mathbf{x})) - \mathbf{y}|$$

$$\geq 7\delta - t|\mathbf{g}_i(\mathbf{x}) - \mathbf{g}_{i-1}(\mathbf{x})| > 7\delta - t3\delta \geq 4\delta > \delta.$$

This proves the lemma.

We make the following definition of a set of functions.

Definition 16.11 *For each $\varepsilon > 0$, let*

$$\phi_\varepsilon \in C_c^\infty(B(\mathbf{0}, \varepsilon)), \phi_\varepsilon \geq 0, \quad \int \phi_\varepsilon dx = 1.$$

Lemma 16.12 *Let $\mathbf{g} \in \mathcal{U}_{\mathbf{y}} \cap C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ and let $\mathbf{y}$ be a regular value of $\mathbf{g}$. Then according to the definition of degree given in (16.3) and (16.4),*

$$d(\mathbf{g}, \Omega, \mathbf{y}) = \int_\Omega \phi_\varepsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \det D\mathbf{g}(\mathbf{x})\, dx \tag{16.10}$$

whenever ε is small enough. Also $\mathbf{y} + \mathbf{v}$ is a regular value of $\mathbf{g}$ whenever $|\mathbf{v}|$ is small enough.

Proof: Let the points in $\mathbf{g}^{-1}(\mathbf{y})$ be $\{\mathbf{x}_i\}_{i=1}^m$. By the inverse function theorem, there exist disjoint open sets, $U_i, \mathbf{x}_i \in U_i$, such that $\mathbf{g}$ is one to one on U_i with $\det(D\mathbf{g}(x))$ having constant sign on U_i and $\mathbf{g}(U_i)$ is an open set containing $\mathbf{y}$. Then let ε be small enough that

$$B(\mathbf{y}, \varepsilon) \subseteq \cap_{i=1}^m \mathbf{g}(U_i)$$

and let $V_i \equiv \mathbf{g}^{-1}(B(\mathbf{y}, \varepsilon)) \cap U_i$. Therefore, for any ε this small,

$$\int_\Omega \phi_\varepsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \det D\mathbf{g}(\mathbf{x})\, dx = \sum_{i=1}^m \int_{V_i} \phi_\varepsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \det D\mathbf{g}(\mathbf{x})\, dx$$

The reason for this is as follows. The integrand is nonzero only if $\mathbf{g}(\mathbf{x}) - \mathbf{y} \in B(\mathbf{0}, \varepsilon)$ which occurs only if $\mathbf{g}(\mathbf{x}) \in B(\mathbf{y}, \varepsilon)$ which is the same as $\mathbf{x} \in \mathbf{g}^{-1}(B(\mathbf{y}, \varepsilon))$. Therefore, the integrand is nonzero only if $\mathbf{x}$ is contained in exactly one of the sets, V_i. Now using the change of variables theorem,

$$= \sum_{i=1}^m \int_{\mathbf{g}(V_i) - \mathbf{y}} \phi_\varepsilon(\mathbf{z}) \det D\mathbf{g}\left(\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})\right) \left|\det D\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})\right| dz$$

By the chain rule, $I = D\mathbf{g}\left(\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})\right) D\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})$ and so

$$\det D\mathbf{g}\left(\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})\right) \left|\det D\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})\right| = \operatorname{sign}\left(\det D\mathbf{g}\left(\mathbf{g}^{-1}(\mathbf{y} + \mathbf{z})\right)\right)$$

$$= \text{sign}(\det D\mathbf{g}(\mathbf{x})) = \text{sign}(\det D\mathbf{g}(\mathbf{x}_i)).$$

Therefore, this reduces to

$$\sum_{i=1}^{m} \text{sign}(\det D\mathbf{g}(\mathbf{x}_i)) \int_{\mathbf{g}(V_i)-\mathbf{y}} \phi_\varepsilon(\mathbf{z})\, dz =$$

$$\sum_{i=1}^{m} \text{sign}(\det D\mathbf{g}(\mathbf{x}_i)) \int_{B(\mathbf{0},\varepsilon)} \phi_\varepsilon(\mathbf{z})\, dz = \sum_{i=1}^{m} \text{sign}(\det D\mathbf{g}(\mathbf{x}_i)).$$

In case $\mathbf{g}^{-1}(\mathbf{y}) = \emptyset$, there exists $\varepsilon > 0$ such that $\mathbf{g}\left(\overline{\Omega}\right) \cap B(\mathbf{y}, \varepsilon) = \emptyset$ and so for ε this small,

$$\int_\Omega \phi_\varepsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \det D\mathbf{g}(\mathbf{x})\, dx = 0.$$

The last assertion of the lemma follows from the observation that $\mathbf{g}(S)$ is a compact set and so its complement is an open set. This proves the lemma.

Now we are ready to prove a lemma which will complete the task of showing the above definition of the degree is well defined. In the following lemma, and elsewhere, a comma followed by an index indicates the partial derivative with respect to the indicated variable. Thus, $f_{,j}$ will mean $\frac{\partial f}{\partial x_j}$.

Lemma 16.13 *Suppose* $\mathbf{f}, \mathbf{g}$ *are two functions in* $C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ *and*

$$B(\mathbf{y}, \varepsilon) \cap ((1-t)\mathbf{f} + t\mathbf{g})(\partial\Omega) = \emptyset \tag{16.11}$$

for all $t \in [0,1]$. *Then*

$$\int_\Omega \phi_\varepsilon(\mathbf{f}(\mathbf{x}) - \mathbf{y}) \det(D\mathbf{f}(\mathbf{x}))\, dx = \int_\Omega \phi_\varepsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \det(D\mathbf{g}(\mathbf{x}))\, dx. \tag{16.12}$$

Proof: Define for $t \in [0,1]$,

$$H(t) \equiv \int_\Omega \phi_\varepsilon(\mathbf{f} - \mathbf{y} + t(\mathbf{g} - \mathbf{f})) \det(D(\mathbf{f} + t(\mathbf{g} - \mathbf{f})))\, dx.$$

Then if $t \in (0,1)$,

$$H'(t) = \int_\Omega \sum_\alpha \phi_{\varepsilon,\alpha}(\mathbf{f}(\mathbf{x}) - \mathbf{y} + t(\mathbf{g}(\mathbf{x}) - \mathbf{f}(\mathbf{x}))) \cdot$$

$$(g_\alpha(\mathbf{x}) - f_\alpha(\mathbf{x})) \det D(\mathbf{f} + t(\mathbf{g} - \mathbf{f}))\, dx$$

$$+ \int_\Omega \phi_\varepsilon(\mathbf{f} - \mathbf{y} + t(\mathbf{g} - \mathbf{f})) \cdot$$

$$\sum_{\alpha,j} \det D(\mathbf{f} + t(\mathbf{g} - \mathbf{f}))_{,\alpha j} (g_\alpha - f_\alpha)_{,j}\, dx \equiv \mathbf{A} + \mathbf{B}.$$

In this formula, the function det is considered as a function of the n^2 entries in the $n \times n$ matrix. Now as in the proof of Lemma 16.4,

$$\det D(\mathbf{f} + t(\mathbf{g} - \mathbf{f}))_{,\alpha j} = (\text{cof } D(\mathbf{f} + t(\mathbf{g} - \mathbf{f})))_{\alpha j}$$

and so

$$\mathbf{B} = \int_\Omega \sum_\alpha \sum_j \phi_\varepsilon \left(\mathbf{f} - \mathbf{y} + t\left(\mathbf{g} - \mathbf{f}\right)\right) \cdot$$

$$\left(\text{cof } D\left(\mathbf{f}+t\left(\mathbf{g} - \mathbf{f}\right)\right)\right)_{\alpha j} \left(g_\alpha - f_\alpha\right)_{,j} dx.$$

By hypothesis

$$\mathbf{x} \rightarrow \phi_\varepsilon \left(\mathbf{f}\left(\mathbf{x}\right) - \mathbf{y} + t\left(\mathbf{g}\left(\mathbf{x}\right) - \mathbf{f}\left(\mathbf{x}\right)\right)\right) \left(\text{cof } D\left(\mathbf{f}\left(\mathbf{x}\right) + t\left(\mathbf{g}\left(\mathbf{x}\right) - \mathbf{f}\left(\mathbf{x}\right)\right)\right)\right)_{\alpha j}$$

is in $C_c^1(\Omega)$ because if $\mathbf{x} \in \partial\Omega$, it follows by (16.11) that

$$\mathbf{f}\left(\mathbf{x}\right) - \mathbf{y} + t\left(\mathbf{g}\left(\mathbf{x}\right) - \mathbf{f}\left(\mathbf{x}\right)\right) \notin B\left(\mathbf{0}, \varepsilon\right).$$

Therefore, we may integrate by parts and write

$$\mathbf{B} = -\int_\Omega \sum_\alpha \sum_j \frac{\partial}{\partial x_j} \left(\phi_\varepsilon \left(\mathbf{f} - \mathbf{y} + t\left(\mathbf{g} - \mathbf{f}\right)\right)\right) \cdot$$

$$\left(\text{cof } D\left(\mathbf{f}+t\left(\mathbf{g} - \mathbf{f}\right)\right)\right)_{\alpha j} \left(g_\alpha - f_\alpha\right) dx +$$

$$-\int_\Omega \sum_\alpha \sum_j \phi_\varepsilon \left(\mathbf{f} - \mathbf{y} + t\left(\mathbf{g} - \mathbf{f}\right)\right) \left(\text{cof } D\left(\mathbf{f}+t\left(\mathbf{g} - \mathbf{f}\right)\right)\right)_{\alpha j,j} \left(g_\alpha - f_\alpha\right) dx.$$

The second term equals zero by Lemma 16.4. Simplifying the first term yields

$$\mathbf{B} = -\int_\Omega \sum_\alpha \sum_j \sum_\beta \phi_{\varepsilon,\beta} \left(\mathbf{f} - \mathbf{y} + t\left(\mathbf{g} - \mathbf{f}\right)\right) \cdot$$

$$\left(f_{\beta,j} + t\left(g_{\beta,j} - f_{\beta,j}\right)\right) \left(\text{cof } D\left(\mathbf{f}+t\left(\mathbf{g} - \mathbf{f}\right)\right)\right)_{\alpha j} \left(g_\alpha - f_\alpha\right) dx$$

$$= -\int_\Omega \sum_\alpha \sum_\beta \phi_{\varepsilon,\beta} \left(\mathbf{f} - \mathbf{y} + t\left(\mathbf{g} - \mathbf{f}\right)\right) \delta_{\beta\alpha} \det\left(D\left(\mathbf{f}+t\left(\mathbf{g} - \mathbf{f}\right)\right)\right) \left(g_\alpha - f_\alpha\right) dx$$

$$= -\int_\Omega \sum_\alpha \phi_{\varepsilon,\alpha} \left(\mathbf{f} - \mathbf{y} + t\left(\mathbf{g} - \mathbf{f}\right)\right) \det\left(D\left(\mathbf{f}+t\left(\mathbf{g} - \mathbf{f}\right)\right)\right) \left(g_\alpha - f_\alpha\right) dx = -\mathbf{A}.$$

Therefore, $H'(t) = 0$ and so H is a constant. This proves the lemma.

Now we are ready to prove that the Brouwer degree is well defined.

Proposition 16.14 *Definition 16.8 is well defined.*

Proof: We only need to verify that for $\mathbf{k}, \mathbf{l} \in \mathcal{U}_\mathbf{y}$ $\mathbf{k} \sim \mathbf{l}$, $\mathbf{k}, \mathbf{l} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$, and $\mathbf{y}$ a regular value of $\mathbf{k}, \mathbf{l}$, $d\left(\mathbf{k}, \Omega, \mathbf{y}\right) = d\left(\mathbf{l}, \Omega, \mathbf{y}\right)$ as given in the first part of Definition 16.8. Let the functions, $\mathbf{g}_i$, $i = 1, \cdots, m$ be as described in Lemma 16.10. By Lemma 16.8 we may take $\varepsilon > 0$ small enough that equation (16.10) holds for $\mathbf{g} = \mathbf{k}, \mathbf{l}$. Then by Lemma 16.13 and letting $\mathbf{g}_0 \equiv \mathbf{k}$, and $\mathbf{g}_{m+1} = \mathbf{l}$,

$$\int_\Omega \phi_\varepsilon \left(\mathbf{g}_i\left(\mathbf{x}\right) - \mathbf{y}\right) \det D\mathbf{g}_i\left(\mathbf{x}\right) dx = \int_\Omega \phi_\varepsilon \left(\mathbf{g}_{i-1}\left(\mathbf{x}\right) - \mathbf{y}\right) \det D\mathbf{g}_{i-1}\left(\mathbf{x}\right) dx$$

for $i = 1, \cdots, m+1$. In particular $d\left(\mathbf{k}, \Omega, \mathbf{y}\right) = d\left(\mathbf{l}, \Omega, \mathbf{y}\right)$ proving the proposition.

The degree has many wonderful properties. We begin with a simple technical lemma which will allow us to establish them.

Lemma 16.15 *Let* $\mathbf{y}_1 \notin \mathbf{f}(\partial\Omega)$. *Then* $d(\mathbf{f},\Omega,\mathbf{y}_1) = d(\mathbf{f},\Omega,\mathbf{y})$ *whenever* $\mathbf{y}$ *is close enough to* $\mathbf{y}_1$. *Also,* $d(\mathbf{f},\Omega,\mathbf{y})$ *equals an integer.*

Proof: It follows immediately from the definition of the degree that

$$d(\mathbf{f},\Omega,\mathbf{y})$$

is an integer. Let $\mathbf{g} \in C^2\left(\overline{\Omega};\mathbb{R}^n\right)$ for which $\mathbf{y}_1$ is a regular value and let $\mathbf{h} : \overline{\Omega}\times[0,1] \to \mathbb{R}^n$ be a continuous function for which $\mathbf{h}(\mathbf{x},0) = \mathbf{f}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x},1) = \mathbf{g}(\mathbf{x})$ such that $\mathbf{h}(\partial\Omega,t)$ does not contain $\mathbf{y}_1$ for any $t \in [0,1]$. Then let ε_1 be small enough that

$$B(\mathbf{y}_1,\varepsilon_1) \cap \mathbf{h}(\partial\Omega\times[0,1]) = \emptyset.$$

From Lemma 16.12, we may take $\varepsilon < \varepsilon_1$ small enough that whenever $|\mathbf{v}| < \varepsilon$, $\mathbf{y}_1 + \mathbf{v}$ is a regular value of $\mathbf{g}$ and

$$d(\mathbf{g},\Omega,\mathbf{y}_1) = \sum\left\{\text{sign } D\mathbf{g}(\mathbf{x}) : \mathbf{x} \in \mathbf{g}^{-1}(\mathbf{y}_1)\right\}$$

$$= \sum\left\{\text{sign } D\mathbf{g}(\mathbf{x}) : \mathbf{x} \in \mathbf{g}^{-1}(\mathbf{y}_1+\mathbf{v})\right\} = d(\mathbf{g},\Omega,\mathbf{y}_1+\mathbf{v}).$$

The second equal sign above needs some justification. We know $\mathbf{g}^{-1}(\mathbf{y}_1) = \{\mathbf{x}_1,\cdots,\mathbf{x}_m\}$ and by the inverse function theorem, there are open sets, U_i such that $\mathbf{x}_i \in U_i \subseteq \Omega$ and $\mathbf{g}$ is one to one on U_i having an inverse on the open set $\mathbf{g}(U_i)$ which is also C^2. We want to say that for $|\mathbf{v}|$ small enough, $\mathbf{g}^{-1}(\mathbf{y}_1+\mathbf{v}) \subseteq \cup_{j=1}^m U_i$. If not, there exists $\mathbf{v}_k \to \mathbf{0}$ and

$$\mathbf{z}_k \in \mathbf{g}^{-1}(\mathbf{y}_1+\mathbf{v}_k) \setminus \cup_{j=1}^m U_i.$$

But then, taking a subsequence, still denoted by $\mathbf{z}_k$ we could have $\mathbf{z}_k \to \mathbf{z} \notin \cup_{j=1}^m U_i$ and so

$$\mathbf{g}(\mathbf{z}) = \lim_{k\to\infty} \mathbf{g}(\mathbf{z}_k) = \lim_{k\to\infty}(\mathbf{y}_1+\mathbf{v}_k) = \mathbf{y}_1,$$

contradicting the fact that $\mathbf{g}^{-1}(\mathbf{y}_1) \subseteq \cup_{j=1}^m U_i$. This justifies the second equal sign.

For the above homotopy of $\mathbf{f}$ and $\mathbf{g}$, if $\mathbf{x} \in \partial\Omega$,

$$|\mathbf{h}(\mathbf{x},t) - (\mathbf{y}_1+\mathbf{v})| \geq |\mathbf{h}(\mathbf{x},t) - \mathbf{y}_1| - |\mathbf{v}| > \varepsilon_1 - \varepsilon > 0.$$

Therefore, by the definition of the degree,

$$d(\mathbf{f},\Omega,\mathbf{y}) = d(\mathbf{g},\Omega,\mathbf{y}) = d(\mathbf{g},\Omega,\mathbf{y}_1+\mathbf{v}) = d(\mathbf{f},\Omega,\mathbf{y}_1+\mathbf{v}).$$

This proves the lemma.

Here are some important properties of the degree.

Theorem 16.16 *The degree satisfies the following properties. In what follows,* $id(\mathbf{x}) = \mathbf{x}$.

1. $d(id,\Omega,\mathbf{y}) = 1$ *if* $\mathbf{y} \in \Omega$.
2. *If* $\Omega_i \subseteq \Omega, \Omega_i$ *open, and* $\Omega_1 \cap \Omega_2 = \emptyset$ *and if* $\mathbf{y} \notin \mathbf{f}\left(\overline{\Omega}\setminus(\Omega_1\cup\Omega_2)\right)$, *then* $d(\mathbf{f},\Omega_1,\mathbf{y}) + d(\mathbf{f},\Omega_2,\mathbf{y}) = d(\mathbf{f},\Omega,\mathbf{y})$.
3. *If* $\mathbf{y} \notin \mathbf{f}\left(\overline{\Omega}\setminus\Omega_1\right)$ *and* Ω_1 *is an open subset of* Ω, *then*

$$d(\mathbf{f},\Omega,\mathbf{y}) = d(\mathbf{f},\Omega_1,\mathbf{y}).$$

4. $d(\mathbf{f},\Omega,\mathbf{y}) \neq 0$ *implies* $\mathbf{f}^{-1}(\mathbf{y}) \neq \emptyset$.
5. *If* $\mathbf{f},\mathbf{g}$ *are homotopic with a homotopy,* $\mathbf{h} : \overline{\Omega} \times [0,1]$ *for which* $\mathbf{h}(\partial\Omega, t)$ *does not contain* $\mathbf{y}$, *then* $d(\mathbf{g},\Omega,\mathbf{y}) = d(\mathbf{f},\Omega,\mathbf{y})$.
6. $d(\cdot,\Omega,\mathbf{y})$ *is defined and constant on*
$$\left\{\mathbf{g} \in C\left(\overline{\Omega};\mathbb{R}^n\right) : ||\mathbf{g}-\mathbf{f}||_\infty < r\right\}$$
where $r = dist(\mathbf{y}, \mathbf{f}(\partial\Omega))$.
7. $d(\mathbf{f},\Omega,\cdot)$ *is constant on every connected component of* $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$.
8. $d(\mathbf{g},\Omega,\mathbf{y}) = d(\mathbf{f},\Omega,\mathbf{y})$ *if* $\mathbf{g}|_{\partial\Omega} = \mathbf{f}|_{\partial\Omega}$.

Proof: The first property follows immediately from the definition of the degree.

To obtain the second property, let δ be small enough that
$$B(\mathbf{y},3\delta) \cap \mathbf{f}\left(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)\right) = \emptyset.$$
Next pick $\mathbf{g} \in C^2\left(\overline{\Omega};\mathbb{R}^n\right)$ such that $||\mathbf{f}-\mathbf{g}||_\infty < \delta$. Letting $\mathbf{y}_1$ be a regular value of $\mathbf{g}$ with $|\mathbf{y}_1 - \mathbf{y}| < \delta$, and defining $\mathbf{g}_1(\mathbf{x}) \equiv \mathbf{g}(\mathbf{x}) + \mathbf{y} - \mathbf{y}_1$, we see that $||\mathbf{g}-\mathbf{g}_1||_\infty < \delta$ and $\mathbf{y}$ is a regular value of $\mathbf{g}_1$. Then if $\mathbf{x} \in \overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)$, it follows that for $t \in [0,1]$,
$$|\mathbf{f}(\mathbf{x}) + t(\mathbf{g}_1(\mathbf{x}) - \mathbf{f}(\mathbf{x})) - \mathbf{y}| \geq |\mathbf{f}(\mathbf{x}) - \mathbf{y}| - t\,||\mathbf{g}_1 - \mathbf{f}||_\infty > 3\delta - 2\delta > 0$$
and so $\mathbf{g}_1\left(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2)\right)$ does not contain $\mathbf{y}$. Hence $\mathbf{g}_1^{-1}(\mathbf{y}) \subseteq \Omega_1 \cup \Omega_2$ and $\mathbf{g}_1 \sim \mathbf{f}$. Therefore, from the definition of the degree,
$$\begin{aligned} d(\mathbf{f},\Omega,\mathbf{y}) &\equiv d(\mathbf{g},\Omega,\mathbf{y}) = d(\mathbf{g},\Omega_1,\mathbf{y}) + d(\mathbf{g},\Omega_2,\mathbf{y}) \\ &\equiv d(\mathbf{f},\Omega_1,\mathbf{y}) + d(\mathbf{f},\Omega_2,\mathbf{y}) \end{aligned}$$

The third property follows from the second if we let $\Omega_2 = \emptyset$. In the above formula, $d(\mathbf{g},\Omega_2,\mathbf{y}) = 0$ in this case.

Now consider the fourth property. If $\mathbf{f}^{-1}(\mathbf{y}) = \emptyset$, then for $\delta > 0$ small enough, $B(\mathbf{y},3\delta) \cap \mathbf{f}\left(\overline{\Omega}\right) = \emptyset$. Let $\mathbf{g}$ be in C^2 and $||\mathbf{f}-\mathbf{g}||_\infty < \delta$ with $\mathbf{y}$ a regular point of $\mathbf{g}$. Then $d(\mathbf{f},\Omega,\mathbf{y}) = d(\mathbf{g},\Omega,\mathbf{y})$ and $\mathbf{g}^{-1}(\mathbf{y}) = \emptyset$ so $d(\mathbf{g},\Omega,\mathbf{y}) = 0$.

From the definition of degree, there exists $\mathbf{k} \in C^2\left(\overline{\Omega};\mathbb{R}^n\right)$ for which $\mathbf{y}$ is a regular point which is homotopic to $\mathbf{g}$ and $d(\mathbf{g},\Omega,\mathbf{y}) = d(\mathbf{k},\Omega,\mathbf{y})$. But the property of being homotopic is an equivalence relation and so from the definition of degree again, $d(\mathbf{k},\Omega,\mathbf{y}) = d(\mathbf{f},\Omega,\mathbf{y})$. This verifies the fifth property.

The sixth property follows from the fifth. Let the homotopy be
$$\mathbf{h}(\mathbf{x},t) \equiv \mathbf{f}(\mathbf{x}) + t(\mathbf{g}(\mathbf{x}) - \mathbf{f}(\mathbf{x}))$$
for $t \in [0,1]$. Then for $\mathbf{x} \in \partial\Omega$,
$$|\mathbf{h}(\mathbf{x},t) - \mathbf{y}| \geq |\mathbf{f}(\mathbf{x}) - \mathbf{y}| - t\,||\mathbf{g}-\mathbf{f}||_\infty > r - r = 0.$$

The seventh property follows from Lemma 16.15. The connected components are open connected sets. This lemma implies $d(\mathbf{f},\Omega,\cdot)$ is continuous. However, this function is also integer valued. If it is not constant on K, a connected component, there exists a number, r such that the values of this function are contained in $(-\infty,r) \cup (r,\infty)$ with the function having values in both of these disjoint open sets. But then we could consider the open sets $A \equiv \{\mathbf{z} \in K : d(\mathbf{f},\Omega,\mathbf{z}) > r\}$ and $B \equiv \{\mathbf{z} \in K : d(\mathbf{f},\Omega,\mathbf{z}) < r\}$. Now $K = A \cup B$ and we see that K is not connected.

The last property results from the homotopy
$$\mathbf{h}(\mathbf{x},t) = \mathbf{f}(\mathbf{x}) + t(\mathbf{g}(\mathbf{x}) - \mathbf{f}(\mathbf{x})).$$
Since $\mathbf{g} = \mathbf{f}$ on $\partial\Omega$, it follows that $\mathbf{h}(\partial\Omega,t)$ does not contain $\mathbf{y}$ and so the conclusion follows from property 5.

Definition 16.17 *We say that a bounded open set, Ω is symmetric if $-\Omega = \Omega$. We say a continuous function, $\mathbf{f} : \overline{\Omega} \to \mathbb{R}^n$ is odd if $\mathbf{f}(-\mathbf{x}) = -\mathbf{f}(\mathbf{x})$.*

Suppose Ω is symmetric and $\mathbf{g} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ is an odd map for which $\mathbf{0}$ is a regular point. Then the chain rule implies $D\mathbf{g}(-\mathbf{x}) = D\mathbf{g}(\mathbf{x})$ and so $d(\mathbf{g}, \Omega, \mathbf{0})$ must equal an odd integer because if $\mathbf{x} \in \mathbf{g}^{-1}(\mathbf{0})$, it follows that $-\mathbf{x} \in \mathbf{g}^{-1}(\mathbf{0})$ also and since $D\mathbf{g}(-\mathbf{x}) = D\mathbf{g}(\mathbf{x})$, it follows the overall contribution to the degree from $\mathbf{x}$ and $-\mathbf{x}$ must be an even integer. We also have $\mathbf{0} \in \mathbf{g}^{-1}(\mathbf{0})$ and so we have that the degree equals an even integer added to sign $(\det D\mathbf{g}(\mathbf{0}))$, an odd integer. It seems reasonable to expect that this would hold for an arbitrary continuous odd function defined on symmetric Ω. In fact this is the case and we will show this next. The following lemma is the key result used.

Lemma 16.18 *Let $\mathbf{g} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ be an odd map. Then for every $\varepsilon > 0$, there exists $\mathbf{h} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ such that $\mathbf{h}$ is also an odd map, $||\mathbf{h} - \mathbf{g}||_\infty < \varepsilon$, and $\mathbf{0}$ is a regular point of $\mathbf{h}$.*

Proof: Let $\mathbf{h}_0(\mathbf{x}) = \mathbf{g}(\mathbf{x}) + \delta\mathbf{x}$ where δ is chosen such that $\det D\mathbf{h}_0(\mathbf{0}) \neq 0$ and $\delta < \frac{\varepsilon}{2}$. Now let $\Omega_i \equiv \{\mathbf{x} \in \Omega : x_i \neq 0\}$. Define $\mathbf{h}_1(\mathbf{x}) \equiv \mathbf{h}_0(\mathbf{x}) - \mathbf{y}_1 x_1^3$ where $|\mathbf{y}_1| < \eta$ and $\mathbf{y}_1$ is a regular value of the function,

$$\mathbf{x} \to \frac{\mathbf{h}_0(\mathbf{x})}{x_1^3}$$

for $\mathbf{x} \in \Omega_1$. Thus $\mathbf{h}_1(\mathbf{x}) = \mathbf{0}$ if and only if $\mathbf{y}_1 = \frac{\mathbf{h}_0(\mathbf{x})}{x_1^3}$. Since $\mathbf{y}_1$ is a regular value,

$$\det\left(\frac{h_{0i,j}(\mathbf{x})\, x_1^3 - \frac{\partial}{\partial x_j}\left(x_1^3\right) h_{0i}(\mathbf{x})}{x_1^6}\right) =$$

$$\det\left(\frac{h_{0i,j}(\mathbf{x})\, x_1^3 - \frac{\partial}{\partial x_j}\left(x_1^3\right) y_{1i} x_1^3}{x_1^6}\right) \neq 0$$

implying that

$$\det\left(h_{0i,j}(\mathbf{x}) - \frac{\partial}{\partial x_j}\left(x_1^3\right) y_{1i}\right) = \det\left(D\mathbf{h}_1(\mathbf{x})\right) \neq 0.$$

We have shown that $\mathbf{0}$ is a regular value of $\mathbf{h}_1$ on the set Ω_1. Now we define $\mathbf{h}_2(\mathbf{x}) \equiv \mathbf{h}_1(\mathbf{x}) - \mathbf{y}_2 x_2^3$ where $|\mathbf{y}_2| < \eta$ and $\mathbf{y}_2$ is a regular value of

$$\mathbf{x} \to \frac{\mathbf{h}_1(\mathbf{x})}{x_2^3}$$

for $\mathbf{x} \in \Omega_2$. Thus, as in the step going from $\mathbf{h}_0$ to $\mathbf{h}_1$, for such $\mathbf{x} \in \mathbf{h}_2^{-1}(\mathbf{0})$,

$$\det\left(h_{1i,j}(\mathbf{x}) - \frac{\partial}{\partial x_j}\left(x_2^3\right) y_{2i}\right) = \det\left(D\mathbf{h}_2(\mathbf{x})\right) \neq 0.$$

Actually, $\det(D\mathbf{h}_2(\mathbf{x})) \neq 0$ for $\mathbf{x} \in (\Omega_1 \cup \Omega_2) \cap \mathbf{h}_2^{-1}(\mathbf{0})$ because if $\mathbf{x} \in (\Omega_1 \setminus \Omega_2) \cap \mathbf{h}_2^{-1}(\mathbf{0})$, then $x_2 = 0$. From the above formula for $\det(D\mathbf{h}_2(\mathbf{x}))$, we see that in this case,

$$\det\left(D\mathbf{h}_2(\mathbf{x})\right) = \det\left(D\mathbf{h}_1(\mathbf{x})\right) = \det\left(h_{1i,j}(\mathbf{x})\right).$$

We continue in this way, finding a sequence of odd functions, $\mathbf{h}_i$ where

$$\mathbf{h}_{i+1}(\mathbf{x}) = \mathbf{h}_i(\mathbf{x}) - \mathbf{y}_i x_i^3$$

for $|\mathbf{y}_i| < \eta$ and $\mathbf{0}$ a regular value of $\mathbf{h}_i$ on $\cup_{j=1}^i \Omega_j$. Let $\mathbf{h}_n \equiv \mathbf{h}$. Then $\mathbf{0}$ is a regular value of $\mathbf{h}$ for $\mathbf{x} \in \cup_{j=1}^n \Omega_j$. The point of Ω which is not in $\cup_{j=1}^n \Omega_j$ is $\mathbf{0}$. If $\mathbf{x} = \mathbf{0}$, then from the construction, $D\mathbf{h}(\mathbf{0}) = D\mathbf{h}_0(\mathbf{0})$ and so $\mathbf{0}$ is a regular value of $\mathbf{h}$ for $\mathbf{x} \in \Omega$. By choosing η small enough, we see that $||\mathbf{h} - \mathbf{g}||_\infty < \varepsilon$. This proves the lemma.

Theorem 16.19 *(Borsuk) Let* $\mathbf{f} \in C\left(\overline{\Omega};\mathbb{R}^n\right)$ *be odd and let* Ω *be symmetric with* $\mathbf{0} \notin \mathbf{f}(\partial\Omega)$*. Then* $d(\mathbf{f},\Omega,\mathbf{0})$ *equals an odd integer.*

Proof: Let δ be small enough that $B(\mathbf{0},3\delta) \cap \mathbf{f}(\partial\Omega) = \emptyset$. Let

$$\mathbf{g}_1 \in C^2\left(\overline{\Omega};\mathbb{R}^n\right)$$

be such that $||\mathbf{f}-\mathbf{g}_1||_\infty < \delta$ and let $\mathbf{g}$ denote the odd part of $\mathbf{g}_1$. Thus

$$\mathbf{g}(\mathbf{x}) \equiv \frac{1}{2}\left(\mathbf{g}_1(\mathbf{x}) - \mathbf{g}_1(-\mathbf{x})\right).$$

Then since $\mathbf{f}$ is odd, it follows that $||\mathbf{f}-\mathbf{g}||_\infty < \delta$ also. By Lemma 16.18 there exists odd $\mathbf{h} \in C^2\left(\overline{\Omega};\mathbb{R}^n\right)$ for which $\mathbf{0}$ is a regular value and $||\mathbf{h}-\mathbf{g}||_\infty < \delta$. Therefore, $||\mathbf{f}-\mathbf{h}||_\infty < 2\delta$ and from Theorem 16.16 $d(\mathbf{f},\Omega,\mathbf{0}) = d(\mathbf{h},\Omega,\mathbf{0})$. However, since $\mathbf{0}$ is a regular point of $\mathbf{h}$, $\mathbf{h}^{-1}(\mathbf{0}) = \{\mathbf{x}_i, -\mathbf{x}_i, \mathbf{0}\}_{i=1}^m$, and since $\mathbf{h}$ is odd, $D\mathbf{h}(-\mathbf{x}_i) = D\mathbf{h}(\mathbf{x}_i)$ and so $d(\mathbf{h},\Omega,\mathbf{0}) \equiv \sum_{i=1}^m \text{sign } \det(D\mathbf{h}(\mathbf{x}_i)) + \sum_{i=1}^m \text{sign } \det(D\mathbf{h}(-\mathbf{x}_i)) + \text{sign } \det(D\mathbf{h}(\mathbf{0}))$, an odd integer.

16.3 Applications

With these theorems it is possible to give easy proofs of some very important and difficult theorems.

Definition 16.20 *If* $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$*, we say* $\mathbf{f}$ *is locally one to one if for every* $\mathbf{x} \in U$*, there exists* $\delta > 0$ *such that* $\mathbf{f}$ *is one to one on* $B(\mathbf{x},\delta)$*.*

To begin with we consider the Invariance of domain theorem.

Theorem 16.21 *(Invariance of domain)Let* Ω *be any open subset of* $\mathbb{R}^n$ *and let* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *be continuous and locally one to one. Then* $\mathbf{f}$ *maps open subsets of* Ω *to open sets in* $\mathbb{R}^n$*.*

Proof: Suppose not. Then there exists an open set, $U \subseteq \Omega$ with U open but $\mathbf{f}(U)$ is not open . This means that there exists $\mathbf{y}_0$, not an interior point of $\mathbf{f}(U)$ where $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ for $\mathbf{x}_0 \in U$. Let $B(\mathbf{x}_0,r) \subseteq U$ be such that $\mathbf{f}$ is one to one on $\overline{B(\mathbf{x}_0,r)}$.

Let $\widetilde{\mathbf{f}}(\mathbf{x}) \equiv \mathbf{f}(\mathbf{x}+\mathbf{x}_0) - \mathbf{y}_0$. Then $\widetilde{\mathbf{f}} : B(\mathbf{0},r) \to \mathbb{R}^n, \widetilde{\mathbf{f}}(\mathbf{0}) = \mathbf{0}$. If $\mathbf{x} \in \partial B(\mathbf{0},r)$ and $t \in [0,1]$, then

$$\widetilde{\mathbf{f}}\left(\frac{\mathbf{x}}{1+t}\right) - \widetilde{\mathbf{f}}\left(\frac{-t\mathbf{x}}{1+t}\right) \neq 0 \tag{16.13}$$

because if this quantity were to equal zero, then since $\widetilde{\mathbf{f}}$ is one to one on $\overline{B(\mathbf{0},r)}$, it would follow that

$$\frac{\mathbf{x}}{1+t} = \frac{-t\mathbf{x}}{1+t}$$

which is not so unless $\mathbf{x} = \mathbf{0} \notin \partial B(\mathbf{0},r)$. Let $\mathbf{h}(\mathbf{x},t) \equiv \widetilde{\mathbf{f}}\left(\frac{\mathbf{x}}{1+t}\right) - \widetilde{\mathbf{f}}\left(\frac{-t\mathbf{x}}{1+t}\right)$. Then we just showed that $\mathbf{0} \notin \mathbf{h}(\partial\Omega,t)$ for all $t \in [0,1]$. By Borsuk's theorem, $d(\mathbf{h}(1,\cdot), B(\mathbf{0},r), \mathbf{0})$ equals an odd integer. Also by part 5 of Theorem 16.16, the homotopy invariance assertion,

$$d(\mathbf{h}(1,\cdot), B(\mathbf{0},r), \mathbf{0}) = d(\mathbf{h}(0,\cdot), B(\mathbf{0},r), \mathbf{0}) = d\left(\widetilde{\mathbf{f}}, B(\mathbf{0},r), \mathbf{0}\right).$$

Now from the case where $t = 0$ in (16.13), there exists $\delta > 0$ such that $B(\mathbf{0},\delta) \cap \widetilde{\mathbf{f}}(\partial B(\mathbf{0},r)) = \emptyset$. Therefore, $B(\mathbf{0},\delta)$ is a subset of a component of $\mathbb{R}^n \setminus \widetilde{\mathbf{f}}(\partial B(\mathbf{0},r))$ and so

$$d\left(\widetilde{\mathbf{f}}, B(\mathbf{0},r), \mathbf{z}\right) = d\left(\widetilde{\mathbf{f}}, B(\mathbf{0},r), \mathbf{0}\right) \neq 0$$

for all $\mathbf{z} \in B(\mathbf{0},\delta)$. It follows that for all $\mathbf{z} \in B(\mathbf{0},\delta)$, $\widetilde{\mathbf{f}}^{-1}(\mathbf{z}) \cap B(\mathbf{0},r) \neq \emptyset$. In terms of $\mathbf{f}$, this says that for all $|\mathbf{z}| < \delta$, there exists $\mathbf{x} \in B(\mathbf{0},r)$ such that

$$\mathbf{f}(\mathbf{x}+\mathbf{x}_0) - \mathbf{y}_0 = \mathbf{z}.$$

In other words, $\mathbf{f}(U) \supseteq \mathbf{f}(B(\mathbf{x}_0, r)) \supseteq B(\mathbf{y}_0, \delta)$ showing that $\mathbf{y}_0$ is an interior point of $\mathbf{f}(U)$ after all. This proves the theorem..

Corollary 16.22 *If $n > m$ there does not exist a continuous one to one map from $\mathbb{R}^n$ to $\mathbb{R}^m$.*

Proof: Suppose not and let $\mathbf{f}$ be such a continuous map, $\mathbf{f}(\mathbf{x}) \equiv (f_1(\mathbf{x}), \cdots, f_m(\mathbf{x}))^T$. Then let $\mathbf{g}(\mathbf{x}) \equiv (f_1(\mathbf{x}), \cdots, f_m(\mathbf{x}), 0, \cdots, 0)^T$ where there are $n-m$ zeros added in. Then $\mathbf{g}$ is a one to one continuous map from $\mathbb{R}^n$ to $\mathbb{R}^n$ and so $\mathbf{g}(\mathbb{R}^n)$ would have to be open from the invariance of domain theorem and this is not the case. This proves the corollary.

Corollary 16.23 *If $\mathbf{f}$ is locally one to one, $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$, and*

$$\lim_{|\mathbf{x}| \to \infty} |\mathbf{f}(\mathbf{x})| = \infty,$$

then $\mathbf{f}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

Proof: By the invariance of domain theorem, $\mathbf{f}(\mathbb{R}^n)$ is an open set. If $\mathbf{f}(\mathbf{x}_k) \to \mathbf{y}$, the growth condition ensures that $\{\mathbf{x}_k\}$ is a bounded sequence. Taking a subsequence which converges to $\mathbf{x} \in \mathbb{R}^n$ and using the continuity of $\mathbf{f}$, we see that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$. Thus $\mathbf{f}(\mathbb{R}^n)$ is both open and closed which implies $\mathbf{f}$ must be an onto map since otherwise, $\mathbb{R}^n$ would not be connected.

The next theorem is the famous Brouwer fixed point theorem.

Theorem 16.24 *(Brouwer fixed point) Let $B = \overline{B(\mathbf{0}, r)} \subseteq \mathbb{R}^n$ and let $\mathbf{f} : B \to B$ be continuous. Then there exists a point, $\mathbf{x} \in B$, such that $\mathbf{f}(\mathbf{x}) = \mathbf{x}$.*

Proof: Consider $\mathbf{h}(\mathbf{x}, t) \equiv t\mathbf{f}(\mathbf{x}) - \mathbf{x}$ for $t \in [0,1]$. Then if there is no fixed point in B for $\mathbf{f}$, it follows that $\mathbf{0} \notin \mathbf{h}(\partial B, t)$ for all t. Therefore, by the homotopy invariance,

$$0 = d(\mathbf{f} - id, B, \mathbf{0}) = d(-id, B, \mathbf{0}) = (-1)^n,$$

a contradiction.

Corollary 16.25 *(Brouwer fixed point) Let K be any convex compact set in $\mathbb{R}^n$ and let $\mathbf{f} : K \to K$ be continuous. Then $\mathbf{f}$ has a fixed point.*

Proof: Let $B \equiv B(\mathbf{0}, R)$ where R is large enough that $B \supseteq K$, and let P denote the projection map onto K. Let $\mathbf{g} : B \to B$ be defined as $\mathbf{g}(\mathbf{x}) \equiv \mathbf{f}(P(\mathbf{x}))$. Then $\mathbf{g}$ is continuous and so by Theorem 16.24 it has a fixed point, $\mathbf{x}$. But $\mathbf{x} \in K$ and so $\mathbf{x} = \mathbf{f}(P(\mathbf{x})) = \mathbf{f}(\mathbf{x})$.

Definition 16.26 *We say $\mathbf{f}$ is a retraction of $\overline{B(\mathbf{0}, r)}$ onto $\partial B(0, r)$ if $\mathbf{f}$ is continuous, $\mathbf{f}\left(\overline{B(\mathbf{0}, r)}\right) \subseteq \partial B(\mathbf{0},r)$, and $\mathbf{f}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial B(\mathbf{0},r)$.*

Theorem 16.27 *There does not exist a retraction of $\overline{B(\mathbf{0}, r)}$ onto $\partial B(\mathbf{0}, r)$.*

Proof: Suppose $\mathbf{f}$ were such a retraction. Then for all $\mathbf{x} \in \partial\Omega$, $\mathbf{f}(\mathbf{x}) = \mathbf{x}$ and so from the properties of the degree,

$$1 = d(id, \Omega, \mathbf{0}) = d(\mathbf{f}, \Omega, \mathbf{0})$$

which is clearly impossible because $\mathbf{f}^{-1}(\mathbf{0}) = \emptyset$ which implies $d(\mathbf{f}, \Omega, \mathbf{0}) = 0$.

In the next two theorems we make use of the Tietze extension theorem which states that in a metric space (more generally a normal topological space) every continuous function defined on a closed subset of the space having values in $[a, b]$ may be extended to a continuous function defined on the whole space having values in $[a, b]$. For a discussion of this important theorem and an outline of its proof see Problems 9 - 11 of Chapter 4.

Theorem 16.28 *Let Ω be a symmetric open set in $\mathbb{R}^n$ such that $\mathbf{0} \in \Omega$ and let $\mathbf{f} : \partial\Omega \to V$ be continuous where V is an m dimensional subspace of $\mathbb{R}^n, m < n$. Then $\mathbf{f}(-\mathbf{x}) = \mathbf{f}(\mathbf{x})$ for some $\mathbf{x} \in \partial\Omega$.*

Proof: Suppose not. Using the Tietze extension theorem, extend $\mathbf{f}$ to all of $\overline{\Omega}$, $\mathbf{f}\left(\overline{\Omega}\right) \subseteq V$. Let $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - \mathbf{f}(-\mathbf{x})$. Then $\mathbf{0} \notin \mathbf{g}(\partial\Omega)$ and so for some $r > 0$, $B(\mathbf{0},r) \subseteq \mathbb{R}^n \setminus \mathbf{g}(\partial\Omega)$. For $\mathbf{z} \in B(\mathbf{0},r)$,

$$d(\mathbf{g}, \Omega, \mathbf{z}) = d(\mathbf{g}, \Omega, \mathbf{0}) \neq 0$$

because $B(\mathbf{0},r)$ is contained in a component of $\mathbb{R}^n \setminus \mathbf{g}(\partial\Omega)$ and Borsuk's theorem implies that $d(\mathbf{g}, \Omega, \mathbf{0}) \neq 0$ since $\mathbf{g}$ is odd. Hence

$$V \supseteq \mathbf{g}(\Omega) \supseteq B(\mathbf{0},r)$$

and this is a contradiction because V is m dimensional. This proves the theorem.

This theorem is called the Borsuk Ulam theorem. Note that it implies that there exist two points on opposite sides of the surface of the earth that have the same atmospheric pressure and temperature. The next theorem is an amusing result which is like combing hair. It gives the existence of a "cowlick".

Theorem 16.29 *Let n be odd and let Ω be an open bounded set in $\mathbb{R}^n$ with $\mathbf{0} \in \Omega$. Suppose $\mathbf{f} : \partial\Omega \to \mathbb{R}^n \setminus \{\mathbf{0}\}$ is continuous. Then for some $\mathbf{x} \in \partial\Omega$ and $\lambda \neq 0$, $\mathbf{f}(\mathbf{x}) = \lambda\mathbf{x}$.*

Proof: Using the Tietze extension theorem, extend $\mathbf{f}$ to all of $\overline{\Omega}$. Suppose for all $\mathbf{x} \in \partial\Omega$, $\mathbf{f}(\mathbf{x}) \neq \lambda\mathbf{x}$ for all $\lambda \in \mathbb{R}$. Then

$$\mathbf{0} \notin t\mathbf{f}(\mathbf{x}) + (1-t)\mathbf{x},\ (\mathbf{x}, t) \in \partial\Omega \times [0, 1]$$

$$\mathbf{0} \notin t\mathbf{f}(\mathbf{x}) - (1-t)\mathbf{x},\ (\mathbf{x}, t) \in \partial\Omega \times [0, 1].$$

Then by the homotopy invariance of degree,

$$d(\mathbf{f}, \Omega, \mathbf{0}) = d(id, \Omega, \mathbf{0}),\ d(\mathbf{f}, \Omega, \mathbf{0}) = d(-id, \Omega, \mathbf{0}).$$

But this is impossible because $d(id, \Omega, \mathbf{0}) = 1$ but $d(-id, \Omega, \mathbf{0}) = (-1)^n$. This proves the theorem.

16.4 The Product formula and Jordan separation theorem

In this section we present the product formula for the degree and use it to prove a very important theorem in topology. To begin with we give the following lemma.

Lemma 16.30 *Let $\mathbf{y}_1, \cdots, \mathbf{y}_r$ be points not in $\mathbf{f}(\partial\Omega)$. Then there exists $\widetilde{\mathbf{f}} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ such that $\left|\left|\widetilde{\mathbf{f}} - \mathbf{f}\right|\right|_\infty < \delta$ and $\mathbf{y}_i$ is a regular value for $\widetilde{\mathbf{f}}$ for each i.*

Proof: Let $\mathbf{f}_0 \in C^2\left(\overline{\Omega};\mathbb{R}^n\right)$, $||\mathbf{f}_0-\mathbf{f}||_\infty < \frac{\delta}{2}$. For S_0 the singular set of $\mathbf{f}_0$, pick $\widetilde{\mathbf{y}}_1$ such that $\widetilde{\mathbf{y}}_1 \notin \mathbf{f}_0(S_0)$.($\widetilde{\mathbf{y}}_1$ is a regular value of $\mathbf{f}_0$) and $|\widetilde{\mathbf{y}}_1-\mathbf{y}_1| < \frac{\delta}{3r}$. Let $\mathbf{f}_1(\mathbf{x}) \equiv \mathbf{f}_0(\mathbf{x})+\mathbf{y}_1-\widetilde{\mathbf{y}}_1$. Thus $\mathbf{y}_1$ is a regular value of $\mathbf{f}_1$ and

$$||\mathbf{f}-\mathbf{f}_1||_\infty \leq ||\mathbf{f}-\mathbf{f}_0||_\infty + ||\mathbf{f}_0-\mathbf{f}_1||_\infty < \frac{\delta}{3r}+\frac{\delta}{2}.$$

Letting S_1 be the singular set of $\mathbf{f}_1$, choose $\widetilde{\mathbf{y}}_2$ such that $|\widetilde{\mathbf{y}}_2-\mathbf{y}_2| < \frac{\delta}{3r}$ and

$$\widetilde{\mathbf{y}}_2 \notin \mathbf{f}_1(S_1) \cup (\mathbf{f}_1(S_1)+\mathbf{y}_2-\mathbf{y}_1).$$

Let $\mathbf{f}_2(\mathbf{x}) \equiv \mathbf{f}_2(\mathbf{x})+\mathbf{y}_2-\widetilde{\mathbf{y}}_2$. Thus if $\mathbf{f}_2(\mathbf{x}) = \mathbf{y}_1$, then

$$\mathbf{f}_1(\mathbf{x})+\mathbf{y}_2-\mathbf{y}_1 = \widetilde{\mathbf{y}}_2$$

and so $\mathbf{x} \notin S_1$. Thus $\mathbf{y}_1$ is a regular value of $\mathbf{f}_2$. If $\mathbf{f}_2(\mathbf{x}) = \mathbf{y}_2$, then

$$\mathbf{y}_2 = \mathbf{f}_1(\mathbf{x})+\mathbf{y}_2-\widetilde{\mathbf{y}}_2$$

and so $\mathbf{f}_1(\mathbf{x}) = \widetilde{\mathbf{y}}_2$ implying $\mathbf{x} \notin S_1$ showing $\det D\mathbf{f}_2(\mathbf{x}) = \det D\mathbf{f}_1(\mathbf{x}) \neq 0$. Thus $\mathbf{y}_1$ and $\mathbf{y}_2$ are both regular values of $\mathbf{f}_2$ and

$$||\mathbf{f}_2-\mathbf{f}||_\infty \leq ||\mathbf{f}_2-\mathbf{f}_1||_\infty + ||\mathbf{f}_1-\mathbf{f}||$$

$$< \frac{\delta}{3r}+\frac{\delta}{3r}+\frac{\delta}{2}.$$

We continue in this way. Let $\widetilde{\mathbf{f}} \equiv \mathbf{f}_r$. Then $\left|\left|\widetilde{\mathbf{f}}-\mathbf{f}\right|\right|_\infty < \frac{\delta}{2}+\frac{\delta}{3} < \delta$ and each $\mathbf{y}_i$ is a regular value of $\widetilde{\mathbf{f}}$.

Definition 16.31 *Let the connected components of $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$ be denoted by K_i. We know from the properties of the degree that $d(\mathbf{f},\Omega,\cdot)$ is constant on each of these components. We will denote by $d(\mathbf{f},\Omega,K_i)$ the constant value on the component, K_i.*

Lemma 16.32 *Let $\mathbf{f} \in C\left(\overline{\Omega};\mathbb{R}^n\right), \mathbf{g} \in C^2(\mathbb{R}^n,\mathbb{R}^n)$, and $\mathbf{y} \notin \mathbf{g}(\mathbf{f}(\partial\Omega))$. Suppose also that $\mathbf{y}$ is a regular value of $\mathbf{g}$. Then we have the following product formula where K_i are the bounded components of $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$.*

$$d(\mathbf{g}\circ\mathbf{f},\Omega,\mathbf{y}) = \sum_{i=1}^{\infty} d(\mathbf{f},\Omega,K_i)\, d(\mathbf{g},K_i,\mathbf{y}).$$

Proof: First note that if K_i is unbounded, $d(\mathbf{f},\Omega,K_i) = 0$ because there exists a point, $\mathbf{z} \in K_i$ such that $\mathbf{f}^{-1}(\mathbf{z}) = \emptyset$ due to the fact that $\mathbf{f}\left(\overline{\Omega}\right)$ is compact and is consequently bounded. Thus it makes no difference in the above formula whether we let K_i be general components or insist on their being bounded. Let $\left\{\mathbf{x}_j^i\right\}_{j=1}^{k_i}$ denote the point of $\mathbf{g}^{-1}(\mathbf{y})$ which are contained in K_i, the *ith* component of $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$. Note also that $\mathbf{g}^{-1}(\mathbf{y}) \cap \mathbf{f}\left(\overline{\Omega}\right)$ is a compact set covered by the components of $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$ and so it is covered by finitely many of these components. For the other components, $d(\mathbf{f},\Omega,K_i) = 0$ and so this is actually a finite sum. There are no convergence problems. Now let $\varepsilon > 0$ be small enough that

$$B(\mathbf{y},3\varepsilon) \cap \mathbf{g}(\mathbf{f}(\partial\Omega)) = \emptyset,$$

and for each $\mathbf{x}_j^i \in \mathbf{g}^{-1}(\mathbf{y})$

$$B\left(\mathbf{x}_j^i,3\varepsilon\right) \cap \mathbf{f}(\partial\Omega) = \emptyset.$$

Next choose $\delta > 0$ small enough that $\delta < \varepsilon$ and if $\mathbf{z}_1$ and $\mathbf{z}_2$ are any two points of $\mathbf{f}\left(\overline{\Omega}\right)$ with $|\mathbf{z}_1 - \mathbf{z}_2| < \delta$, it follows that $|\mathbf{g}(\mathbf{z}_1) - \mathbf{g}(\mathbf{z}_2)| < \varepsilon$.

Now choose $\widetilde{\mathbf{f}} \in C^2\left(\overline{\Omega}; \mathbb{R}^n\right)$ such that $\left|\left|\widetilde{\mathbf{f}} - \mathbf{f}\right|\right|_\infty < \delta$ and each point, $\mathbf{x}_j^i$ is a regular value of $\widetilde{\mathbf{f}}$. From the properties of the degree we know that $d(\mathbf{f}, \Omega, K_i) = d\left(\mathbf{f}, \Omega, \mathbf{x}_j^i\right)$ for each $j = 1, \cdot\cdot\cdot, m_i$. For $\mathbf{x} \in \partial\Omega$, and $t \in [0,1]$,

$$\left|\mathbf{f}(\mathbf{x}) + t\left(\widetilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\right) - \mathbf{x}_j^i\right| > 3\varepsilon - t\varepsilon > 0$$

and so

$$d\left(\widetilde{\mathbf{f}}, \Omega, \mathbf{x}_j^i\right) = d\left(\mathbf{f}, \Omega, \mathbf{x}_j^i\right) = d(\mathbf{f}, \Omega, K_i) \tag{16.14}$$

independent of j. Also for $\mathbf{x} \in \partial\Omega$, and $t \in [0,1]$,

$$\left|\mathbf{g}(\mathbf{f}(\mathbf{x})) + t\left(\mathbf{g}\left(\widetilde{\mathbf{f}}(\mathbf{x})\right) - \mathbf{g}(\mathbf{f}(\mathbf{x}))\right) - \mathbf{y}\right| \geq 3\varepsilon - t\varepsilon > 0$$

and so

$$d(\mathbf{g} \circ \mathbf{f}, \Omega, \mathbf{y}) = d\left(\mathbf{g}\circ\widetilde{\mathbf{f}}, \Omega, \mathbf{y}\right). \tag{16.15}$$

Now let $\left\{\mathbf{u}_l^{ij}\right\}_{l=1}^{k_{ij}}$ be the points of $\widetilde{\mathbf{f}}^{-1}\left(\mathbf{x}_j^i\right)$. Therefore, $k_{ij} < \infty$ because $\widetilde{\mathbf{f}}^{-1}\left(\mathbf{x}_j^i\right) \subseteq \Omega$, a bounded open set. It follows from (16.15) that

$$d(\mathbf{g} \circ \mathbf{f}, \Omega, \mathbf{y}) = d\left(\mathbf{g}\circ\widetilde{\mathbf{f}}, \Omega, \mathbf{y}\right)$$

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{m_i}\sum_{l=1}^{k_{ij}} sgn \det D\mathbf{g}\left(\widetilde{\mathbf{f}}\left(\mathbf{u}_l^{ij}\right)\right) sgn \det D\widetilde{\mathbf{f}}\left(\mathbf{u}_l^{ij}\right)$$

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{m_i} \det D\mathbf{g}\left(\mathbf{x}_j^i\right) d\left(\widetilde{\mathbf{f}}, \Omega, \mathbf{x}_j^i\right) = \sum_{i=1}^{\infty} d(\mathbf{g}, K_i, \mathbf{y})\, d\left(\widetilde{\mathbf{f}}, \Omega, \mathbf{x}_j^i\right)$$

$$= \sum_{i=1}^{\infty} d(\mathbf{g}, K_i, \mathbf{y})\, d(\mathbf{f}, \Omega, K_i).$$

With this lemma we are ready to prove the product formula.

Theorem 16.33 *(product formula) Let* $\{K_i\}_{i=1}^{\infty}$ *be the bounded components of* $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$ *for* $\mathbf{f} \in C\left(\overline{\Omega}; \mathbb{R}^n\right)$, *let* $\mathbf{g} \in C(\mathbb{R}^n, \mathbb{R}^n)$, *and suppose that* $\mathbf{y} \notin \mathbf{g}(\mathbf{f}(\partial\Omega))$. *Then*

$$d(\mathbf{g} \circ \mathbf{f}, \Omega, \mathbf{y}) = \sum_{i=1}^{\infty} d(\mathbf{g}, K_i, \mathbf{y})\, d(\mathbf{f}, \Omega, K_i). \tag{16.16}$$

Proof: Let $B(\mathbf{y}, 3\delta) \cap \mathbf{g}(\mathbf{f}(\partial\Omega)) = \emptyset$ and let $\widetilde{\mathbf{g}} \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ be such that

$$\sup\left\{|\widetilde{\mathbf{g}}(\mathbf{z}) - \mathbf{g}(\mathbf{z})| : \mathbf{z} \in \mathbf{f}\left(\overline{\Omega}\right)\right\} < \delta$$

And also $\mathbf{y}$ is a regular value of $\widetilde{\mathbf{g}}$. Then from the above inequality, if $\mathbf{x} \in \partial\Omega$ and $t \in [0,1]$,

$$|\mathbf{g}(\mathbf{f}(\mathbf{x})) + t(\widetilde{\mathbf{g}}(\mathbf{f}(\mathbf{x})) - \mathbf{g}(\mathbf{f}(\mathbf{x}))) - \mathbf{y}| \geq 3\delta - t\delta > 0.$$

It follows that

$$d(\mathbf{g} \circ \mathbf{f}, \Omega, \mathbf{y}) = d(\widetilde{\mathbf{g}} \circ \mathbf{f}, \Omega, \mathbf{y}). \tag{16.17}$$

Now also, $\partial K_i \subseteq \mathbf{f}(\partial\Omega)$ and so if $\mathbf{z} \in \partial K_i$, then $\mathbf{g}(\mathbf{z}) \in \mathbf{g}(\mathbf{f}(\partial\Omega))$. Consequently, for such $\mathbf{z}$,

$$|\mathbf{g}(\mathbf{z}) + t(\widetilde{\mathbf{g}}(\mathbf{z}) - \mathbf{g}(\mathbf{z})) - \mathbf{y}| \geq |\mathbf{g}(\mathbf{z}) - \mathbf{y}| - t\delta > 3\delta - t\delta > 0$$

which shows that

$$d(\mathbf{g}, K_i, \mathbf{y}) = d(\widetilde{\mathbf{g}}, K_i, \mathbf{y}). \tag{16.18}$$

Therefore, by Lemma 16.32,

$$d(\mathbf{g} \circ \mathbf{f}, \Omega, \mathbf{y}) = d(\widetilde{\mathbf{g}} \circ \mathbf{f}, \Omega, \mathbf{y}) = \sum_{i=1}^{\infty} d(\widetilde{\mathbf{g}}, K_i, \mathbf{y})\, d(\mathbf{f}, \Omega, K_i)$$

$$= \sum_{i=1}^{\infty} d(\mathbf{g}, K_i, \mathbf{y})\, d(\mathbf{f}, \Omega, K_i).$$

This proves the product formula. Note there are no convergence problems because these sums are actually finite sums because, as in the previous lemma, $\mathbf{g}^{-1}(\mathbf{y}) \cap \mathbf{f}\left(\overline{\Omega}\right)$ is a compact set covered by the components of $\mathbb{R}^n \setminus \mathbf{f}(\partial\Omega)$ and so it is covered by finitely many of these components. For the other components, $d(\mathbf{f}, \Omega, K_i) = 0$.

With the product formula is possible to give a fairly easy proof of the Jordan separation theorem, a very profound result in the topology of $\mathbb{R}^n$.

Theorem 16.34 *(Jordan separation theorem) Let $\mathbf{f}$ be a homeomorphism of C and $\mathbf{f}(C)$ where C is a compact set in $\mathbb{R}^n$. Then $\mathbb{R}^n \setminus C$ and $\mathbb{R}^n \setminus \mathbf{f}(C)$ have the same number of connected components.*

Proof: Denote by $\mathcal{K}$ the bounded components of $\mathbb{R}^n \setminus C$ and denote by $\mathcal{L}$, the bounded components of $\mathbb{R}^n \setminus \mathbf{f}(C)$. Also let $\overline{\mathbf{f}}$ be an extension of $\mathbf{f}$ to all of $\mathbb{R}^n$ and let $\overline{\mathbf{f}^{-1}}$ denote an extension of $\mathbf{f}^{-1}$ to all of $\mathbb{R}^n$. Pick $K \in \mathcal{K}$ and take $\mathbf{y} \in K$. Let $\mathcal{H}$ denote the set of bounded components of $\mathbb{R}^n \setminus \mathbf{f}(\partial K)$ (note $\partial K \subseteq C$). Since $\overline{\mathbf{f}^{-1}} \circ \overline{\mathbf{f}}$ equals the identity, id, on ∂K, it follows that

$$1 = d(id, K, \mathbf{y}) = d\left(\overline{\mathbf{f}^{-1}} \circ \overline{\mathbf{f}}, K, \mathbf{y}\right).$$

By the product formula,

$$1 = d\left(\overline{\mathbf{f}^{-1}} \circ \overline{\mathbf{f}}, K, \mathbf{y}\right) = \sum_{H \in \mathcal{H}} d\left(\overline{\mathbf{f}}, K, H\right) d\left(\overline{\mathbf{f}^{-1}}, H, \mathbf{y}\right),$$

the sum being a finite sum. Now letting $\mathbf{x} \in L \in \mathcal{L}$, if S is a connected set containing $\mathbf{x}$ and contained in $\mathbb{R}^n \setminus \mathbf{f}(C)$, then it follows S is contained in $\mathbb{R}^n \setminus \mathbf{f}(\partial K)$ because $\partial K \subseteq C$. Therefore, every set of $\mathcal{L}$ is contained in some set of $\mathcal{H}$. Letting $\mathcal{G}_H$ denote those sets of $\mathcal{L}$ which are contained in H, we note that

$$\overline{H} \setminus \cup \mathcal{G}_H \subseteq \mathbf{f}(C).$$

This is because if $\mathbf{z} \notin \cup\mathcal{G}_H$, then $\mathbf{z}$ cannot be contained in any set of $\mathcal{L}$ which has nonempty intersection with H since then, that whole set of $\mathcal{L}$ would be contained in H due to the fact that the sets of $\mathcal{H}$ are disjoint open set and the sets of $\mathcal{L}$ are connected. It follows that $\mathbf{z}$ is either an element of $\mathbf{f}(C)$ which would correspond to being contained in none of the sets of $\mathcal{L}$, or else, $\mathbf{z}$ is contained in some set of $\mathcal{L}$ which has empty intersection with H. But the sets of $\mathcal{L}$ are open and so this point, $\mathbf{z}$ cannot, in this latter case, be contained in $\overline{H}$. Therefore, the above inclusion is verified.

Claim: $\mathbf{y} \notin \overline{\mathbf{f}^{-1}}\left(\overline{H} \setminus \cup\mathcal{G}_H\right)$.

Proof of the claim: If not, then $\overline{\mathbf{f}^{-1}}(\mathbf{z}) = \mathbf{y}$ where $\mathbf{z} \in \overline{H} \setminus \cup\mathcal{G}_H \subseteq \mathbf{f}(C)$ and so $\mathbf{f}^{-1}(\mathbf{z}) = \mathbf{y} \in C$. But $\mathbf{y} \notin C$ and this contradiction proves the claim.

Now every set of $\mathcal{L}$ is contained in some set of $\mathcal{H}$. What about those sets of H which contain no set of $\mathcal{L}$? From the claim, $\mathbf{y} \notin \overline{\mathbf{f}^{-1}}\left(\overline{H}\right)$ and so $d\left(\overline{\mathbf{f}^{-1}}, H, \mathbf{y}\right) = 0$. Therefore, letting $\mathcal{H}_1$ denote those sets of $\mathcal{H}$ which contain some set of $\mathcal{L}$, properties of the degree imply

$$1 = \sum_{H \in \mathcal{H}_1} d\left(\overline{\mathbf{f}}, K, H\right) d\left(\overline{\mathbf{f}^{-1}}, H, \mathbf{y}\right) = \sum_{H \in \mathcal{H}_1} d\left(\overline{\mathbf{f}}, K, H\right) \sum_{L \in \mathcal{G}_H} d\left(\overline{\mathbf{f}^{-1}}, L, \mathbf{y}\right)$$

$$= \sum_{H \in \mathcal{H}_1} \sum_{L \in \mathcal{G}_H} d\left(\overline{\mathbf{f}}, K, L\right) d\left(\overline{\mathbf{f}^{-1}}, L, \mathbf{y}\right) = \sum_{L \in \mathcal{L}} d\left(\overline{\mathbf{f}}, K, L\right) d\left(\overline{\mathbf{f}^{-1}}, L, \mathbf{y}\right)$$

$$= \sum_{L \in \mathcal{L}} d\left(\overline{\mathbf{f}}, K, L\right) d\left(\overline{\mathbf{f}^{-1}}, L, \mathbf{y}\right) = \sum_{L \in \mathcal{L}} d\left(\overline{\mathbf{f}}, K, L\right) d\left(\overline{\mathbf{f}^{-1}}, L, K\right)$$

and each sum is finite. Letting $|\mathcal{K}|$ denote the number of elements in $\mathcal{K}$,

$$|\mathcal{K}| = \sum_{K \in \mathcal{K}} \sum_{L \in \mathcal{L}} d\left(\overline{\mathbf{f}}, K, L\right) d\left(\overline{\mathbf{f}^{-1}}, L, K\right).$$

By symmetry, we may use the above argument to write

$$|\mathcal{L}| = \sum_{L \in \mathcal{L}} \sum_{K \in \mathcal{K}} d\left(\overline{\mathbf{f}}, K, L\right) d\left(\overline{\mathbf{f}^{-1}}, L, K\right).$$

It follows $|\mathcal{K}| = |\mathcal{L}|$ and this proves the theorem because if $n > 1$ there is exactly one unbounded component and if $n = 1$ there are exactly two.

16.5 Integration and the degree

There is a very interesting application of the degree to integration. Recall Lemma 16.12. We will use Theorem 16.16 to generalize this lemma next. In this proposition, we let ϕ_ε be the mollifier of Definition 16.11.

Proposition 16.35 *Let* $\mathbf{g} \in \mathcal{U}_{\mathbf{y}} \cap C^1\left(\overline{\Omega}; \mathbb{R}^n\right)$ *then whenever* $\varepsilon > 0$ *is small enough,*

$$d(\mathbf{g}, \Omega, \mathbf{y}) = \int_\Omega \phi_\varepsilon(\mathbf{g}(\mathbf{x}) - \mathbf{y}) \det D\mathbf{g}(\mathbf{x})\, dx.$$

Proof: Let $\varepsilon_0 > 0$ be small enough that

$$B(\mathbf{y}, 3\varepsilon_0) \cap \mathbf{g}(\partial\Omega) = \emptyset.$$

Now let ψ_m be a mollifier and let

$$\mathbf{g}_m \equiv \mathbf{g}*\psi_m.$$

Thus $\mathbf{g}_m \in C^\infty\left(\overline{\Omega};\mathbb{R}^n\right)$ and

$$||\mathbf{g}_m - \mathbf{g}||_\infty, ||D\mathbf{g}_m - D\mathbf{g}||_\infty \to 0 \tag{16.19}$$

as $m \to \infty$. Choose M such that for $m \geq M$,

$$||\mathbf{g}_m - \mathbf{g}||_\infty < \varepsilon_0. \tag{16.20}$$

Thus $\mathbf{g}_m \in \mathcal{U}_{\mathbf{y}} \cap C^2\left(\overline{\Omega};\mathbb{R}^n\right)$

Letting $\mathbf{z} \in B(\mathbf{y},\varepsilon)$ for $\varepsilon < \varepsilon_0$, and $\mathbf{x} \in \partial\Omega$,

$$|(1-t)\,\mathbf{g}_m(\mathbf{x}) + \mathbf{g}_k(\mathbf{x})\,t - \mathbf{z}| \geq (1-t)\,|\mathbf{g}_m(\mathbf{x}) - \mathbf{z}| + t\,|\mathbf{g}_k(\mathbf{x}) - \mathbf{z}|$$

$$\begin{aligned} &> (1-t)\,|\mathbf{g}(\mathbf{x}) - \mathbf{z}| + t\,|\mathbf{g}(\mathbf{x}) - \mathbf{z}| - \varepsilon_0 \\ &= |\mathbf{g}(\mathbf{x}) - \mathbf{z}| - \varepsilon_0 \\ &\geq |\mathbf{g}(\mathbf{x}) - \mathbf{y}| - |\mathbf{y} - \mathbf{z}| - \varepsilon_0 \\ &> 3\varepsilon_0 - \varepsilon_0 - \varepsilon_0 = \varepsilon_0 > 0 \end{aligned}$$

showing that $B(\mathbf{y},\varepsilon) \cap ((1-t)\,\mathbf{g}_m + t\mathbf{g}_k)(\partial\Omega) = \emptyset$. By Lemma 16.13

$$\int_\Omega \phi_\varepsilon(\mathbf{g}_m(\mathbf{x}) - \mathbf{y}) \det(D\mathbf{g}_m(\mathbf{x}))\,dx =$$

$$\int_\Omega \phi_\varepsilon(\mathbf{g}_k(\mathbf{x}) - \mathbf{y}) \det(D\mathbf{g}_k(\mathbf{x}))\,dx \tag{16.21}$$

for all $k, m \geq M$.

We may assume that $\mathbf{y}$ is a regular value of $\mathbf{g}_m$ since if it is not, we will simply replace $\mathbf{g}_m$ with $\widetilde{\mathbf{g}}_m$ defined by

$$\mathbf{g}_m(\mathbf{x}) \equiv \widetilde{\mathbf{g}}_m(\mathbf{x}) - (\mathbf{y} - \widetilde{\mathbf{y}})$$

where $\widetilde{\mathbf{y}}$ is a regular value of $\mathbf{g}$ chosen close enough to $\mathbf{y}$ such that (16.20) holds for $\widetilde{\mathbf{g}}_m$ in place of $\mathbf{g}_m$. Thus $\widetilde{\mathbf{g}}_m(\mathbf{x}) = \mathbf{y}$ if and only if $\mathbf{g}_m(\mathbf{x}) = \widetilde{\mathbf{y}}$ and so $\mathbf{y}$ is a regular value of $\widetilde{\mathbf{g}}_m$. Therefore,

$$d(\mathbf{y},\Omega,\mathbf{g}_m) = \int_\Omega \phi_\varepsilon(\mathbf{g}_m(\mathbf{x}) - \mathbf{y}) \det(D\mathbf{g}_m(\mathbf{x}))\,dx \tag{16.22}$$

for all ε small enough by Lemma 16.12. For $\mathbf{x} \in \partial\Omega$, and $t \in [0,1]$,

$$\begin{aligned} |(1-t)\,\mathbf{g}(\mathbf{x}) + t\mathbf{g}_m(\mathbf{x}) - \mathbf{y}| &\geq (1-t)\,|\mathbf{g}(\mathbf{x}) - \mathbf{y}| + t\,|\mathbf{g}_m(\mathbf{x}) - \mathbf{y}| \\ &\geq (1-t)\,|\mathbf{g}(\mathbf{x}) - \mathbf{y}| + t\,|\mathbf{g}(x) - \mathbf{y}| - \varepsilon_0 \\ &> 3\varepsilon_0 - \varepsilon_0 > 0 \end{aligned}$$

and so by Theorem 16.16, and (16.22), we can write

$$d(\mathbf{y},\Omega,\mathbf{g}) = d(\mathbf{y},\Omega,\mathbf{g}_m) =$$

$$\int_{\Omega} \phi_{\varepsilon} (\mathbf{g}_m (\mathbf{x}) - \mathbf{y}) \det (D\mathbf{g}_m (\mathbf{x})) \, dx$$

whenever ε is small enough. Fix such an $\varepsilon < \varepsilon_0$ and use (16.21) to conclude the right side of the above equations is independent of $m > M$. Then let $m \to \infty$ and use (16.19) to take a limit as $m \to \infty$ and conclude

$$\begin{aligned} d (\mathbf{y},\Omega, \mathbf{g}) &= \lim_{m \to \infty} \int_{\Omega} \phi_{\varepsilon} (\mathbf{g}_m (\mathbf{x}) - \mathbf{y}) \det (D\mathbf{g}_m (\mathbf{x})) \, dx \\ &= \int_{\Omega} \phi_{\varepsilon} (\mathbf{g} (\mathbf{x}) - \mathbf{y}) \det (D\mathbf{g} (\mathbf{x})) \, dx. \end{aligned}$$

This proves the proposition.

With this proposition, we are ready to present the interesting change of variables theorem. Let U be a bounded open set with the property that ∂U has measure zero and let $\mathbf{f} \in C^1 \left(\overline{U}; \mathbb{R}^n\right)$. Then Theorem 11.13 implies that $\mathbf{f} (\partial U)$ also has measure zero. From Proposition 16.35 we see that for $\mathbf{y} \notin \mathbf{f} (\partial U)$,

$$d (\mathbf{y}, U, \mathbf{f}) = \lim_{\varepsilon \to 0} \int_U \phi_{\varepsilon} (\mathbf{f} (\mathbf{x}) - \mathbf{y}) \det D\mathbf{f} (\mathbf{x}) \, dx,$$

showing that $\mathbf{y} \to d (\mathbf{y}, U, \mathbf{f})$ is a measurable function. Also,

$$\mathbf{y} \to \int_U \phi_{\varepsilon} (\mathbf{f} (\mathbf{x}) - \mathbf{y}) \det D\mathbf{f} (\mathbf{x}) \, dx \tag{16.23}$$

is a function bounded independent of ε because $\det D\mathbf{f} (\mathbf{x})$ is bounded and the integral of ϕ_{ε} equals one. Letting $h \in C_c (\mathbb{R}^n)$, we can therefore, apply the dominated convergence theorem and the above observation that $\mathbf{f} (\partial U)$ has measure zero to write

$$\int h (\mathbf{y}) \, d (\mathbf{y}, U, \mathbf{f}) \, dy = \lim_{\varepsilon \to 0} \int h (\mathbf{y}) \int_U \phi_{\varepsilon} (\mathbf{f} (\mathbf{x}) - \mathbf{y}) \det D\mathbf{f} (\mathbf{x}) \, dx dy.$$

Now we will assume for convenience that ϕ_{ε} has the following simple form.

$$\phi_{\varepsilon} (\mathbf{x}) \equiv \frac{1}{\varepsilon^n} \phi \left(\frac{\mathbf{x}}{\varepsilon} \right)$$

where the support of ϕ is contained in $B (\mathbf{0},1)$, $\phi (\mathbf{x}) \geq 0$, and $\int \phi (\mathbf{x}) \, dx = 1$. Therefore, interchanging the order of integration in the above,

$$\begin{aligned} \int h (\mathbf{y}) \, d (\mathbf{y}, U, \mathbf{f}) \, dy &= \lim_{\varepsilon \to 0} \int_U \det D\mathbf{f} (\mathbf{x}) \int h (\mathbf{y}) \, \phi_{\varepsilon} (\mathbf{f} (\mathbf{x}) - \mathbf{y}) \, dy dx \\ &= \lim_{\varepsilon \to 0} \int_U \det D\mathbf{f} (\mathbf{x}) \int_{B(\mathbf{0},1)} h (\mathbf{f} (\mathbf{x}) - \varepsilon \mathbf{u}) \, \phi (\mathbf{u}) \, du dx. \end{aligned}$$

Now

$$\left| \int_U \det D\mathbf{f} (\mathbf{x}) \int_{B(\mathbf{0},1)} h (\mathbf{f} (\mathbf{x}) - \varepsilon \mathbf{u}) \, \phi (\mathbf{u}) \, du dx - \int_U \det D\mathbf{f} (\mathbf{x}) \, h (\mathbf{f} (\mathbf{x})) \, dx \right| =$$

$$\left| \int_U \det D\mathbf{f} (\mathbf{x}) \int_{B(\mathbf{0},1)} h (\mathbf{f} (\mathbf{x}) - \varepsilon \mathbf{u}) \, \phi (\mathbf{u}) \, du dx - \right.$$

$$\left. \int_U \det D\mathbf{f}(\mathbf{x}) \int_{B(\mathbf{0},1)} h(\mathbf{f}(\mathbf{x})) \phi(\mathbf{u})\, du dx \right| \leq$$

$$\left| \int_{B(\mathbf{0},1)} \int_U \det D\mathbf{f}(\mathbf{x}) \left| h(\mathbf{f}(\mathbf{x}) - \varepsilon \mathbf{u}) - h(\mathbf{f}(\mathbf{x})) \right| dx \phi(\mathbf{u})\, du \right|$$

By the uniform continuity of h we see this converges to zero as $\varepsilon \to 0$. Consequently,

$$\begin{aligned} \int h(\mathbf{y})\, d(\mathbf{y}, U, \mathbf{f})\, dy &= \lim_{\varepsilon \to 0} \int_U \det D\mathbf{f}(\mathbf{x}) \int h(\mathbf{y}) \phi_\varepsilon(\mathbf{f}(\mathbf{x}) - \mathbf{y})\, dy dx \\ &= \int_U \det D\mathbf{f}(\mathbf{x})\, h(\mathbf{f}(\mathbf{x}))\, dx \end{aligned}$$

which proves the following lemma.

Lemma 16.36 *Let $h \in C_c(\mathbb{R}^n)$ and let $\mathbf{f} \in C^1(\overline{U}; \mathbb{R}^n)$ where ∂U has measure zero for U a bounded open set. Then everything is measurable which needs to be and we have the following formula.*

$$\int h(\mathbf{y})\, d(\mathbf{y}, U, \mathbf{f})\, dy = \int_U \det D\mathbf{f}(\mathbf{x})\, h(\mathbf{f}(\mathbf{x}))\, dx.$$

Next we give a simple corollary which replaces $C_c(\mathbb{R}^n)$ with $L^1(\mathbb{R}^n)$.

Corollary 16.37 *Let $h \in L^\infty(\mathbb{R}^n)$ and let $\mathbf{f} \in C^1(\overline{U}; \mathbb{R}^n)$ where ∂U has measure zero for U a bounded open set. Then everything is measurable which needs to be and we have the following formula.*

$$\int h(\mathbf{y})\, d(\mathbf{y}, U, \mathbf{f})\, dy = \int_U \det D\mathbf{f}(\mathbf{x})\, h(\mathbf{f}(\mathbf{x}))\, dx.$$

Proof: For all $\mathbf{y} \notin \mathbf{f}(\partial U)$ a set of measure zero, $d(\mathbf{y}, U, \mathbf{f})$ is bounded by some constant which is independent of $\mathbf{y} \notin U$ due to boundedness of the formula (16.23). The integrand of the integral on the left equals zero off some bounded set because if $\mathbf{y} \notin \mathbf{f}(U), d(\mathbf{y}, U, \mathbf{f}) = 0$. Therefore, we can modify h off a bounded set and assume without loss of generality that $h \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$. Letting h_k be a sequence of functions of $C_c(\mathbb{R}^n)$ which converges pointwise a.e. to h and in $L^1(\mathbb{R}^n)$, in such a way that $|h_k(\mathbf{y})| \leq ||h||_\infty + 1$ for all $\mathbf{x}$,

$$|\det D\mathbf{f}(\mathbf{x})\, h_k(\mathbf{f}(\mathbf{x}))| \leq |\det D\mathbf{f}(\mathbf{x})| \left(||h||_\infty + 1\right)$$

and

$$\int_U |\det D\mathbf{f}(\mathbf{x})| \left(||h||_\infty + 1\right) dx < \infty$$

because U is bounded and $h \in L^\infty(\mathbb{R}^n)$. Therefore, we may apply the dominated convergence theorem to the equation

$$\int h_k(\mathbf{y})\, d(\mathbf{y}, U, \mathbf{f})\, dy = \int_U \det D\mathbf{f}(\mathbf{x})\, h_k(\mathbf{f}(\mathbf{x}))\, dx$$

and obtain the desired result.

Differential forms

17.1 Manifolds

Manifolds are sets which resemble $\mathbb{R}^n$ locally. To make this more precise, we need some definitions.

Definition 17.1 *Let $X \subseteq Y$ where (Y, τ) is a topological space. Then we define the relative topology on X to be the set of all intersections with X of sets from τ. We say these relatively open sets are open in X. Similarly, we say a subset of X is closed in X if it is closed with respect to the relative topology on X.*

We leave as an easy exercise the following lemma.

Lemma 17.2 *Let X and Y be defined as in Definition 17.1. Then the relative topology defined there is a topology for X. Furthermore, the sets closed in X consist of the intersections of closed sets from Y with X.*

With the above lemma and definition, we are ready to define manifolds.

Definition 17.3 *A closed and bounded subset of $\mathbb{R}^m$, Ω, will be called an n dimensional manifold with boundary if there are finitely many sets, U_i, open in Ω and continuous one to one functions, $\mathbf{R}_i : U_i \to \mathbf{R}_i U_i \subseteq \mathbb{R}^n$ such that $\mathbf{R}_i$ and $\mathbf{R}_i^{-1}$ both are continuous, $\mathbf{R}_i U_i$ is open in $\mathbb{R}^n_{\leq} \equiv \left\{\mathbf{u} \in \mathbb{R}^n : u^1 \leq 0\right\}$, and $\Omega = \cup_{i=1}^p U_i$. These mappings, $\mathbf{R}_i$, together with their domains, U_i, are called charts and the totality of all the charts, $(U_i, \mathbf{R}_i)$ just described is called an atlas for the manifold. We also define $int\,(\Omega) \equiv \{\mathbf{x} \in \Omega : \text{ for some } i, \mathbf{R}_i \mathbf{x} \in \mathbb{R}^n_<\}$ where $\mathbb{R}^n_< \equiv \left\{\mathbf{u} \in \mathbb{R}^n : u^1 < 0\right\}$. We define $\partial\Omega \equiv \{\mathbf{x} \in \Omega : \text{ for some } i, \mathbf{R}_i \mathbf{x} \in \mathbb{R}^n_0\}$ where $\mathbb{R}^n_0 \equiv \left\{\mathbf{u} \in \mathbb{R}^n : u^1 = 0\right\}$ and we refer to $\partial\Omega$ as the boundary of Ω.*

This definition is a little too restrictive. In general we do not require the collection of sets, U_i to be finite. However, in the case where Ω is closed and bounded, we can always reduce to this because of the compactness of Ω and since this is the case of most interest to us here, the assumption that the collection of sets, U_i, is finite is made.

Lemma 17.4 *Let $\partial\Omega$ and $int\,(\Omega)$ be as defined above. Then $int\,(\Omega)$ is open in Ω and $\partial\Omega$ is closed in Ω. Furthermore, $\partial\Omega \cap int\,(\Omega) = \emptyset$, $\Omega = \partial\Omega \cup int\,(\Omega)$, and $\partial\Omega$ is an $n-1$ dimensional manifold for which $\partial\,(\partial\Omega) = \emptyset$. In addition to this, the property of being in $int\,(\Omega)$ or $\partial\Omega$ does not depend on the choice of atlas.*

Proof: It is clear that $\Omega = \partial\Omega \cup int\,(\Omega)$. We show that $\partial\Omega \cap int\,(\Omega) = \emptyset$. Suppose this does not happen. Then there would exist $\mathbf{x} \in \partial\Omega \cap int\,(\Omega)$. Therefore, there would exist two mappings $\mathbf{R}_i$ and $\mathbf{R}_j$ such that $\mathbf{R}_j \mathbf{x} \in \mathbb{R}^n_0$ and $\mathbf{R}_i \mathbf{x} \in \mathbb{R}^n_<$ with $\mathbf{x} \in U_i \cap U_j$. Now consider the map, $\mathbf{R}_j \circ \mathbf{R}_i^{-1}$, a continuous one to one map from $\mathbb{R}^n_{\leq}$ to $\mathbb{R}^n_{\leq}$ having a continuous inverse. Choosing $r > 0$ small enough, we may obtain that

$$\mathbf{R}_i^{-1} B\,(\mathbf{R}_i \mathbf{x}, r) \subseteq U_i \cap U_j.$$

Therefore, $\mathbf{R}_j \circ \mathbf{R}_i^{-1}\,(B\,(\mathbf{R}_i \mathbf{x}, r)) \subseteq \mathbb{R}^n_{\leq}$ and contains a point on $\mathbb{R}^n_0$. However, this cannot occur because it contradicts the theorem on invariance of domain, Theorem 16.21, which requires that $\mathbf{R}_j \circ \mathbf{R}_i^{-1}\,(B\,(\mathbf{R}_i \mathbf{x}, r))$

must be an open subset of $\mathbb{R}^n$. Therefore, we have shown that $\partial\Omega \cap int(\Omega) = \emptyset$ as claimed. This same argument shows that the property of being in $int(\Omega)$ or $\partial\Omega$ does not depend on the choice of the atlas. To verify that $\partial(\partial\Omega) = \emptyset$, let $\mathbf{P}_1 : \mathbb{R}^n \to \mathbb{R}^{n-1}$ be defined by $\mathbf{P}_1(u_1, \cdots, u_n) = (u_2, \cdots, u_n)$ and consider the maps $\mathbf{P}_1\mathbf{R}_i - k\mathbf{e}_1$ where k is large enough that the images of these maps are in $\mathbb{R}^{n-1}_<$. Here $\mathbf{e}_1$ refers to $\mathbb{R}^{n-1}$.

We now show that $int(\Omega)$ is open in Ω. If $\mathbf{x} \in int(\Omega)$, then for some i, $\mathbf{R}_i\mathbf{x} \in \mathbb{R}^n_<$ and so whenever, $r > 0$ is small enough, $B(\mathbf{R}_i\mathbf{x}, r) \subseteq \mathbb{R}^n_<$ and $\mathbf{R}_i^{-1}(B(\mathbf{R}_i\mathbf{x}, r))$ is a set open in Ω and contained in U_i. Therefore, all the points of $\mathbf{R}_i^{-1}(B(\mathbf{R}_i\mathbf{x}, r))$ are in $int(\Omega)$ which shows that $int(\Omega)$ is open in Ω as claimed. Now it follows that $\partial\Omega$ is closed because $\partial\Omega = \Omega \setminus int(\Omega)$.

Definition 17.5 *Let $V \subseteq \mathbb{R}^n$. We denote by $C^k\left(\overline{V}; \mathbb{R}^m\right)$ the set of functions which are restrictions to V of some function defined on $\mathbb{R}^n$ which has k continuous derivatives and compact support.*

Definition 17.6 *We will say an n dimensional manifold with boundary, Ω is a C^k manifold with boundary for some $k \geq 1$ if $\mathbf{R}_j \circ \mathbf{R}_i^{-1} \in C^k\left(\overline{\mathbf{R}_i(U_i \cap U_j)}; \mathbb{R}^n\right)$ and $\mathbf{R}_i^{-1} \in C^k\left(\overline{\mathbf{R}_i U_i}; \mathbb{R}^m\right)$. We say Ω is orientable if in addition to this there exists an atlas, $(U_r, \mathbf{R}_r)$, such that whenever $U_i \cap U_j \neq \emptyset$,*

$$\det\left(D\left(\mathbf{R}_j \circ \mathbf{R}_i^{-1}\right)\right)(\mathbf{u}) > 0 \tag{17.1}$$

whenever $\mathbf{u} \in \mathbf{R}_i(U_i \cap U_j)$. The mappings, $\mathbf{R}_i \circ \mathbf{R}_j^{-1}$ are called the overlap maps. We will refer to an atlas satisfying (17.1) as an oriented atlas. We will also assume that if an oriented n manifold has nonempty boundary, then $n \geq 2$. Thus we are not defining the concept of an oriented one manifold with boundary.

The following lemma is immediate from the definitions.

Lemma 17.7 *If Ω is a C^k oriented manifold with boundary, then $\partial\Omega$ is also a C^k oriented manifold with empty boundary.*

Proof: We simply let an atlas consist of $(V_r, \mathbf{S}_r)$ where V_r is the intersection of U_r with $\partial\Omega$ and $\mathbf{S}_r$ is of the form

$$\begin{aligned}\mathbf{S}_r(\mathbf{x}) &\equiv P_1\mathbf{R}_r(\mathbf{x}) - k_r\mathbf{e}_1 \\ &= (u_2 - k_r, \cdots, u_n)\end{aligned}$$

where P_1 is defined above in Lemma 17.4, $\mathbf{e}_1$ refers to $\mathbb{R}^{n-1}$ and k_r is large enough that for all $\mathbf{x} \in V_r$, $\mathbf{S}_r(\mathbf{x}) \in \mathbb{R}^{n-1}_<$.

When we refer to an oriented manifold, Ω, we will always regard $\partial\Omega$ as an oriented manifold according to the construction of Lemma 17.7.

The study of manifolds is really a generalization of something with which everyone who has taken a normal calculus course is familiar. We think of a point in three dimensional space in two ways. There is a geometric point and there are coordinates associated with this point. Remember, there are lots of different coordinate systems which describe a point. There are spherical coordinates, cylindrical coordinates and rectangular coordinates to name the three most popular coordinate systems. These coordinates are like the vector $\mathbf{u}$. The point, $\mathbf{x}$ is like the geometric point although we are always assuming $\mathbf{x}$ has rectangular coordinates in $\mathbb{R}^m$ for some m. Under fairly general conditions, it has been shown there is no loss of generality in making such an assumption and so we are doing so.

17.2 The integration of differential forms on manifolds

In this section we consider the integration of differential forms on manifolds. This topic is a generalization of what you did in calculus when you found the work done by a force field on an object which moves over some path. There you evaluated line integrals. Differential forms are just a generalization of this idea and it turns out they are what it makes sense to integrate on manifolds. The following lemma, used in establishing the definition of the degree and proved in that chapter is also the fundamental result in discussing the integration of differential forms.

Lemma 17.8 *Let $\mathbf{g} : U \to V$ be C^2 where U and V are open subsets of $\mathbb{R}^n$. Then*

$$\sum_{j=1}^{n} (cof(D\mathbf{g}))_{ij,j} = 0,$$

where here $(D\mathbf{g})_{ij} \equiv g_{i,j} \equiv \frac{\partial g_i}{\partial x_j}$.

We will also need the following fundamental lemma on partitions of unity which is also discussed earlier, Corollary 12.24.

Lemma 17.9 *Let K be a compact set in $\mathbb{R}^n$ and let $\{U_i\}_{i=1}^{\infty}$ be an open cover of K. Then there exists functions, $\psi_k \in C_c^{\infty}(U_i)$ such that $\psi_i \prec U_i$ and*

$$\sum_{i=1}^{\infty} \psi_i(\mathbf{x}) = 1.$$

The following lemma will also be used.

Lemma 17.10 *Let $\{i_1, \cdots, i_n\} \subseteq \{1, \cdots, m\}$ and let $\mathbf{R} \in C^1\left(\overline{V}; \mathbb{R}^m\right)$. Letting $\mathbf{x} = \mathbf{R}\mathbf{u}$, we define*

$$\frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(u^1 \cdots u^n\right)}$$

to be the following determinant.

$$\det\begin{pmatrix} \frac{\partial x^{i_1}}{\partial u^1} & \cdots & \frac{\partial x^{i_1}}{\partial u^n} \\ \vdots & & \vdots \\ \frac{\partial x^{i_n}}{\partial u^1} & \cdots & \frac{\partial x^{i_n}}{\partial u^n} \end{pmatrix}.$$

Then letting $\mathbf{R}_1 \in C^1\left(\overline{W}; \mathbb{R}^m\right)$ and $\mathbf{x} = \mathbf{R}_1\mathbf{v} = \mathbf{R}\mathbf{u}$, we have the following formula.

$$\frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(v^1 \cdots v^n\right)} = \frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(u^1 \cdots u^n\right)} \frac{\partial\left(u^1 \cdots u^n\right)}{\partial\left(v^1 \cdots v^n\right)}.$$

Proof: We define for $I \equiv \{i_1, \cdots, i_n\}$, the mapping $\mathbf{P}_I : \mathbb{R}^m \to span\left(\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_n}\right)$ by

$$\mathbf{P}_I\mathbf{x} \equiv \begin{pmatrix} x^{i_1} \\ \vdots \\ x^{i_n} \end{pmatrix}$$

Thus

$$\frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(u^1 \cdots u^n\right)} = \det\left(D\left(\mathbf{P}_I\mathbf{R}\right)(\mathbf{u})\right).$$

since $\mathbf{R}_1(\mathbf{v}) = \mathbf{R}(\mathbf{u}) = \mathbf{x}$,

$$\mathbf{P}_I\mathbf{R}_1(\mathbf{v}) = \mathbf{P}_I\mathbf{R}(\mathbf{u})$$

and so the chain rule implies

$$D\left(\mathbf{P}_I\mathbf{R}_1\right)(\mathbf{v}) = D\left(\mathbf{P}_I\mathbf{R}\right)(\mathbf{u})\, D\left(\mathbf{R}^{-1} \circ \mathbf{R}_1\right)(\mathbf{v})$$

and so

$$\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(v^1\cdots v^n\right)}=\det\left(D\left(\mathbf{P}_I\mathbf{R}_1\right)(\mathbf{v})\right)=$$

$$\det\left(D\left(\mathbf{P}_I\mathbf{R}\right)(\mathbf{u})\right)\det\left(D\left(\mathbf{R}^{-1}\circ\mathbf{R}_1\right)(\mathbf{v})\right)=$$

$$\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)}\frac{\partial\left(u^1\cdots u^n\right)}{\partial\left(v^1\cdots v^n\right)}$$

as claimed.

With these three lemmas, we first define what a differential form is and then describe how to integrate one.

Definition 17.11 *We will let I denote an ordered list of n indices from the set, $\{1,\cdots,m\}$. Thus $I=(i_1,\cdots,i_n)$. We say it is an ordered list because the order matters. Thus if we switch two indices, I would be changed. A differential form of order n in $\mathbb{R}^m$ is a formal expression,*

$$\omega=\sum_I a_I(\mathbf{x})\,d\mathbf{x}^I$$

where a_I is at least Borel measurable or continuous if you wish $d\mathbf{x}^I$ is short for the expression

$$dx^{i_1}\wedge\cdots\wedge dx^{i_n},$$

and the sum is taken over all ordered lists of indices taken from the set, $\{1,\cdots,m\}$. For Ω an orientable n dimensional manifold with boundary, we define

$$\int_\Omega\omega \tag{17.2}$$

according to the following procedure. We let $(U_i,\mathbf{R}_i)$ be an oriented atlas for Ω. Each U_i is the intersection of an open set in $\mathbb{R}^m$, with Ω and so there exists a C^∞ partition of unity subordinate to the open cover, $\{O_i\}$ which sums to 1 on Ω. Thus $\psi_i\in C_c^\infty(O_i)$, has values in $[0,1]$ and satisfies $\sum_i\psi_i(\mathbf{x})=1$ for all $\mathbf{x}\in\Omega$. We call this a partition of unity subordinate to $\{U_i\}$ in this context. Then we define (17.2) by

$$\int_\Omega\omega\equiv\sum_{i=1}^p\sum_I\int_{\mathbf{R}_iU_i}\psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right)a_I\left(\mathbf{R}_i^{-1}(\mathbf{u})\right)\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)}du \tag{17.3}$$

Of course there are all sorts of questions related to whether this definition is well defined. The formula (17.2) makes no mention of partitions of unity or a particular atlas. What if we picked a different atlas and a different partition of unity? Would we get the same number for $\int_\Omega\omega$? In general, the answer is no. However, there is a sense in which (17.2) is well defined. This involves the concept of orientation.

Definition 17.12 *Suppose Ω is an n dimensional C^k orientable manifold with boundary and let $(U_i,\mathbf{R}_i)$ and $(V_i,\mathbf{S}_i)$ be two oriented atlass of Ω. We say they have the same orientation if whenever $U_i\cap V_j\neq\emptyset$,*

$$\det\left(D\left(\mathbf{R}_i\circ\mathbf{S}_j^{-1}\right)(\mathbf{v})\right)>0 \tag{17.4}$$

for all $\mathbf{v}\in\mathbf{S}_j(U_i\cap V_j)$.

Note that by the chain rule, (17.4) is equivalent to saying $\det\left(D\left(\mathbf{S}_j\circ\mathbf{R}_i^{-1}\right)(\mathbf{u})\right)>0$ for all $\mathbf{u}\in\mathbf{R}_i(U_i\cap V_j)$.

Now we are ready to discuss the manner in which (17.2) is well defined.

Theorem 17.13 *Suppose Ω is an n dimensional C^k orientable manifold with boundary and let $(U_i, \mathbf{R}_i)$ and $(V_i, \mathbf{S}_i)$ be two oriented atlass of Ω. Suppose the two atlass have the same orientation. Then if $\int_\Omega \omega$ is computed with respect to the two atlass the same number is obtained.*

Proof: In Definition 17.11 let $\{\psi_i\}$ be a partition of unity as described there which is associated with the atlas $(U_i, \mathbf{R}_i)$ and let $\{\eta_i\}$ be a partition of unity associated in the same manner with the atlas $(V_i, \mathbf{S}_i)$. Then since the orientations are the same, letting $\mathbf{u} = \left(\mathbf{R}_i \circ \mathbf{S}_j^{-1}\right) \mathbf{v}$,

$$\det \left(D \left(\mathbf{R}_i \circ \mathbf{S}_j^{-1}\right)(\mathbf{v})\right) \equiv \frac{\partial \left(u^1 \cdots u^n\right)}{\partial \left(v^1 \cdots v^n\right)} > 0$$

and so using the change of variables formula,

$$\sum_I \int_{\mathbf{R}_i U_i} \psi_i \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) a_I \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(u^1 \cdots u^n\right)} du = \tag{17.5}$$

$$\sum_{j=1}^{q} \sum_I \int_{\mathbf{R}_i (U_i \cap V_j)} \eta_j \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) \psi_i \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) a_I \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(u^1 \cdots u^n\right)} du =$$

$$\sum_{j=1}^{q} \sum_I \int_{\mathbf{S}_j (U_i \cap V_j)} \eta_j \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \psi_i \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) a_I \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(u^1 \cdots u^n\right)} \frac{\partial \left(u^1 \cdots u^n\right)}{\partial \left(v^1 \cdots v^n\right)} dv$$

which by Lemma 17.10 equals

$$= \sum_{j=1}^{q} \sum_I \int_{\mathbf{S}_j (U_i \cap V_j)} \eta_j \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \psi_i \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) a_I \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(v^1 \cdots v^n\right)} dv. \tag{17.6}$$

We sum over i in (17.6) and (17.5) to obtain

$$\text{the definition of } \int \omega \text{ using } (U_i, \mathbf{R}_i) \equiv$$

$$\sum_{i=1}^{p} \sum_I \int_{\mathbf{R}_i U_i} \psi_i \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) a_I \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(u^1 \cdots u^n\right)} du =$$

$$\sum_{i=1}^{p} \sum_{j=1}^{q} \sum_I \int_{\mathbf{S}_j (U_i \cap V_j)} \eta_j \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \psi_i \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) a_I \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(v^1 \cdots v^n\right)} dv$$

$$= \sum_{j=1}^{q} \sum_I \int_{\mathbf{S}_j (U_i \cap V_j)} \eta_j \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) a_I \left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \frac{\partial \left(x^{i_1} \cdots x^{i_n}\right)}{\partial \left(v^1 \cdots v^n\right)} dv =$$

$$\text{the definition of } \int \omega \text{ using } (V_i, \mathbf{S}_i).$$

This proves the theorem.

17.3 Some examples of orientable manifolds

We show in this section that there are lots of orientable manifolds. The following simple proposition will give abundant examples.

Proposition 17.14 *Suppose Ω is a bounded open subset of $\mathbb{R}^n$ with $n \geq 2$ having the property that for all $\mathbf{p} \in \partial\Omega \equiv \overline{\Omega} \setminus \Omega$, there exists an open set, $\widetilde{U}$, containing $\mathbf{p}$, an open interval, (a,b), an open set, $B \subseteq \mathbb{R}^{n-1}$, and a function, $g \in C^k\left(\overline{B};\mathbb{R}\right)$ such that for some $k \in \{1,\cdots,n\}$*

$$x^k = g\left(\widehat{\mathbf{x}_k}\right)$$

whenever $\mathbf{x} \in \partial\Omega \cap \widetilde{U}$, and $\Omega \cap \widetilde{U}$ equals either

$$\left\{\mathbf{x} \in \mathbb{R}^n : \widehat{\mathbf{x}_k} \in B \text{ and } x^k \in \left(a, g\left(\widehat{\mathbf{x}_k}\right)\right)\right\} \tag{17.7}$$

or

$$\left\{\mathbf{x} \in \mathbb{R}^n : \widehat{\mathbf{x}_k} \in B \text{ and } x^k \in \left(g\left(\widehat{\mathbf{x}_k}\right), b\right)\right\}. \tag{17.8}$$

Then Ω is an orientable C^k manifold with boundary. Here

$$\widehat{\mathbf{x}_k} \equiv \left(x^1,\cdots,x^{k-1}x^{k+1},\cdots,x^n\right)^T.$$

Proof: Let $\widetilde{U}$ and g be as described above. In the case of (17.7) define

$$\mathbf{R}\left(\mathbf{x}\right) \equiv \begin{pmatrix} x^k - g\left(\widehat{\mathbf{x}_k}\right) & -x^2 & \cdots & x^1 & \cdots & x^n \end{pmatrix}^T$$

where the x^1 is in the k^{th} slot. Then it is a simple exercise to verify that $\det\left(D\mathbf{R}\left(\mathbf{x}\right)\right) = 1$. Now in case (17.8) holds, we let

$$\mathbf{R}\left(\mathbf{x}\right) \equiv \begin{pmatrix} g\left(\widehat{\mathbf{x}_k}\right) - x^k & x^2 & \cdots & x^1 & \cdots & x^n \end{pmatrix}^T$$

We see that in this case we also have $\det\left(D\mathbf{R}\left(\mathbf{x}\right)\right) = 1$.

Also, in either case we see that $\mathbf{R}$ is one to one and k times continuously differentiable mapping into $\mathbb{R}^n_{\leq}$. In case (17.7) the inverse is given by

$$\mathbf{R}^{-1}\left(\mathbf{u}\right) = \begin{pmatrix} u^1 & -u^2 & \cdots & u^k + g\left(\widehat{\mathbf{u}_k}\right) & \cdots & u^n \end{pmatrix}^T$$

and in case (17.8), there is also a simple formula for $\mathbf{R}^{-1}$. Now modifying these functions outside of suitable compact sets, we may assume they are all of the sort needed in the definition of a C^k manifold.

The set, $\partial\Omega$ is compact and so there are p of these sets, $\widetilde{U_j}$ covering $\partial\Omega$ along with functions $\mathbf{R}_j$ as just described. Let U_0 satisfy

$$\Omega \setminus \cup_{i=1}^p U_i \subseteq U_0 \subseteq \overline{U_0} \subseteq \Omega$$

and let $\mathbf{R}_0\left(\mathbf{x}\right) \equiv \begin{pmatrix} x^1 - k & x^2 & \cdots & x^n \end{pmatrix}$ where k is chosen large enough that $\mathbf{R}_0$ maps U_0 into $\mathbb{R}^n_{<}$. Modifying this function off some compact set containing $\overline{U_0}$, to equal zero off this set, we see $(U_r, \mathbf{R}_r)$ is an oriented atlas for Ω if we define $U_r \equiv \widetilde{U_r} \cap \Omega$. The chain rule shows the derivatives of the overlap maps have positive determinants.

For example, a ball of radius $r > 0$ is an oriented n manifold with boundary because it satisfies the conditions of the above proposition. This proposition gives examples of n manifolds in $\mathbb{R}^n$ but we want to have examples of n manifolds in $\mathbb{R}^m$ for $m > n$. The following lemma will help.

Lemma 17.15 *Suppose O is a bounded open subset of $\mathbb{R}^n$ and let $\mathbf{F} : O \to \mathbb{R}^m$ be a function in $C^k\left(\overline{O};\mathbb{R}^m\right)$ where $m \geq n$ with the property that for all $\mathbf{x} \in O$, $D\mathbf{F}(\mathbf{x})$ has rank n. Then if $\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0)$, there exists a bounded open set in $\mathbb{R}^m, W$, which contains $\mathbf{y}_0$, a bounded open set, $U \subseteq O$ which contains $\mathbf{x}_0$ and a function $\mathbf{G} : W \to U$ such that $\mathbf{G}$ is in $C^k\left(\overline{W};\mathbb{R}^n\right)$ and for all $\mathbf{x} \in U$,*

$$\mathbf{G}(\mathbf{F}(\mathbf{x})) = \mathbf{x}.$$

Furthermore, $\mathbf{G} = \mathbf{G}_1 \circ \mathbf{P}$ on W where $\mathbf{P}$ is a map of the form

$$\mathbf{P}(\mathbf{y}) = \left(y^{i_1}, \cdots, y^{i_n}\right)$$

for some list of indices, $i_1 < \cdots < i_n$.

Proof: Consider the system

$$\mathbf{F}(\mathbf{x}) = \mathbf{y}.$$

Since $D\mathbf{F}(\mathbf{x}_0)$ has rank n, the inverse function theorem can be applied to some system of the form

$$F^{i_j}(\mathbf{x}) = y^{i_j}, j = 1, \cdots, n$$

to obtain open sets, U_1 and V_1, subsets of $\mathbb{R}^n$ and a function, $\mathbf{G}_1 \in C^k(V_1; U_1)$ such that $\mathbf{G}_1$ is one to one and onto and is the inverse of the function $\mathbf{F}^I$ where $\mathbf{F}^I(\mathbf{x})$ is the vector valued function whose j^{th} component is $F^{i_j}(\mathbf{x})$. If we restrict the open sets, U_1 and V_1, calling the restricted open sets, U and V respectively, we can modify $\mathbf{G}_1$ off a compact set to obtain $\mathbf{G}_1 \in C^k\left(\overline{V};\mathbb{R}^n\right)$ and $\mathbf{G}_1$ is the inverse of $\mathbf{F}^I$. Now let $\mathbf{P}$ and $\mathbf{G}$ be as defined above and let $W = V \times Z$ where Z is a bounded open set in $\mathbb{R}^{m-n}$ such that W contains $\mathbf{F}(U)$. (If $n = m$, we let $W = V$.) We can modify $\mathbf{P}$ off a compact set which contains W so the resulting function, still denoted by $\mathbf{P}$ is in $C^k\left(\overline{W};\mathbb{R}^n\right)$ Then for $\mathbf{x} \in U$

$$\mathbf{G}(\mathbf{F}(\mathbf{x})) = \mathbf{G}_1(\mathbf{P}(\mathbf{F}(\mathbf{x}))) = \mathbf{G}_1\left(\mathbf{F}^I(\mathbf{x})\right) = \mathbf{x}.$$

This proves the lemma.

With this lemma we can give a theorem which will provide many other examples.

Theorem 17.16 *Let Ω be an n manifold with boundary in $\mathbb{R}^n$ and suppose $\Omega \subseteq O$, an open bounded subset of $\mathbb{R}^n$. Suppose $\mathbf{F} \in C^k\left(\overline{O};\mathbb{R}^m\right)$ is one to one on O and $D\mathbf{F}(\mathbf{x})$ has rank n for all $\mathbf{x} \in O$. Then $\mathbf{F}(\Omega)$ is also a manifold with boundary and $\partial\mathbf{F}(\Omega) = \mathbf{F}(\partial\Omega)$. If Ω is a C^l manifold for $l \leq k$, then so is $\mathbf{F}(\Omega)$. If Ω is orientable, then so is $\mathbf{F}(\Omega)$.*

Proof: Let $(U_r, \mathbf{R}_r)$ be an atlas for Ω and suppose $U_r = O_r \cap \Omega$ where O_r is an open subset of O. Let $\mathbf{x}_0 \in U_r$. By Lemma 17.15 there exists an open set, $W_{\mathbf{x}_0}$ in $\mathbb{R}^m$ containing $\mathbf{F}(\mathbf{x}_0)$, an open set in $\mathbb{R}^n, \widetilde{U_{\mathbf{x}_0}}$ containing $\mathbf{x}_0$, and $\mathbf{G}_{\mathbf{x}_0} \in C^k\left(\overline{W_{\mathbf{x}_0}};\mathbb{R}^n\right)$ such that

$$\mathbf{G}_{\mathbf{x}_0}(\mathbf{F}(\mathbf{x})) = \mathbf{x}$$

for all $\mathbf{x} \in \widetilde{U_{\mathbf{x}_0}}$. Let $U_{\mathbf{x}_0} \equiv U_r \cap \widetilde{U_{\mathbf{x}_0}}$.

Claim: $\mathbf{F}(U_{\mathbf{x}_0})$ is open in $\mathbf{F}(\Omega)$.

Proof: Let $\mathbf{x} \in U_{\mathbf{x}_0}$. If $\mathbf{F}(\mathbf{x}_1)$ is close enough to $\mathbf{F}(\mathbf{x})$ where $\mathbf{x}_1 \in \Omega$, then $\mathbf{F}(\mathbf{x}_1) \in W_{\mathbf{x}_0}$ and so

$$\begin{aligned} |\mathbf{x} - \mathbf{x}_1| &= |\mathbf{G}_{\mathbf{x}_0}(\mathbf{F}(\mathbf{x})) - \mathbf{G}_{\mathbf{x}_0}(\mathbf{F}(\mathbf{x}_1))| \\ &\leq K|(\mathbf{F}(\mathbf{x})) - \mathbf{F}(\mathbf{x}_1)| \end{aligned}$$

where K is some constant which depends only on

$$\max\{||D\mathbf{G}_{\mathbf{x}_0}(\mathbf{y})|| : \mathbf{y} \in \mathbb{R}^m\}.$$

Therefore, if $\mathbf{F}(\mathbf{x}_1)$ is close enough to $\mathbf{F}(\mathbf{x})$, it follows we can conclude $|\mathbf{x}-\mathbf{x}_1|$ is very small. Since $U_{\mathbf{x}_0}$ is open in Ω it follows that whenever $\mathbf{F}(\mathbf{x}_1)$ is sufficiently close to $\mathbf{F}(\mathbf{x})$, we have $\mathbf{x}_1 \in U_{\mathbf{x}_0}$. Consequently $\mathbf{F}(\mathbf{x}_1) \in \mathbf{F}(U_{\mathbf{x}_0})$. This shows $\mathbf{F}(U_{\mathbf{x}_0})$ is open in $\mathbf{F}(\Omega)$ and proves the claim.

With this claim it follows that $(\mathbf{F}(U_{\mathbf{x}_0}), \mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_0})$ is a chart. The inverse map of $\mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_0}$ being $\mathbf{F} \circ \mathbf{R}_r^{-1}$. Since Ω is compact there are finitely many of these sets, $\mathbf{F}(U_{\mathbf{x}_i})$ covering Ω. This yields an atlas for $\mathbf{F}(\Omega)$ of the form $(\mathbf{F}(U_{\mathbf{x}_i}), \mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_i})$ where $\mathbf{x}_i \in U_r$ and proves the first part. If the $\mathbf{R}_r^{-1}$ are in $C^l\left(\overline{\mathbf{R}_r U_r}; \mathbb{R}^n\right)$, then the overlap map for two of these charts is of the form,

$$\left(\mathbf{R}_s \circ \mathbf{G}_{\mathbf{x}_j}\right) \circ \left(\mathbf{F} \circ \mathbf{R}_r^{-1}\right) = \mathbf{R}_s \circ \mathbf{R}_r^{-1}$$

while the inverse of one of the maps in the chart is of the form

$$\mathbf{F} \circ \mathbf{R}_r^{-1}$$

showing that if Ω is a C^l manifold, then $\mathbf{F}(\Omega)$ is also. This also verifies the claim that if $(U_r, \mathbf{R}_r)$ is an oriented atlas for Ω, then $\mathbf{F}(\Omega)$ also has an oriented atlas since the overlap maps described above are all of the form $\mathbf{R}_s \circ \mathbf{R}_r^{-1}$.

It remains to verify the assertion about boundaries. $\mathbf{y} \in \partial \mathbf{F}(\Omega)$ if and only if for some $\mathbf{x}_i \in U_r$,

$$\mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_i}(\mathbf{y}) \in \mathbb{R}_0^n$$

if and only if

$$\mathbf{G}_{\mathbf{x}_i}(\mathbf{y}) \in \partial\Omega$$

if and only if

$$\mathbf{G}_{\mathbf{x}_i}(\mathbf{F}(\mathbf{x})) = \mathbf{x} \in \partial\Omega$$

where $\mathbf{F}(\mathbf{x}) = \mathbf{y}$ if and only if $\mathbf{y} \in \mathbf{F}(\partial\Omega)$. This proves the theorem.

A function $\mathbf{F}$ satisfying the condifions listed in Theorem17.16 is called a regular mapping.

17.4 Stokes theorem

One of the most important theorems in this subject is Stokes theorem which relates an integral of a differential form on an oriented manifold with boundary to another integral of a differential form on the boundary.

Lemma 17.17 *Let $(U_i, \mathbf{R}_i)$ be an oriented atlas for Ω, an oriented manifold. Also let $\omega = \sum a_I d\mathbf{x}^I$ be a differential form for which a_I has compact support contained in U_r for each I. Then*

$$\sum_I \int_{\mathbf{R}_r U_r} a_I \circ \mathbf{R}_r^{-1}(\mathbf{u}) \frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)} du = \int_\Omega \omega. \tag{17.9}$$

Proof: Let $K \subseteq U_r$ be a compact set for which $a_I = 0$ off K for all I. Then consider the atlas, $(U_i', \mathbf{R}_i)$ where $U_i' \equiv U_i \cap K^C$ for all $i \neq r, U_r \equiv U_r'$. Thus $(U_i', \mathbf{R}_i)$ is also an oriented atlas. Now let $\{\psi_i\}$be a partition of unity subordinate to the sets $\{U_i'\}$. Then if $i \neq r$

$$\psi_i(\mathbf{x})\, a_I(\mathbf{x}) = 0$$

for all I. Therefore,

$$\begin{aligned}
\int_\Omega \omega &\equiv \sum_i \sum_I \int_{\mathbf{R}_i U_i} (\psi_i a_I) \circ \mathbf{R}_i^{-1}(\mathbf{u}) \frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)} du \\
&= \sum_I \int_{\mathbf{R}_r U_r} a_I \circ \mathbf{R}_r^{-1}(\mathbf{u}) \frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)} du
\end{aligned}$$

and this proves the lemma.

Before proving Stokes theorem we need a definition. (This subject has a plethora of definitions.)

Definition 17.18 *Let $\omega = \sum_I a_I(\mathbf{x})\, dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}}$ be a differential form of order $n-1$ where a_I is in $C_c^1(\mathbb{R}^n)$. Then we define $d\omega$, a differential form of order n by replacing $a_I(\mathbf{x})$ with*

$$da_I(\mathbf{x}) \equiv \sum_{k=1}^{m} \frac{\partial a_I(\mathbf{x})}{\partial x^k} dx^k \tag{17.10}$$

and putting a wedge after the dx^k. Therefore,

$$d\omega \equiv \sum_I \sum_{k=1}^{m} \frac{\partial a_I(\mathbf{x})}{\partial x^k} dx^k \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}}. \tag{17.11}$$

Having wallowed in definitions, we are finally ready to prove Stoke's theorem. The proof is very elementary, amounting to a computation which comes from the definitions.

Theorem 17.19 *(Stokes theorem) Let Ω be a C^2 orientable manifold with boundary and let $\omega \equiv \sum_I a_I(\mathbf{x})\, dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}}$ be a differential form of order $n-1$ for which a_I is C^1. Then*

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega. \tag{17.12}$$

Proof: We let $(U_r, \mathbf{R}_r)$, $r \in \{1, \cdots, p\}$ be an oriented atlas for Ω and let $\{\psi_r\}$ be a partition of unity subordinate to $\{U_r\}$. Let $B \equiv \{r : \mathbf{R}_r U_r \cap \mathbb{R}_0^n \neq \emptyset\}$.

$$\int_{\Omega} d\omega \equiv \sum_{j=1}^{m} \sum_I \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r} \left(\psi_r \frac{\partial a_I}{\partial x^j} \circ \mathbf{R}_r^{-1} \right)(\mathbf{u}) \frac{\partial \left(x^j x^{i_1} \cdots x^{i_n - 1}\right)}{\partial \left(u^1 \cdots u^n\right)} du$$

Now

$$\psi_r \frac{\partial a_I}{\partial x^j} = \frac{\partial (\psi_r a_I)}{\partial x^j} - \frac{\partial \psi_r}{\partial x^j} a_I$$

Therefore,

$$\int_{\Omega} d\omega = \sum_{j=1}^{m} \sum_I \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r} \left(\frac{\partial (\psi_r a_I)}{\partial x^j} \circ \mathbf{R}_r^{-1} \right)(\mathbf{u}) \frac{\partial \left(x^j x^{i_1} \cdots x^{i_n - 1}\right)}{\partial \left(u^1 \cdots u^n\right)} du$$

$$- \sum_{j=1}^{m} \sum_I \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r} \left(\frac{\partial \psi_r}{\partial x^j} a_I \circ \mathbf{R}_r^{-1} \right)(\mathbf{u}) \frac{\partial \left(x^j x^{i_1} \cdots x^{i_n - 1}\right)}{\partial \left(u^1 \cdots u^n\right)} du \tag{17.13}$$

Consider the second line in (17.13). The expression, $\frac{\partial \psi_r}{\partial x^j} a_I$ has compact support in U_r and so by Lemma 17.17, this equals

$$- \sum_{r=1}^{p} \int_{\Omega} \sum_{j=1}^{m} \sum_I \frac{\partial \psi_r}{\partial x^j} a_I dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}} =$$

$$- \int_{\Omega} \sum_{j=1}^{m} \sum_I \sum_{r=1}^{p} \frac{\partial \psi_r}{\partial x^j} a_I dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}} =$$

$$- \int_{\Omega} \sum_{j=1}^{m} \sum_I \frac{\partial}{\partial x^j} \left(\sum_{r=1}^{p} \psi_r \right) a_I dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}} = 0$$

because $\sum_r \psi_r = 1$ on Ω. Thus we are left to consider the first line in (17.13).

$$\frac{\partial\left(x^j x^{i_1}\cdots x^{i_n-1}\right)}{\partial\left(u^1\cdots u^n\right)} = \sum_{k=1}^n \frac{\partial x^j}{\partial u^k} A^{1k}$$

where A^{1k} denotes the cofactor of $\frac{\partial x^j}{\partial u^k}$. Thus, letting $\mathbf{x} = \mathbf{R}_r^{-1}(\mathbf{u})$ and using the chain rule,

$$\begin{aligned}
\int_\Omega d\omega &= \sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbf{R}_r U_r} \left(\frac{\partial(\psi_r a_I)}{\partial x^j}\circ \mathbf{R}_r^{-1}\right)(\mathbf{u})\,\frac{\partial\left(x^j x^{i_1}\cdots x^{i_n-1}\right)}{\partial\left(u^1\cdots u^n\right)} du \\
&= \sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbf{R}_r U_r} \left(\frac{\partial(\psi_r a_I)}{\partial x^j}(\mathbf{x})\right) \sum_{k=1}^n \frac{\partial x^j}{\partial u^k} A^{1k} du \\
&= \sum_{j=1}^m \sum_I \sum_{r=1}^p \sum_{k=1}^n \int_{\mathbf{R}_r U_r} \left(\frac{\partial\left(\psi_r a_I\circ \mathbf{R}_r^{-1}\right)}{\partial u^k}\right) A^{1k} du \\
&= \sum_{k=1}^n \left(\sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbb{R}^n_\leq} \left(\frac{\partial\left(\psi_r a_I\circ \mathbf{R}_r^{-1}\right)}{\partial u^k}\right) A^{1k} du\right)
\end{aligned}$$

There are two cases here on the kth term in the above sum over k. The first case is where $k \neq 1$. In this case we integrate by parts and obtain the kth term is

$$\sum_{j=1}^m \sum_I \sum_{r=1}^p \left(-\int_{\mathbb{R}^n_\leq} \psi_r a_I \circ \mathbf{R}_r^{-1} \frac{\partial A^{1k}}{\partial u^k} du\right)$$

In the case where $k = 1$, we integrate by parts and obtain the kth term equals

$$\sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbb{R}^{n-1}} \left(\psi_r a_I \circ \mathbf{R}_r^{-1} A^{11}|_{-\infty}^0\right) du_2\cdots du_n - \int_{\mathbb{R}^n_\leq} \psi_r a_I \circ \mathbf{R}_r^{-1} \frac{\partial A^{11}}{\partial u^1} du$$

Adding all these terms up and using Lemma 17.8, we finally obtain

$$\int_\Omega d\omega = \sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbb{R}^{n-1}} \left(\psi_r a_I \circ \mathbf{R}_r^{-1} A^{11}|_{-\infty}^0\right) =$$

$$\begin{aligned}
& \sum_{j=1}^m \sum_I \sum_{r\in B} \int_{\mathbb{R}^{n-1}} \left(\psi_r a_I \circ \mathbf{R}_r^{-1}\right) \frac{\partial\left(x^{i_1}\cdots x^{i_{n-1}}\right)}{\partial\left(u^2\cdots u^n\right)}\left(0,u^2,\cdots,u^n\right) du_2\cdots du_n \\
= & \sum_{j=1}^m \sum_I \sum_{r\in B} \int_{\mathbf{R}_r U_r \cap \mathbb{R}^n_0} \left(\psi_r a_I \circ \mathbf{R}_r^{-1}\right) \frac{\partial\left(x^{i_1}\cdots x^{i_{n-1}}\right)}{\partial\left(u^2\cdots u^n\right)}\left(0,u^2,\cdots,u^n\right) du_2\cdots du_n \\
= & \int_{\partial\Omega} \omega.
\end{aligned}$$

This proves Stoke's theorem.

17.5 A generalization

We used that $\mathbf{R}_i^{-1}$ is C^2 in the proof of Stokes theorem but the end result is a formula which involves only the first derivatives of $\mathbf{R}_i^{-1}$. This suggests that it is not necessary to make this assumption. This is in fact the case. We give an argument which shows that Stoke's theorem holds for oriented C^1manifolds. For a still more general theorem see Section 20.7. We do not present this more general result here because it depends on too many hard theorems which are not proved until later.

Now suppose Ω is only a C^1 orientable manifold. Then in the proof of Stoke's theorem, we can say there exists some subsequence, $n \to \infty$ such that

$$\int_\Omega d\omega \equiv \lim_{n\to\infty} \sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbf{R}_r U_r} \left(\psi_r \frac{\partial a_I}{\partial x^j} \circ \left(\mathbf{R}_r^{-1} * \phi_N \right) \right) (\mathbf{u}) \frac{\partial \left(x^j x^{i_1} \cdots x^{i_n - 1} \right)}{\partial \left(u^1 \cdots u^n \right)} du$$

where ϕ_N is a mollifier and

$$\frac{\partial \left(x^j x^{i_1} \cdots x^{i_n - 1} \right)}{\partial \left(u^1 \cdots u^n \right)}$$

is obtained from

$$\mathbf{x} = \mathbf{R}_r^{-1} * \phi_N (\mathbf{u}).$$

The reason we can assert this limit is that from the dominated convergence theorem, it is routine to show

$$\left(\mathbf{R}_r^{-1} * \phi_N \right)_{,i} = \left(\mathbf{R}_r^{-1} \right)_{,i} * \phi_N$$

and by results presented in Chapter 12 using Minkowski's inequality, we see $\lim_{N\to\infty} \left(\mathbf{R}_r^{-1} * \phi_N \right)_{,i} = \left(\mathbf{R}_r^{-1} \right)_{,i}$ in $L^p (\mathbf{R}_r U_r)$ for every p. Taking an appropriate subsequence, we can obtain, in addition to this, almost everywhere convergence for every partial derivative and every $\mathbf{R}_r$.We may also arrange to have $\sum \psi_r = 1$ near Ω. We may do this as follows. If $U_r = O_r \cap \Omega$ where O_r is open in $\mathbb{R}^m$, we see that the compact set, Ω is covered by the open sets, O_r. Consider the compact set, $\overline{\Omega + B(\mathbf{0}, \delta)} \equiv K$ where $\delta < dist \left(\Omega, \mathbb{R}^m \setminus \cup_{i=1}^p O_i \right)$. Then take a partition of unity subordinate to the open cover $\{O_i\}$ which sums to 1 on K. Then for N large enough, $\mathbf{R}_r^{-1} * \phi_N (\mathbf{R}_r U_i)$ will lie in this set, K, where $\sum \psi_r = 1$.

Then we do the computations as in the proof of Stokes theorem. Using the same computations, with $\mathbf{R}_r^{-1} * \phi_N$ in place of $\mathbf{R}_r^{-1}$, along with the dominated convergence theorem,

$$\int_\Omega d\omega =$$

$$\lim_{n\to\infty} \sum_{j=1}^m \sum_I \sum_{r\in B} \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \left(\psi_r a_I \circ \left(\mathbf{R}_r^{-1} * \phi_N \right) \right) \frac{\partial \left(x^{i_1} \cdots x^{i_{n-1}} \right)}{\partial \left(u^2 \cdots u^n \right)} \left(0, u^2, \cdots, u^n \right) du_2 \cdots du_n$$

$$= \sum_{j=1}^m \sum_I \sum_{r\in B} \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \left(\psi_r a_I \circ \mathbf{R}_r^{-1} \right) \frac{\partial \left(x^{i_1} \cdots x^{i_{n-1}} \right)}{\partial \left(u^2 \cdots u^n \right)} \left(0, u^2, \cdots, u^n \right) du_2 \cdots du_n \equiv \int_{\partial\Omega} \omega.$$

This yields the following generalization of Stoke's theorem to the case of C^1 manifolds.

Theorem 17.20 *(Stokes theorem) Let Ω be a C^1 oriented manifold with boundary and let $\omega \equiv \sum_I a_I (\mathbf{x}) dx^{i_1} \wedge \cdots \wedge dx^{i_{n-1}}$ be a differential form of order $n-1$ for which a_I is C^1. Then*

$$\int_{\partial\Omega} \omega = \int_\Omega d\omega. \tag{17.14}$$

17.6 Surface measures

Let Ω be a C^1 manifold in $\mathbb{R}^m$, oriented or not. Let f be a continuous function defined on Ω, and let $(U_i, \mathbf{R}_i)$ be an atlas and let $\{\psi_i\}$ be a C^∞ partition of unity subordinate to the sets, U_i as described earlier. If $\omega = \sum_I a_I(\mathbf{x})\, d\mathbf{x}^I$ is a differential form, we may always assume

$$d\mathbf{x}^I = dx^{i_1} \wedge \cdots \wedge dx^{i_n}$$

where $i_1 < i_2 < \cdots < i_n$. The reason for this is that in taking an integral used to integrate the differential form, a switch in two of the dx^j results in switching two rows in the determinant, $\frac{\partial(x^{i_1} \cdots x^{i_n})}{\partial(u^1 \cdots u^n)}$, implying that any two of these differ only by a multiple of -1. Therefore, there is no loss of generality in assuming from now on that in the sum for ω, I is always a list of indices which are strictly increasing. The case where some term of ω has a repeat, $dx^{i_r} = dx^{i_s}$ can be ignored because such terms deliver zero in integrating the differential form because they involve a determinant having two equal rows. We emphasize again that from now on I will refer to an increasing list of indices.

Let

$$J_i(\mathbf{u}) \equiv \left[\sum_I \left(\frac{\partial (x^{i_1} \cdots x^{i_n})}{\partial (u^1 \cdots u^n)} \right)^2 \right]^{1/2}$$

where here the sum is taken over all possible increasing lists of n indices, I, from $\{1, \cdots, m\}$ and $\mathbf{x} = \mathbf{R}_i^{-1}\mathbf{u}$. Thus there are $\binom{m}{n}$ terms in the sum. Note that if $m = n$ we obtain only one term in the sum, the absolute value of the determinant of $D\mathbf{x}(\mathbf{u})$. We define a positive linear functional, Λ on $C(\Omega)$ as follows:

$$\Lambda f \equiv \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} f \psi_i \left(\mathbf{R}_i^{-1}(\mathbf{u}) \right) J_i(\mathbf{u})\, du. \tag{17.15}$$

We will now show this is well defined.

Lemma 17.21 *The functional defined in (17.15) does not depend on the choice of atlas or the partition of unity.*

Proof: In (17.15), let $\{\psi_i\}$ be a partition of unity as described there which is associated with the atlas $(U_i, \mathbf{R}_i)$ and let $\{\eta_i\}$ be a partition of unity associated in the same manner with the atlas $(V_i, \mathbf{S}_i)$. Using the change of variables formula with $\mathbf{u} = \left(\mathbf{R}_i \circ \mathbf{S}_j^{-1}\right)\mathbf{v}$

$$\sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f \left(\mathbf{R}_i^{-1}(\mathbf{u}) \right) J_i(\mathbf{u})\, du = \tag{17.16}$$

$$\sum_{i=1}^{p} \sum_{j=1}^{q} \int_{\mathbf{R}_i U_i} \eta_j \psi_i f \left(\mathbf{R}_i^{-1}(\mathbf{u}) \right) J_i(\mathbf{u})\, du =$$

$$\sum_{i=1}^{p} \sum_{j=1}^{q} \int_{\mathbf{S}_j(U_i \cap V_j)} \eta_j \left(\mathbf{S}_j^{-1}(\mathbf{v}) \right) \psi_i \left(\mathbf{S}_j^{-1}(\mathbf{v}) \right) f \left(\mathbf{S}_j^{-1}(\mathbf{v}) \right) J_i(\mathbf{u}) \left| \frac{\partial (u^1 \cdots u^n)}{\partial (v^1 \cdots v^n)} \right| dv$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{q} \int_{\mathbf{S}_j(U_i \cap V_j)} \eta_j \left(\mathbf{S}_j^{-1}(\mathbf{v}) \right) \psi_i \left(\mathbf{S}_j^{-1}(\mathbf{v}) \right) f \left(\mathbf{S}_j^{-1}(\mathbf{v}) \right) J_j(\mathbf{v})\, dv. \tag{17.17}$$

This yields

$$\text{the definition of } \Lambda f \text{ using } (U_i, \mathbf{R}_i) \equiv$$

$$\sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du =$$

$$\sum_{i=1}^{p} \sum_{j=1}^{q} \int_{\mathbf{S}_j(U_i \cap V_j)} \eta_j\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) \psi_i\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) f\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) J_j(\mathbf{v})\, dv$$

$$= \sum_{j=1}^{q} \int_{\mathbf{S}_j(V_j)} \eta_j\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) f\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) J_j(\mathbf{v})\, dv$$

$$\text{the definition of } \Lambda f \text{ using } (V_i, \mathbf{S}_i).$$

This proves the lemma.

This lemma implies the following theorem.

Theorem 17.22 *Let Ω be a C^k manifold with boundary. Then there exists a unique Radon measure, μ, defined on Ω such that whenever f is a continuous function defined on Ω and $(U_i, \mathbf{R}_i)$ denotes an atlas and $\{\psi_i\}$ a partition of unity subordinate to this atlas, we have*

$$\Lambda f = \int_{\Omega} f d\mu = \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du. \tag{17.18}$$

Furthermore, for any $f \in L^1(\Omega, \mu)$,

$$\int_{\Omega} f d\mu = \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \tag{17.19}$$

and a subset, A, of Ω is μ measurable if and only if for all $r, \mathbf{R}_r(U_r \cap A)$ is $J_r(\mathbf{u})\, du$ measurable.

Proof: We begin by proving the following claim.

Claim :A set, $S \subseteq U_i$, has μ measure zero in U_i, if and only if $\mathbf{R}_i S$ has measure zero in $\mathbf{R}_i U_i$ with respect to the measure, $J_i(\mathbf{u})\, du$.

Proof of the claim: Let $\varepsilon > 0$ be given. By outer regularity, there exists a set, $V \subseteq U_i$, open in Ω such that $\mu(V) < \varepsilon$ and $S \subseteq V \subseteq U_i$. Then $\mathbf{R}_i V$ is open in $\mathbb{R}^n_{\leq}$ and contains $\mathbf{R}_i S$. Letting $h \prec O$, where $O \cap \mathbb{R}^n_{\leq} = \mathbf{R}_i V$ and $m_n(O) < m_n(\mathbf{R}_i V) + \varepsilon$, and letting $h_1(\mathbf{x}) \equiv h(\mathbf{R}_i(\mathbf{x}))$ for $\mathbf{x} \in U_i$, we see $h_1 \prec V$. By Corollary 12.24, we can also choose our partition of unity so that $spt(h_1) \subseteq \{\mathbf{x} \in \mathbb{R}^m : \psi_i(\mathbf{x}) = 1\}$. Thus $\psi_j h_1\left(\mathbf{R}_j^{-1}(u)\right) = 0$ unless $j = i$ when this reduces to $h_1\left(\mathbf{R}_i^{-1}(u)\right)$. Thus

$$\begin{aligned}
\varepsilon &\geq \mu(V) \geq \int_V h_1 d\mu = \int_{\Omega} h_1 d\mu = \sum_{j=1}^{p} \int_{\mathbf{R}_j U_j} \psi_j h_1\left(\mathbf{R}_j^{-1}(\mathbf{u})\right) J_j(\mathbf{u})\, du \\
&= \int_{\mathbf{R}_i U_i} h_1\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du = \int_{\mathbf{R}_i U_i} h(\mathbf{u}) J_i(\mathbf{u})\, du = \int_{\mathbf{R}_i V} h(\mathbf{u}) J_i(\mathbf{u})\, du \\
&\geq \int_O h(\mathbf{u}) J_i(\mathbf{u})\, du - K_i \varepsilon,
\end{aligned}$$

where $K_i \geq ||J_i||_\infty$. Now this holds for all $h \prec O$ and so

$$\int_{\mathbf{R}_i S} J_i(\mathbf{u})\,du \leq \int_{\mathbf{R}_i V} J_i(\mathbf{u})\,du \leq \int_O J_i(\mathbf{u})\,du \leq \varepsilon(1+K_i).$$

Since ε is arbitrary, this shows $\mathbf{R}_i S$ has mesure zero with respect to the measure, $J_i(\mathbf{u})\,du$ as claimed.

Now we prove the converse. Suppose $\mathbf{R}_i S$ has $J_r(\mathbf{u})\,du$ measure zero. Then there exists an open set, O such that $O \supseteq \mathbf{R}_i S$ and

$$\int_O J_i(\mathbf{u})\,du < \varepsilon.$$

Thus $\mathbf{R}_i^{-1}(O \cap \mathbf{R}_i U_i)$ is open in Ω and contains S. Let $h \prec \mathbf{R}_i^{-1}(O \cap \mathbf{R}_i U_i)$ be such that

$$\int_\Omega h d\mu + \varepsilon > \mu\left(\mathbf{R}_i^{-1}(O \cap \mathbf{R}_i U_i)\right) \geq \mu(S)$$

As in the first part, we can choose our partition of unity such that $h(\mathbf{x}) = 0$ off the set,

$$\{\mathbf{x} \in \mathbb{R}^m : \psi_i(\mathbf{x}) = 1\}$$

and so as in this part of the argument,

$$\begin{aligned}
\int_\Omega h d\mu &\equiv \sum_{j=1}^p \int_{\mathbf{R}_j U_j} \psi_j h\left(\mathbf{R}_j^{-1}(\mathbf{u})\right) J_j(\mathbf{u})\,du \\
&= \int_{\mathbf{R}_i U_i} h\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\,du \\
&= \int_{O \cap \mathbf{R}_i U_i} h\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\,du \\
&\leq \int_O J_i(\mathbf{u})\,du < \varepsilon
\end{aligned}$$

and so $\mu(S) \leq 2\varepsilon$. Since ε is arbitrary, this proves the claim.

Now let $A \subseteq U_r$ be μ measurable. By the regularity of the measure, there exists an F_σ set, F and a G_δ set, G such that $U_r \supseteq G \supseteq A \supseteq F$ and $\mu(G \setminus F) = 0$.(Recall a G_δ set is a countable intersection of open sets and an F_σ set is a countable union of closed sets.) Then since Ω is compact, it follows each of the closed sets whose union equals F is a compact set. Thus if $F = \cup_{k=1}^\infty F_k$ we know $\mathbf{R}_r(F_k)$ is also a compact set and so $\mathbf{R}_r(F) = \cup_{k=1}^\infty \mathbf{R}_r(F_k)$ is a Borel set. Similarly, $\mathbf{R}_r(G)$ is also a Borel set. Now by the claim,

$$\int_{\mathbf{R}_r(G \setminus F)} J_r(\mathbf{u})\,du = 0.$$

We also see that since $\mathbf{R}_r$ is one to one,

$$\mathbf{R}_r G \setminus \mathbf{R}_r F = \mathbf{R}_r(G \setminus F)$$

and so

$$\mathbf{R}_r(F) \subseteq \mathbf{R}_r(A) \subseteq \mathbf{R}_r(G)$$

where $\mathbf{R}_r(G) \setminus \mathbf{R}_r(F)$ has measure zero. By completeness of the measure, $J_i(\mathbf{u})\,du$, we see $\mathbf{R}_r(A)$ is measurable. It follows that if $A \subseteq \Omega$ is μ measurable, then $\mathbf{R}_r(U_r \cap A)$ is $J_r(\mathbf{u})\,du$ measurable for all r. The converse is entirely similar.

Letting $f \in L^1(\Omega, \mu)$, we use the fact that μ is a Radon mesure to obtain a sequence of continuous functions, $\{f_k\}$ which converge to f in $L^1(\Omega, \mu)$ and for μ a.e. $\mathbf{x}$. Therefore, the sequence $\{f_k(\mathbf{R}_i^{-1}(\cdot))\}$ is a Cauchy sequence in $L^1(\mathbf{R}_i U_i; \psi_i(\mathbf{R}_i^{-1}(\mathbf{u})) J_i(\mathbf{u}) du)$. It follows there exists

$$g \in L^1\left(\mathbf{R}_i U_i; \psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du\right)$$

such that $f_k(\mathbf{R}_i^{-1}(\cdot)) \to g$ in $L^1(\mathbf{R}_i U_i; \psi_i(\mathbf{R}_i^{-1}(\mathbf{u})) J_i(\mathbf{u}) du)$. By the pointwise convergence, $g(\mathbf{u}) = f(\mathbf{R}_i^{-1}(\mathbf{u}))$ for μ a.e. $\mathbf{R}_i^{-1}(\mathbf{u}) \in U_i$. By the above claim, $g = f \circ \mathbf{R}_i^{-1}$ for a.e. $\mathbf{u} \in \mathbf{R}_i U_i$ and so

$$f \circ \mathbf{R}_i^{-1} \in L^1(\mathbf{R}_i U_i; J_i(\mathbf{u})\, du)$$

and we can write

$$\begin{aligned}
\int_\Omega f d\mu &= \lim_{k\to\infty} \int_\Omega f_k d\mu = \lim_{k\to\infty} \sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i f_k\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \\
&= \sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) g(\mathbf{u}) J_i(\mathbf{u})\, du \\
&= \sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du.
\end{aligned}$$

This proves the theorem.

Corollary 17.23 *Let $f \in L^1(\Omega; \mu)$ and suppose $f(\mathbf{x}) = 0$ for all $x \notin U_r$ where $(U_r, \mathbf{R}_r)$ is a chart in a C^k atlas for Ω. Then*

$$\int_\Omega f d\mu = \int_{U_r} f d\mu = \int_{\mathbf{R}_r U_r} f\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du. \tag{17.20}$$

Proof: Using regularity of the measures, we can pick a compact subset, K, of U_r such that

$$\left| \int_{U_r} f d\mu - \int_K f d\mu \right| < \varepsilon.$$

Now by Corollary 12.24, we can choose the partition of unity such that $K \subseteq \{\mathbf{x} \in \mathbb{R}^m : \psi_r(\mathbf{x}) = 1\}$. Then

$$\begin{aligned}
\int_K f d\mu &= \sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i f \mathcal{X}_K\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \\
&= \int_{\mathbf{R}_r U_r} f \mathcal{X}_K\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du.
\end{aligned}$$

Therefore, letting $K_l \uparrow \mathbf{R}_r U_r$ we can take a limit and conclude

$$\left| \int_{U_r} f d\mu - \int_{\mathbf{R}_r U_r} f\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du \right| \le \varepsilon.$$

Since ε is arbitrary, this proves the corollary.

17.7 Divergence theorem

What about writing the integral of a differential form in terms of this measure? This is a useful idea because it allows us to obtain various important formulas such as the divergence theorem which are traditionally written not in terms of differential forms but in terms of measure on the surface and outer normals. Let ω be a differential form,

$$\omega(\mathbf{x}) = \sum_I a_I(\mathbf{x})\, d\mathbf{x}^I$$

where a_I is continuous and the sum is taken over all increasing lists from $\{1, \cdots, m\}$. We assume Ω is a C^k manifold which is orientable and that $(U_r, \mathbf{R}_r)$ is an oriented atlas for Ω while, $\{\psi_r\}$ is a C^∞ partition of unity subordinate to the U_r.

Lemma 17.24 *Consider the set,*

$$S \equiv \left\{\mathbf{x} \in \Omega : \text{ for some } r, \mathbf{x} = \mathbf{R}_r^{-1}(\mathbf{u}) \text{ where } \mathbf{x} \in U_r \text{ and } J_r(\mathbf{u}) = 0 \right\}.$$

Then $\mu(S) = 0$.

Proof: Let S_r denote those points, $\mathbf{x}$, of U_r for which $\mathbf{x} = \mathbf{R}_r^{-1}(\mathbf{u})$ and $J_r(\mathbf{u}) = 0$. Thus $S = \cup_{r=1}^p S_r$. From Corollary 17.23

$$\int_\Omega \mathcal{X}_{S_r} d\mu = \int_{\mathbf{R}_r U_r} \mathcal{X}_{S_r}\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du = 0$$

and so

$$\mu(S) \leq \sum_{k=1}^p \mu(S_k) = 0.$$

This proves the lemma.

With respect to the above atlas, we define a function of $\mathbf{x}$ in the following way. For $I = (i_1, \cdots, i_n)$ an increasing list of indices,

$$o^I(\mathbf{x}) \equiv \begin{cases} \left(\frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial(u^1 \cdots u^n)}\right) / J_r(\mathbf{u}), \text{ if } \mathbf{x} \in U_r \setminus S \\ 0 \text{ if } \mathbf{x} \in S \end{cases}$$

Now it follows from Lemma 17.10 that if we had used a different atlas having the same orientation, then $o^I(\mathbf{x})$ would be unchanged. We define a vector in $\mathbb{R}^{\binom{m}{n}}$ by letting the I^{th} component of $\mathbf{o}(\mathbf{x})$ be defined by $o^I(\mathbf{x})$. Also note that since $\mu(S) = 0$,

$$\sum_I o^I(\mathbf{x})^2 = 1 \ \mu \text{ a.e.}$$

Define

$$\omega(\mathbf{x}) \cdot \mathbf{o}(\mathbf{x}) \equiv \sum_I a_I(\mathbf{x})\, o^I(\mathbf{x}),$$

From the definition of what we mean by the integral of a differential form, Definition 17.11, it follows that

$$\begin{aligned} \int_\Omega \omega &\equiv \sum_{r=1}^p \sum_I \int_{\mathbf{R}_r U_r} \psi_r\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) a_I\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(u^1 \cdots u^n\right)} du \\ &= \sum_{r=1}^p \int_{\mathbf{R}_r U_r} \psi_r\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \omega\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \cdot \mathbf{o}\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du \\ &\equiv \int_\Omega \omega \cdot \mathbf{o}\, d\mu \end{aligned} \tag{17.21}$$

Note that $\omega \cdot \mathbf{o}$ is bounded and measurable so is in L^1.

Lemma 17.25 *Let Ω be a C^k oriented manifold in $\mathbb{R}^n$ with an oriented atlas, $(U_r, \mathbf{R}_r)$. Letting $\mathbf{x} = \mathbf{R}_r^{-1}\mathbf{u}$ and letting $2 \leq j \leq n$, we have*

$$\sum_{i=1}^{n} \frac{\partial x^i}{\partial u^j} (-1)^{i+1} \frac{\partial \left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial \left(u^2 \cdots u^n\right)} = 0 \tag{17.22}$$

for each r. Here, $\widehat{x^i}$ means this is deleted. If for each r,

$$\frac{\partial \left(x^1 \cdots x^n\right)}{\partial \left(u^1 \cdots u^n\right)} \geq 0 \ ,$$

then for each r

$$\sum_{i=1}^{n} \frac{\partial x^i}{\partial u^1} (-1)^{i+1} \frac{\partial \left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial \left(u^2 \cdots u^n\right)} \geq 0 \ a.e. \tag{17.23}$$

Proof: (17.23) follows from the observation that

$$\frac{\partial \left(x^1 \cdots x^n\right)}{\partial \left(u^1 \cdots u^n\right)} = \sum_{i=1}^{n} \frac{\partial x^i}{\partial u^1} (-1)^{i+1} \frac{\partial \left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial \left(u^2 \cdots u^n\right)}$$

by expanding the determinant,

$$\frac{\partial \left(x^1 \cdots x^n\right)}{\partial \left(u^1 \cdots u^n\right)},$$

along the first column. Formula (17.22) follows from the observation that the sum in (17.22) is just the determinant of a matrix which has two equal columns. This proves the lemma.

With this lemma, it is easy to verify a general form of the divergence theorem from Stoke's theorem. First we recall the definition of the divergence of a vector field.

Definition 17.26 *Let O be an open subset of $\mathbb{R}^n$ and let $\mathbf{F}(\mathbf{x}) \equiv \sum_{k=1}^n F^k(\mathbf{x}) \mathbf{e}_k$ be a vector field for which $F^k \in C^1(O)$. Then*

$$div(\mathbf{F}) \equiv \sum_{k=1}^{n} \frac{\partial F_k}{\partial x^k}.$$

Theorem 17.27 *Let Ω be an orientable C^k n manifold with boundary in $\mathbb{R}^n$ having an oriented atlas, $(U_r, \mathbf{R}_r)$ which satisfies*

$$\frac{\partial \left(x^1 \cdots x^n\right)}{\partial \left(u^1 \cdots u^n\right)} \geq 0 \tag{17.24}$$

for all r. Then letting $\mathbf{n}(\mathbf{x})$ be the vector field whose i^{th} component taken with respect to the usual basis of $\mathbb{R}^n$ is given by

$$n^i(\mathbf{x}) \equiv \begin{cases} (-1)^{i+1} \frac{\partial \left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial \left(u^2 \cdots u^n\right)} / J_r(\mathbf{u}) \ \text{if } J_r(\mathbf{u}) \neq 0 \\ 0 \ \text{if } J_r(\mathbf{u}) = 0 \end{cases} \tag{17.25}$$

for $\mathbf{x} \in U_r \cap \partial\Omega$, it follows $\mathbf{n}(\mathbf{x})$ is independent of the choice of atlas provided the orientation remains unchanged. Also $\mathbf{n}(\mathbf{x})$ is a unit vector for a.e. $\mathbf{x} \in \partial\Omega$. Let $\mathbf{F} \in C^1\left(\overline{\Omega}; \mathbb{R}^n\right)$. Then we have the following formula which is called the divergence theorem.

$$\int_{\Omega} div(\mathbf{F})\, dx = \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n}\, d\mu, \tag{17.26}$$

where μ is the surface measure on $\partial\Omega$ defined above.

Proof: Recall that on $\partial\Omega$

$$J_r(\mathbf{u}) = \left[\sum_{i=1}^n \left(\frac{\partial\left(x^1\cdots\widehat{x^i}\cdots x^n\right)}{\partial(u^2\cdots u^n)}\right)^2\right]^{1/2}$$

From Lemma 17.10 and the definition of two atlass having the same orientation, we see that aside from sets of measure zero, the assertion about the independence of choice of atlas for the normal, $\mathbf{n}(\mathbf{x})$ is verified. Also, by Lemma 17.24, we know $J_r(\mathbf{u}) \neq 0$ off some set of measure zero for each atlas and so $\mathbf{n}(\mathbf{x})$ is a unit vector for μ a.e. $\mathbf{x}$.

Now we define the differential form,

$$\omega \equiv \sum_{i=1}^n (-1)^{i+1} F_i(\mathbf{x})\, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n.$$

Then from the definition of $d\omega$,

$$d\omega = div(\mathbf{F})\, dx^1 \wedge \cdots \wedge dx^n.$$

Now let $\{\psi_r\}$ be a partition of unity subordinate to the U_r. Then using (17.24) and the change of variables formula,

$$\begin{aligned}
\int_\Omega d\omega &\equiv \sum_{r=1}^p \int_{\mathbf{R}_r U_r} (\psi_r div(\mathbf{F}))\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \frac{\partial\left(x^1\cdots x^n\right)}{\partial(u^1\cdots u^n)} du \\
&= \sum_{r=1}^p \int_{U_r} (\psi_r div(\mathbf{F}))(\mathbf{x})\, dx = \int_\Omega div(\mathbf{F})\, dx.
\end{aligned}$$

Now

$$\begin{aligned}
\int_{\partial\Omega} \omega &\equiv \sum_{i=1}^n (-1)^{i+1} \sum_{r=1}^p \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \psi_r F_i\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \frac{\partial\left(x^1\cdots\widehat{x^i}\cdots x^n\right)}{\partial(u^2\cdots u^n)} du^2\cdots du^n \\
&= \sum_{r=1}^p \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \psi_r \left(\sum_{i=1}^n F_i n^i\right)\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du^2\cdots du^n \\
&\equiv \int_{\partial\Omega} \mathbf{F}\cdot\mathbf{n}\, d\mu.
\end{aligned}$$

By Stoke's theorem,

$$\int_\Omega div(\mathbf{F})\, dx = \int_\Omega d\omega = \int_{\partial\Omega} \omega = \int_{\partial\Omega} \mathbf{F}\cdot\mathbf{n}\, d\mu$$

and this proves the theorem.

Definition 17.28 *In the context of the divergence theorem, the vector,* $\mathbf{n}$ *is called the unit outer normal.*

Since we did not assume $\frac{\partial(x^1\cdots x^n)}{\partial(u^1\cdots u^n)} \neq 0$ for all $\mathbf{x} = \mathbf{R}_r^{-1}(\mathbf{u})$, this is about all we can say about the geometric significance of $\mathbf{n}$. However, it is usually the case that we are in a situation where this determinant is non zero. This is the case in the context of Proposition 17.14 for example, when this determinant was equal to one. The next proposition shows that $\mathbf{n}$ does what we would hope it would do if it really does deserve to be called a unit outer normal when $\frac{\partial(x^1\cdots x^n)}{\partial(u^1\cdots u^n)} \neq 0$.

Proposition 17.29 *Let* $\mathbf{n}$ *be given by (17.25) at a point of the boundary,*

$$\mathbf{x}_0 = \mathbf{R}_r^{-1}(\mathbf{u}_0), \mathbf{u}_0 \in \mathbb{R}_0^n,$$

where $(U_r, \mathbf{R}_r)$ *is a chart for the manifold,* Ω, *of Theorem 17.27 and*

$$\frac{\partial (x^1 \cdots x^n)}{\partial (u^1 \cdots u^n)}(\mathbf{u}_0) > 0 \tag{17.27}$$

at this point, then $|\mathbf{n}| = 1$ *and for all* $t > 0$ *small enough,*

$$\mathbf{x}_0 + t\mathbf{n} \notin \Omega \tag{17.28}$$

and

$$\mathbf{n} \cdot \frac{\partial \mathbf{x}}{\partial u^j} = 0 \tag{17.29}$$

for all $j = 2, \cdots, n$.

Proof: First note that (17.27) implies $J_r(\mathbf{u}_0) > 0$ and that we have already verified that $|\mathbf{n}| = 1$ and (17.29). Suppose the proposition is not true. Then there exists a sequence, $\{t_j\}$ such that $t_j > 0$ and $t_j \to 0$ for which

$$\mathbf{R}_r^{-1}(\mathbf{u}_0) + t_j \mathbf{n} \in \Omega.$$

Since U_r is open in Ω, it follows that for all j large enough,

$$\mathbf{R}_r^{-1}(\mathbf{u}_0) + t_j \mathbf{n} \in U_r.$$

Therefore, there exists $\mathbf{u}_j \in \mathbb{R}_{\leq}^n$ such that

$$\mathbf{R}_r^{-1}(\mathbf{u}_j) = \mathbf{R}_r^{-1}(\mathbf{u}_0) + t_j \mathbf{n}.$$

Now by the inverse function theorem, this implies

$$\begin{aligned}
\mathbf{u}_j &= \mathbf{R}_r\left(\mathbf{R}_r^{-1}(\mathbf{u}_0) + t_j \mathbf{n}\right) \\
&= D\mathbf{R}_r\left(\mathbf{R}_r^{-1}(\mathbf{u}_0)\right)\mathbf{n}t_j + \mathbf{u}_0 + \mathbf{o}(t_j) \\
&= D\mathbf{R}_r^{-1}(\mathbf{u}_0)\mathbf{n}t_j + \mathbf{u}_0 + \mathbf{o}(t_j).
\end{aligned}$$

At this point we take the first component of both sides and utilize the fact that the first component of $\mathbf{u}_0$ equals zero. Then

$$0 \geq u_j^1 = t_j \left(D\mathbf{R}_r^{-1}(\mathbf{u}_0)\mathbf{n}\right) \cdot \mathbf{e}_1 + o(t_j). \tag{17.30}$$

We consider the quantity, $\left(D\mathbf{R}_r^{-1}(\mathbf{u}_0)\mathbf{n}\right) \cdot \mathbf{e}_1$. From the formula for the inverse in terms of the transpose of the cofactor matrix, we obtain

$$\left(D\mathbf{R}_r^{-1}(\mathbf{u}_0)\mathbf{n}\right) \cdot \mathbf{e}_1 = \frac{1}{\det\left(D\mathbf{R}_r^{-1}(\mathbf{u}_0)\right)} (-1)^{1+j} M_{j1} n^j$$

where

$$M^{j1} = \frac{\partial \left(x^1 \cdots \widehat{x^j} \cdots x^n\right)}{\partial (u^2 \cdots u^n)}$$

is the determinant of the matrix obtained from deleting the j^{th} row and the first column of the matrix,

$$\begin{pmatrix} \frac{\partial x^1}{\partial u^1}(\mathbf{u}_0) & \cdots & \frac{\partial x^1}{\partial u^n}(\mathbf{u}_0) \\ \vdots & & \vdots \\ \frac{\partial x^n}{\partial u^1}(\mathbf{u}_0) & \cdots & \frac{\partial x^n}{\partial u^n}(\mathbf{u}_0) \end{pmatrix}$$

which is just the matrix of the linear transformation, $D\mathbf{R}_r^{-1}(\mathbf{u}_0)$ taken with respect to the usual basis vectors. Now using the given formula for n^j we see that $\left(D\mathbf{R}_r^{-1}(\mathbf{u}_0)\,\mathbf{n}\right)\cdot\mathbf{e}_1 > 0$. Therefore, we can divide by the positive number t_j in (17.30) and choose j large enough that

$$\left|\frac{o(t_j)}{t_j}\right| < \frac{\left(D\mathbf{R}_r^{-1}(\mathbf{u}_0)\,\mathbf{n}\right)\cdot\mathbf{e}_1}{2}$$

to obtain a contradiction. This proves the proposition and gives a geometric justification of the term "unit outer normal" applied to $\mathbf{n}$.

17.8 Exercises

1. In the section on surface measure we used

$$\left[\sum_I \left(\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)}\right)^2\right]^{1/2} \equiv J_i(\mathbf{u})$$

   What if we had used instead

$$\left[\sum_I \left|\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)}\right|^p\right]^{1/p} \equiv J_i(\mathbf{u})?$$

   Would everything have worked out the same? Why is there a preference for the exponent, 2?

2. Suppose Ω is an oriented $C^1 n$ manifold in $\mathbb{R}^n$and that for one of the charts, $\frac{\partial\left(x^1\cdots x^n\right)}{\partial\left(u^1\cdots u^n\right)} > 0$. Can it be concluded that this condition holds for all the charts? What if we also assume Ω is connected?

3. We defined manifolds with boundary in terms of the maps, $\mathbf{R}_i$ mapping into the special half space, $\{\mathbf{u}\in\mathbb{R}^n : u_1 \le 0\}$. We retained this special half space in the discussion of oriented manifolds. However, we could have used any half space in our definition. Show that if $n \ge 2$, there was no loss of generality in restricting our attention to this special half space. Is there a problem in defining oriented manifolds in this manner using this special half space in the case where $n = 1$?

4. Let Ω be an oriented Lipschitz or C^k n manifold with boundary in $\mathbb{R}^n$ and let $\mathbf{f}\in C^1\left(\overline{\Omega};\mathbb{R}^k\right)$. Show that

$$\int_\Omega \frac{\partial \mathbf{f}}{\partial x^j}dx = \int_{\partial\Omega} \mathbf{f} n_j d\mu$$

   where μ is the surface measure on $\partial\Omega$ discussed above. This says essentially that we can exchange differentiation with respect to x^j on Ω with multiplication by the j^{th} component of the exterior normal on $\partial\Omega$. Compare to the divergence theorem.

5. Recall the vector valued function, $\mathbf{o}(\mathbf{x})$, for a C^1 oriented manifold which has values in $\mathbb{R}^{\binom{m}{n}}$. Show that for an orientable manifold this function is continuous as well as having its length in $\mathbb{R}^{\binom{m}{n}}$ equal to one where the length is measured in the usual way.

6. In the proof of Lemma 17.4 we used a very hard result, the invariance of domain theorem. Assume the manifold in question is a C^1 manifold and give a much easier proof based on the inverse function theorem.

7. Suppose Ω is an oriented, C^1 2 manifold in $\mathbb{R}^3$. And consider the function
$$\mathbf{n}(\mathbf{x}) \equiv \left(\frac{\partial(x^2x^3)}{\partial(u^1u^2)}/J_i(\mathbf{u})\right)\mathbf{e}_1 - \left(\frac{\partial(x^1x^3)}{\partial(u^1u^2)}/J_i(\mathbf{u})\right)\mathbf{e}_2 + \left(\frac{\partial(x^1x^2)}{\partial(u^1u^2)}/J_i(\mathbf{u})\right)\mathbf{e}_3. \tag{17.31}$$
Show this function has unit length in $\mathbb{R}^3$, is independent of the choice of atlas having the same orientation, and is a continuous function of $\mathbf{x} \in \Omega$. Also show this function is perpendicular to Ω at every point by verifying its dot product with $\partial\mathbf{x}/\partial u^i$ equals zero. To do this last thing, observe the following determinant.
$$\begin{vmatrix} \frac{\partial x^1}{\partial u^i} & \frac{\partial x^1}{\partial u^1} & \frac{\partial x^3}{\partial u^2} \\ \frac{\partial x^2}{\partial u^i} & \frac{\partial x^2}{\partial u^1} & \frac{\partial x^3}{\partial u^2} \\ \frac{\partial x^3}{\partial u^i} & \frac{\partial x^3}{\partial u^1} & \frac{\partial x^3}{\partial u^2} \end{vmatrix}$$

8. $\uparrow$Take a long rectangular piece of paper, put one twist in it and then glue or tape the ends together. This is called a Moebus band. Take a pencil and draw a line down the center. If you keep drawing, you will be able to return to your starting point without ever taking the pencil off the paper. In other words, the shape you have obtained has only one side. Now if we consider the pencil as a normal vector to the plane, can you explain why the Moebus band is not orientable? For more fun with scissors and paper, cut the Moebus band down the center line and see what happens. You might be surprised.

9. Let Ω be a C^k 2 manifold with boundary in $\mathbb{R}^3$ and let $\omega = a_1(\mathbf{x})\,dx^1 + a_2(\mathbf{x})\,dx^2 + a_3(\mathbf{x})\,dx^3$ be a one form where the a_i are C^1 functions. Show that
$$d\omega = \left(\frac{\partial a_2}{\partial x^1} - \frac{\partial a_1}{\partial x^2}\right)dx^1 \wedge dx^2 +$$
$$\left(\frac{\partial a_3}{\partial x^1} - \frac{\partial a_1}{\partial x^3}\right)dx^1 \wedge dx^3 + \left(\frac{\partial a_3}{\partial x^2} - \frac{\partial a_2}{\partial x^3}\right)dx^2 \wedge dx^3.$$
Stoke's theorm would say that $\int_{\partial\Omega}\omega = \int_\Omega d\omega$. This is the classical form of Stoke's theorem.

10. $\uparrow$In the context of 9, Stoke's theorem is usually written in terms of vector notation rather than differential form notation. This involves the curl of a vector field and a normal to the given 2 manifold. Let $\mathbf{n}$ be given as in (17.31) and let a C^1 vector field be given by $\mathbf{a}(\mathbf{x}) \equiv a_1(\mathbf{x})\,\mathbf{e}^1 + a_2(\mathbf{x})\,\mathbf{e}^2 + a_3(\mathbf{x})\,\mathbf{e}^3$where the $\mathbf{e}^j$ are the standard unit vectors. Recall from elementary calculus courses that
$$\begin{aligned} curl\,(\mathbf{a}) &= \begin{vmatrix} \mathbf{e}^1 & \mathbf{e}^2 & \mathbf{e}^3 \\ \frac{\partial}{\partial x^1} & \frac{\partial}{\partial x^2} & \frac{\partial}{\partial x^3} \\ a_1(\mathbf{x}) & a_2(\mathbf{x}) & a_3(\mathbf{x}) \end{vmatrix} \\ &= \left(\frac{\partial a_3}{\partial x^2} - \frac{\partial a_2}{\partial x^3}\right)\mathbf{e}^1 + \\ &\quad \left(\frac{\partial a_1}{\partial x^3} - \frac{\partial a_3}{\partial x^1}\right)\mathbf{e}^2 + \left(\frac{\partial a_2}{\partial x^1} - \frac{\partial a_1}{\partial x^2}\right)\mathbf{e}^3. \end{aligned}$$
Letting μ be the surface measure on Ω and μ_1 the surface measure on $\partial\Omega$ defined above, show that
$$\int_\Omega d\omega = \int_\Omega curl\,(\mathbf{a}) \cdot \mathbf{n}\,d\mu$$

and so Stoke's formula takes the form

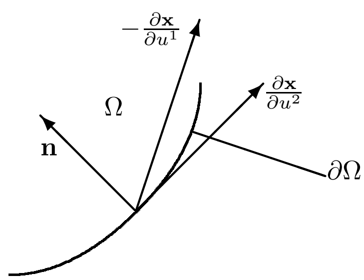
$$\begin{aligned}\int_{\Omega} curl\,(\mathbf{a}) \cdot \mathbf{n} d\mu &= \int_{\partial\Omega} a_1(\mathbf{x})\,dx^1 + a_2(\mathbf{x})\,dx^2 + a_3(\mathbf{x})\,dx^3 \\ &= \int_{\partial\Omega} \mathbf{a} \cdot \mathbf{T} d\mu_1\end{aligned}$$

where $\mathbf{T}$ is a unit tangent vector to $\partial\Omega$ given by

$$\mathbf{T}(\mathbf{x}) \equiv \left(\frac{\partial \mathbf{x}}{\partial u^2}\right) / \left|\frac{\partial \mathbf{x}}{\partial u^2}\right|.$$

Assume $\left|\frac{\partial \mathbf{x}}{\partial u^2}\right| \neq 0$. This means you have a well defined unit tangent vector to $\partial\Omega$.

11. $\uparrow$ It is nice to understand the geometric relationship between $\mathbf{n}$ and $\mathbf{T}$. Show that $-\frac{\partial \mathbf{x}}{\partial u^1}$ points into Ω while $\frac{\partial \mathbf{x}}{\partial u^2}$ points along $\partial\Omega$ and that $\mathbf{n} \times \frac{\partial \mathbf{x}}{\partial u^2} \cdot \left(-\frac{\partial \mathbf{x}}{\partial u^1}\right) = J_i(\mathbf{u})^2 > 0$. Using the geometric description of the cross product from elementary calculus, show $\mathbf{n}$ is the direction of a person walking arround $\partial\Omega$ with Ω on his left hand. The following picture is illustrative of the situation.



Representation Theorems

18.1 Radon Nikodym Theorem

This chapter is on various representation theorems. The first theorem, the Radon Nikodym Theorem, is a representation theorem for one measure in terms of another. The approach given here is due to Von Neumann and depends on the Riesz representation theorem for Hilbert space.

Definition 18.1 *Let μ and λ be two measures defined on a σ-algebra, $\mathcal{S}$, of subsets of a set, Ω. We say that λ is absolutely continuous with respect to μ and write $\lambda << \mu$ if $\lambda(E) = 0$ whenever $\mu(E) = 0$.*

Theorem 18.2 *(Radon Nikodym) Let λ and μ be finite measures defined on a σ-algebra, $\mathcal{S}$, of subsets of Ω. Suppose $\lambda << \mu$. Then there exists $f \in L^1(\Omega, \mu)$ such that $f(x) \geq 0$ and*

$$\lambda(E) = \int_E f \, d\mu.$$

Proof: Let $\Lambda : L^2(\Omega, \mu + \lambda) \to \mathbb{C}$ be defined by

$$\Lambda g = \int_\Omega g \, d\lambda.$$

By Holder's inequality,

$$|\Lambda g| \leq \left(\int_\Omega 1^2 d\lambda \right)^{1/2} \left(\int_\Omega |g|^2 \, d(\lambda + \mu) \right)^{1/2} = \lambda(\Omega)^{1/2} \, ||g||_2$$

and so $\Lambda \in (L^2(\Omega, \mu + \lambda))'$. By the Riesz representation theorem in Hilbert space, Theorem 15.11, there exists $h \in L^2(\Omega, \mu + \lambda)$ with

$$\Lambda g = \int_\Omega g \, d\lambda = \int_\Omega hg d(\mu + \lambda). \tag{18.1}$$

Letting $E =\{x \in \Omega : \operatorname{Im} h(x) > 0\}$, and letting $g = \mathcal{X}_E$, (18.1) implies

$$\lambda(E) = \int_E (\operatorname{Re} h + i \ \operatorname{Im} h) d(\mu + \lambda). \tag{18.2}$$

Since the left side of (18.2) is real, this shows $(\mu + \lambda)(E) = 0$. Similarly, if

$$E = \{x \in \Omega : \operatorname{Im} h(x) < 0\},$$

then $(\mu+\lambda)(E)=0$. Thus we may assume h is real-valued. Now let $E=\{x:h(x)<0\}$ and let $g=\mathcal{X}_E$. Then from (18.2)

$$\lambda(E)=\int_E h\ d(\mu+\lambda).$$

Since $h(x)<0$ on E, it follows $(\mu+\lambda)(E)=0$ or else the right side of this equation would be negative. Thus we can take $h\geq 0$. Now let $E=\{x:h(x)\geq 1\}$ and let $g=\mathcal{X}_E$. Then

$$\lambda(E)=\int_E h\ d(\mu+\lambda)\geq\mu(E)+\lambda(E).$$

Therefore $\mu(E)=0$. Since $\lambda<<\mu$, it follows that $\lambda(E)=0$ also. Thus we can assume

$$0\leq h(x)<1$$

for all x. From (18.1), whenever $g\in L^2(\Omega,\mu+\lambda)$,

$$\int_\Omega g(1-h)d\lambda=\int_\Omega hgd\mu. \tag{18.3}$$

Let $g(x)=\sum_{i=0}^n h^i(x)\mathcal{X}_E(x)$ in (18.3). This yields

$$\int_E(1-h^{n+1}(x))d\lambda=\int_E\sum_{i=1}^{n+1}h^i(x)d\mu. \tag{18.4}$$

Let $f(x)=\sum_{i=1}^\infty h^i(x)$ and use the Monotone Convergence theorem in (18.4) to let $n\to\infty$ and conclude

$$\lambda(E)=\int_E f\ d\mu.$$

We know $f\in L^1(\Omega,\mu)$ because λ is finite. This proves the theorem.

Note that the function, f is unique μ a.e. because, if g is another function which also serves to represent λ, we could consider the set,

$$E\equiv\{x:f(x)-g(x)>\varepsilon>0\}$$

and conclude that

$$0=\int_E f(x)-g(x)\,d\mu\geq\varepsilon\mu(E).$$

Since this holds for every $\varepsilon>0$, it must be the case that the set where f is larger than g has measure zero. Similarly, the set where g is larger than f has measure zero. The f in the theorem is sometimes denoted by

$$\frac{d\lambda}{d\mu}.$$

The next corollary is a generalization to σ finite measure spaces.

Corollary 18.3 *Suppose $\lambda<<\mu$ and there exist sets $S_n\in\mathcal{S}$ with*

$$S_n\cap S_m=\emptyset,\ \cup_{n=1}^\infty S_n=\Omega,$$

and $\lambda(S_n)$, $\mu(S_n)<\infty$. Then there exists $f\geq 0$, where f is μ measurable, and

$$\lambda(E)=\int_E f\ d\mu$$

for all $E\in\mathcal{S}$. The function f is $\mu+\lambda$ a.e. unique.

Proof: Let $\mathcal{S}_n = \{E \cap S_n : E \in \mathcal{S}\}$. Clearly $\mathcal{S}_n$ is a σ algebra of subsets of S_n, λ, μ are both finite measures on $\mathcal{S}_n$, and $\lambda << \mu$. Thus, by Theorem 18.2, there exists an $\mathcal{S}_n$ measurable function f_n, $f_n(x) \geq 0$, with

$$\lambda(E) = \int_E f_n d\mu$$

for all $E \in \mathcal{S}_n$. Define $f(x) = f_n(x)$ for $x \in S_n$. Then f is measurable because

$$f^{-1}((a, \infty]) = \cup_{n=1}^{\infty} f_n^{-1}((a, \infty]) \in \mathcal{S}.$$

Also, for $E \in \mathcal{S}$,

$$\begin{aligned} \lambda(E) &= \sum_{n=1}^{\infty} \lambda(E \cap S_n) = \sum_{n=1}^{\infty} \int \mathcal{X}_{E \cap S_n}(x) f_n(x) d\mu \\ &= \sum_{n=1}^{\infty} \int \mathcal{X}_{E \cap S_n}(x) f(x) d\mu \\ &= \int_E f \, d\mu. \end{aligned}$$

To see f is unique, suppose f_1 and f_2 both work and consider

$$E \equiv \{x : f_1(x) - f_2(x) > 0\}.$$

Then

$$0 = \lambda(E \cap S_n) - \lambda(E \cap S_n) = \int_{E \cap S_n} f_1(x) - f_2(x) d\mu.$$

Hence $\mu(E \cap S_n) = 0$ so $\mu(E) = 0$. Hence $\lambda(E) = 0$ also. Similarly

$$(\mu + \lambda)(\{x : f_2(x) - f_1(x) > 0\}) = 0.$$

This version of the Radon Nikodym theorem will suffice for most applications, but more general versions are available. To see one of these, one can read the treatment in Hewitt and Stromberg. This involves the notion of decomposable measure spaces, a generalization of $\sigma-$ finite.

18.2 Vector measures

The next topic will use the Radon Nikodym theorem. It is the topic of vector and complex measures. Here we are mainly concerned with complex measures although a vector measure can have values in any topological vector space.

Definition 18.4 *Let $(V, ||\cdot||)$ be a normed linear space and let $(\Omega, \mathcal{S})$ be a measure space. We call a function $\mu : \mathcal{S} \to V$ a vector measure if μ is countably additive. That is, if $\{E_i\}_{i=1}^{\infty}$ is a sequence of disjoint sets of $\mathcal{S}$,*

$$\mu(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i).$$

Definition 18.5 *Let $(\Omega, \mathcal{S})$ be a measure space and let μ be a vector measure defined on $\mathcal{S}$. A subset, $\pi(E)$, of $\mathcal{S}$ is called a partition of E if $\pi(E)$ consists of finitely many disjoint sets of $\mathcal{S}$ and $\cup \pi(E) = E$. Let*

$$|\mu|(E) = \sup\{ \sum_{F \in \pi(E)} ||\mu(F)|| : \pi(E) \text{ is a partition of } E\}.$$

$|\mu|$ is called the total variation of μ.

The next theorem may seem a little surprising. It states that, if finite, the total variation is a nonnegative measure.

Theorem 18.6 *If $|\mu|(\Omega) < \infty$, then $|\mu|$ is a measure on $\mathcal{S}$.*

Proof: Let $E_1 \cap E_2 = \emptyset$ and let $\{A_1^i \cdots A_{n_i}^i\} = \pi(E_i)$ with

$$|\mu|(E_i) - \varepsilon < \sum_{j=1}^{n_i} ||\mu(A_j^i)|| \ \ i = 1, 2.$$

Let $\pi(E_1 \cup E_2) = \pi(E_1) \cup \pi(E_2)$. Then

$$|\mu|(E_1 \cup E_2) \geq \sum_{F \in \pi(E_1 \cup E_2)} ||\mu(F)|| > |\mu|(E_1) + |\mu|(E_2) - 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that

$$|\mu|(E_1 \cup E_2) \geq |\mu|(E_1) + |\mu|(E_2). \tag{18.5}$$

Let $\{E_j\}_{j=1}^{\infty}$ be a sequence of disjoint sets of $\mathcal{S}$. Let $E_\infty = \cup_{j=1}^{\infty} E_j$ and let

$$\{A_1, \cdots, A_n\} = \pi(E_\infty)$$

be such that

$$|\mu|(E_\infty) - \varepsilon < \sum_{i=1}^{n} ||\mu(A_i)||.$$

But $||\mu(A_i)|| \leq \sum_{j=1}^{\infty} ||\mu(A_i \cap E_j)||$. Therefore,

$$\begin{aligned} |\mu|(E_\infty) - \varepsilon &< \sum_{i=1}^{n} \sum_{j=1}^{\infty} ||\mu(A_i \cap E_j)|| \\ &= \sum_{j=1}^{\infty} \sum_{i=1}^{n} ||\mu(A_i \cap E_j)|| \\ &\leq \sum_{j=1}^{\infty} |\mu|(E_j). \end{aligned}$$

The interchange in order of integration follows from Fubini's theorem or else Theorem 5.44 on the equality of double sums, and the last inequality follows because $A_1 \cap E_j, \cdots, A_n \cap E_j$ is a partition of E_j.

Since $\varepsilon > 0$ is arbitrary, this shows

$$|\mu|(\cup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} |\mu|(E_j).$$

By induction, (18.5) implies that whenever the E_i are disjoint,

$$|\mu|(\cup_{j=1}^{n} E_j) \geq \sum_{j=1}^{n} |\mu|(E_j).$$

Therefore,

$$\sum_{j=1}^{\infty} |\mu|(E_j) \geq |\mu|(\cup_{j=1}^{\infty} E_j) \geq |\mu|(\cup_{j=1}^{n} E_j) \geq \sum_{j=1}^{n} |\mu|(E_j).$$

Since n is arbitrary, this implies

$$|\mu|(\cup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} |\mu|(E_j)$$

which proves the theorem.

In the case where $V = \mathbb{C}$, it is automatically the case that $|\mu|(\Omega) < \infty$. This is proved in Rudin [24]. We will not need to use this fact, so it is left for the interested reader to look up.

Theorem 18.7 *Let $(\Omega, \mathcal{S})$ be a measure space and let $\lambda : \mathcal{S} \to \mathbb{C}$ be a complex vector measure with $|\lambda|(\Omega) < \infty$. Let $\mu : \mathcal{S} \to [0, \mu(\Omega)]$ be a finite measure such that $\lambda << \mu$. Then there exists a unique $f \in L^1(\Omega)$ such that for all $E \in \mathcal{S}$,*

$$\int_E f d\mu = \lambda(E).$$

Proof: It is clear that $\operatorname{Re}\lambda$ and $\operatorname{Im}\lambda$ are real-valued vector measures on $\mathcal{S}$. Since $|\lambda|(\Omega) < \infty$, it follows easily that $|\operatorname{Re}\lambda|(\Omega)$ and $|\operatorname{Im}\lambda|(\Omega) < \infty$. Therefore, each of

$$\frac{|\operatorname{Re}\lambda| + \operatorname{Re}\lambda}{2}, \ \frac{|\operatorname{Re}\lambda| - \operatorname{Re}(\lambda)}{2}, \ \frac{|\operatorname{Im}\lambda| + \operatorname{Im}\lambda}{2}, \text{ and } \frac{|\operatorname{Im}\lambda| - \operatorname{Im}(\lambda)}{2}$$

are finite measures on $\mathcal{S}$. It is also clear that each of these finite measures are absolutely continuous with respect to μ. Thus there exist unique nonnegative functions in $L^1(\Omega)$, $f_1, \ f_2, \ g_1, \ g_2$ such that for all $E \in \mathcal{S}$,

$$\begin{aligned}
\frac{1}{2}(|\operatorname{Re}\lambda| + \operatorname{Re}\lambda)(E) &= \int_E f_1 d\mu, \\
\frac{1}{2}(|\operatorname{Re}\lambda| - \operatorname{Re}\lambda)(E) &= \int_E f_2 d\mu, \\
\frac{1}{2}(|\operatorname{Im}\lambda| + \operatorname{Im}\lambda)(E) &= \int_E g_1 d\mu, \\
\frac{1}{2}(|\operatorname{Im}\lambda| - \operatorname{Im}\lambda)(E) &= \int_E g_2 d\mu.
\end{aligned}$$

Now let $f = f_1 - f_2 + i(g_1 - g_2)$.

The following corollary is about representing a vector measure in terms of its total variation.

Corollary 18.8 *Let λ be a complex vector measure with $|\lambda|(\Omega) < \infty$. Then there exists a unique $f \in L^1(\Omega)$ such that $\lambda(E) = \int_E f d|\lambda|$. Furthermore, $|f| = 1$ $|\lambda|$ a.e. This is called the polar decomposition of λ.*

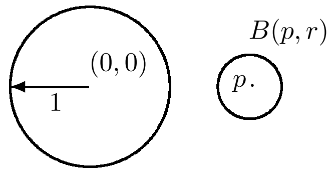
Proof: First we note that $\lambda << |\lambda|$ and so such an L^1 function exists and is unique. We have to show $|f| = 1$ a.e.

Lemma 18.9 *Suppose $(\Omega, \mathcal{S}, \mu)$ is a measure space and f is a function in $L^1(\Omega, \mu)$ with the property that*

$$\left|\int_E f \, d\mu\right| \leq \mu(E)$$

for all $E \in \mathcal{S}$. Then $|f| \leq 1$ a.e.

Proof of the lemma: Consider the following picture.



where $B(p,r) \cap B(0,1) = \emptyset$. Let $E = f^{-1}(B(p,r))$. If $\mu(E) \neq 0$ then

$$\begin{aligned}\left|\frac{1}{\mu(E)}\int_E f\, d\mu - p\right| &= \left|\frac{1}{\mu(E)}\int_E (f-p)d\mu\right| \\ &\leq \frac{1}{\mu(E)}\int_E |f-p|d\mu < r.\end{aligned}$$

Hence

$$|\frac{1}{\mu(E)}\int_E fd\mu| > 1,$$

contradicting the assumption of the lemma. It follows $\mu(E) = 0$. Since $\{z \in \mathbb{C} : |z| > 1\}$ can be covered by countably many such balls, it follows that

$$\mu(f^{-1}(\{z \in \mathbb{C} : |z| > 1\})) = 0.$$

Thus $|f(x)| \leq 1$ a.e. as claimed. This proves the lemma.

To finish the proof of Corollary 18.8, if $|\lambda|(E) \neq 0$,

$$\left|\frac{\lambda(E)}{|\lambda|(E)}\right| = \left|\frac{1}{|\lambda|(E)}\int_E f\, d|\lambda|\right| \leq 1.$$

Therefore $|f| \leq 1,\ |\lambda|$ a.e. Now let

$$E_n = \{x \in \Omega : |f(x)| \leq 1 - \frac{1}{n}\}.$$

Let $\{F_1, \cdots, F_m\}$ be a partition of E_n such that

$$\sum_{i=1}^m |\lambda(F_i)| \geq |\lambda|(E_n) - \varepsilon.$$

Then

$$\begin{aligned}|\lambda|(E_n) - \varepsilon &\leq \sum_{i=1}^m |\lambda(F_i)| \leq \sum_{i=1}^m |\int_{F_i} fd\,|\lambda|\,| \\ &\leq \left(1 - \frac{1}{n}\right)\sum_{i=1}^m |\lambda|\,(F_i) \\ &= \left(1 - \frac{1}{n}\right)|\lambda|\,(E_n)\end{aligned}$$

and so

$$\frac{1}{n}\,|\lambda|\,(E_n) \leq \varepsilon.$$

Since ε was arbitrary, this shows that $|\lambda|\,(E_n) = 0$. But $\{x \in \Omega : |f(x)| < 1\} = \cup_{n=1}^\infty E_n$. So $|\lambda|(\{x \in \Omega : |f(x)| < 1\}) = 0$. This proves Corollary 18.8.

18.3 Representation theorems for the dual space of L^p

In Chapter 14 we discussed the definition of a Banach space and the dual space of a Banach space. We also saw in Chapter 12 that the L^p spaces are Banach spaces. The next topic deals with the dual space of L^p for $p \geq 1$ in the case where the measure space is σ finite or finite.

Theorem 18.10 *(Riesz representation theorem) Let $p > 1$ and let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space. If $\Lambda \in (L^p(\Omega))'$, then there exists a unique $h \in L^q(\Omega)$ ($\frac{1}{p} + \frac{1}{q} = 1$) such that*

$$\Lambda f = \int_\Omega h f d\mu.$$

Proof: (Uniqueness) If h_1 and h_2 both represent Λ, consider

$$f = |h_1 - h_2|^{q-2}(\overline{h_1} - \overline{h_2}),$$

where $\overline{h}$ denotes complex conjugation. By Holder's inequality, it is easy to see that $f \in L^p(\Omega)$. Thus

$$0 = \Lambda f - \Lambda f =$$

$$\int h_1 |h_1 - h_2|^{q-2}(\overline{h_1} - \overline{h_2}) - h_2 |h_1 - h_2|^{q-2}(\overline{h_1} - \overline{h_2}) d\mu$$

$$= \int |h_1 - h_2|^q d\mu.$$

Therefore $h_1 = h_2$ and this proves uniqueness.

Now let $\lambda(E) = \Lambda(\mathcal{X}_E)$. Let $A_1, \cdots, A_n$ be a partition of Ω.

$$|\Lambda \mathcal{X}_{A_i}| = w_i(\Lambda \mathcal{X}_{A_i}) = \Lambda(w_i \mathcal{X}_{A_i})$$

for some $w_i \in \mathbb{C}$, $|w_i| = 1$. Thus

$$\sum_{i=1}^n |\lambda(A_i)| = \sum_{i=1}^n |\Lambda(\mathcal{X}_{A_i})| = \Lambda(\sum_{i=1}^n w_i \mathcal{X}_{A_i})$$

$$\leq ||\Lambda||(\int |\sum_{i=1}^n w_i \mathcal{X}_{A_i}|^p d\mu)^{\frac{1}{p}} = ||\Lambda||(\int_\Omega d\mu)^{\frac{1}{p}} = ||\Lambda||\mu(\Omega)^{\frac{1}{p}}.$$

Therefore $|\lambda|(\Omega) < \infty$. Also, if $\{E_i\}_{i=1}^\infty$ is a sequence of disjoint sets of $\mathcal{S}$, let

$$F_n = \cup_{i=1}^n E_i, \ F = \cup_{i=1}^\infty E_i.$$

Then by the Dominated Convergence theorem,

$$||\mathcal{X}_{F_n} - \mathcal{X}_F||_p \to 0.$$

Therefore,

$$\lambda(F) = \Lambda(\mathcal{X}_F) = \lim_{n\to\infty} \Lambda(\mathcal{X}_{F_n}) = \lim_{n\to\infty} \sum_{k=1}^n \Lambda(\mathcal{X}_{E_k}) = \sum_{k=1}^\infty \lambda(E_k).$$

This shows λ is a complex measure with $|\lambda|$ finite. It is also clear that $\lambda << \mu$. Therefore, by the Radon Nikodym theorem, there exists $h \in L^1(\Omega)$ with

$$\lambda(E) = \int_E h d\mu = \Lambda(\mathcal{X}_E).$$

Now let $s = \sum_{i=1}^m c_i \mathcal{X}_{E_i}$ be a simple function. We have

$$\Lambda(s) = \sum_{i=1}^m c_i \Lambda(\mathcal{X}_{E_i}) = \sum_{i=1}^m c_i \int_{E_i} h d\mu = \int h s d\mu. \tag{18.6}$$

Claim: If f is uniformly bounded and measurable, then

$$\Lambda(f) = \int h f d\mu.$$

Proof of claim: Since f is bounded and measurable, there exists a sequence of simple functions, $\{s_n\}$ which converges to f pointwise and in $L^p(\Omega)$. Then

$$\Lambda(f) = \lim_{n\to\infty} \Lambda(s_n) = \lim_{n\to\infty} \int h s_n d\mu = \int h f d\mu,$$

the first equality holding because of continuity of Λ and the second equality holding by the dominated convergence theorem.

Let $E_n = \{x : |h(x)| \leq n\}$. Thus $|h\mathcal{X}_{E_n}| \leq n$. Then

$$|h\mathcal{X}_{E_n}|^{q-2}(\overline{h}\mathcal{X}_{E_n}) \in L^p(\Omega).$$

By the claim, it follows that

$$||h\mathcal{X}_{E_n}||_q^q = \int h|h\mathcal{X}_{E_n}|^{q-2}(\overline{h}\mathcal{X}_{E_n})d\mu = \Lambda(|h\mathcal{X}_{E_n}|^{q-2}(\overline{h}\mathcal{X}_{E_n}))$$

$$\leq ||\Lambda|| \; || \, |h\mathcal{X}_{E_n}|^{q-2}\overline{h}\mathcal{X}_{E_n}||_p = ||\Lambda|| \; ||h\mathcal{X}_{E_n}||_q^{\frac{q}{p}}.$$

Therefore, since $q - \frac{q}{p} = 1$, it follows that

$$||h\mathcal{X}_{E_n}||_q \leq ||\Lambda||.$$

Letting $n \to \infty$, the Monotone Convergence theorem implies

$$||h||_q \leq ||\Lambda||. \tag{18.7}$$

Now that h has been shown to be in $L^q(\Omega)$, it follows from (18.6) and the density of the simple functions, Theorem 12.8, that

$$\Lambda f = \int h f d\mu$$

for all $f \in L^p(\Omega)$. This proves Theorem 18.10.

Corollary 18.11 *If h is the function of Theorem 18.10 representing Λ, then $||h||_q = ||\Lambda||$.*

Proof: $||\Lambda|| = \sup\{\int hf : ||f||_p \leq 1\} \leq ||h||_q \leq ||\Lambda||$ by (18.7), and Holder's inequality.

To represent elements of the dual space of $L^1(\Omega)$, we need another Banach space.

Definition 18.12 *Let $(\Omega, \mathcal{S}, \mu)$ be a measure space. $L^\infty(\Omega)$ is the vector space of measurable functions such that for some $M > 0$, $|f(x)| \leq M$ for all x outside of some set of measure zero ($|f(x)| \leq M$ a.e.). We define $f = g$ when $f(x) = g(x)$ a.e. and $||f||_\infty \equiv \inf\{M : |f(x)| \leq M \text{ a.e.}\}$.*

Theorem 18.13 *$L^\infty(\Omega)$ is a Banach space.*

Proof: It is clear that $L^\infty(\Omega)$ is a vector space and it is routine to verify that $|| \, ||_\infty$ is a norm. To verify completeness, let $\{f_n\}$ be a Cauchy sequence in $L^\infty(\Omega)$. Let

$$|f_n(x) - f_m(x)| \leq ||f_n - f_m||_\infty$$

for all $x \notin E_{nm}$, a set of measure 0. Let $E = \cup_{n,m} E_{nm}$. Thus $\mu(E) = 0$ and for each $x \notin E$, $\{f_n(x)\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{C}$. Let

$$f(x) = \begin{cases} 0 \text{ if } x \in E \\ \lim_{n\to\infty} f_n(x) \text{ if } x \notin E \end{cases} = \lim_{n\to\infty} \mathcal{X}_{E^C}(x) f_n(x).$$

Then f is clearly measurable because it is the limit of measurable functions. If

$$F_n = \{x : |f_n(x)| > ||f_n||_\infty\}$$

and $F = \cup_{n=1}^\infty F_n$, it follows $\mu(F) = 0$ and that for $x \notin F \cup E$,

$$|f(x)| \leq \lim \inf_{n\to\infty} |f_n(x)| \leq \lim \inf_{n\to\infty} ||f_n||_\infty < \infty$$

because $\{f_n\}$ is a Cauchy sequence. Thus $f \in L^\infty(\Omega)$. Let n be large enough that whenever $m > n$,

$$||f_m - f_n||_\infty < \varepsilon.$$

Thus, if $x \notin E$,

$$\begin{aligned} |f(x) - f_n(x)| &= \lim_{m\to\infty} |f_m(x) - f_n(x)| \\ &\leq \lim_{m\to\infty} \inf \, ||f_m - f_n||_\infty < \varepsilon. \end{aligned}$$

Hence $||f - f_n||_\infty < \varepsilon$ for all n large enough. This proves the theorem.

The next theorem is the Riesz representation theorem for $\left(L^1(\Omega)\right)'$.

Theorem 18.14 *(Riesz representation theorem) Let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space. If $\Lambda \in (L^1(\Omega))'$, then there exists a unique $h \in L^\infty(\Omega)$ such that*

$$\Lambda(f) = \int_\Omega hf \, d\mu$$

for all $f \in L^1(\Omega)$.

Proof: Just as in the proof of Theorem 18.10, there exists a unique $h \in L^1(\Omega)$ such that for all simple functions, s,

$$\Lambda(s) = \int hs \, d\mu. \tag{18.8}$$

To show $h \in L^\infty(\Omega)$, let $\varepsilon > 0$ be given and let

$$E = \{x : |h(x)| \geq ||\Lambda|| + \varepsilon\}.$$

Let $|k| = 1$ and $hk = |h|$. Since the measure space is finite, $k \in L^1(\Omega)$. Let $\{s_n\}$ be a sequence of simple functions converging to k in $L^1(\Omega)$, and pointwise. Also let $|s_n| \leq 1$. Therefore

$$\Lambda(k\mathcal{X}_E) = \lim_{n\to\infty} \Lambda(s_n\mathcal{X}_E) = \lim_{n\to\infty} \int_E hs_n d\mu = \int_E hk d\mu$$

where the last equality holds by the Dominated Convergence theorem. Therefore,

$$\begin{aligned} ||\Lambda||\mu(E) &\geq |\Lambda(k\mathcal{X}_E)| = |\int_\Omega hk\mathcal{X}_E d\mu| = \int_E |h| d\mu \\ &\geq (||\Lambda|| + \varepsilon)\mu(E). \end{aligned}$$

It follows that $\mu(E) = 0$. Since $\varepsilon > 0$ was arbitrary, $||\Lambda|| \geq ||h||_\infty$. Now we have seen that $h \in L^\infty(\Omega)$, the density of the simple functions and (18.8) imply

$$\Lambda f = \int_\Omega hf d\mu \ , \ ||\Lambda|| \geq ||h||_\infty. \tag{18.9}$$

This proves the existence part of the theorem. To verify uniqueness, suppose h_1 and h_2 both represent Λ and let $f \in L^1(\Omega)$ be such that $|f| \leq 1$ and $f(h_1 - h_2) = |h_1 - h_2|$. Then

$$0 = \Lambda f - \Lambda f = \int (h_1 - h_2) f d\mu = \int |h_1 - h_2| d\mu.$$

Thus $h_1 = h_2$.

Corollary 18.15 *If h is the function in $L^\infty(\Omega)$ representing $\Lambda \in (L^1(\Omega))'$, then $||h||_\infty = ||\Lambda||$.*

Proof: $||\Lambda|| = \sup\{|\int hf d\mu| : ||f||_1 \leq 1\} \leq ||h||_\infty \leq ||\Lambda||$ by (18.9).

Next we extend these results to the σ finite case.

Lemma 18.16 *Let $(\Omega, \mathcal{S}, \mu)$ be a measure space and suppose there exists a measurable function, r such that $r(x) > 0$ for all x, there exists M such that $|r(x)| < M$ for all x, and $\int r d\mu < \infty$. Then for*

$$\Lambda \in (L^p(\Omega, \mu))', \ p \geq 1,$$

there exists a unique $h \in L^q(\Omega, \mu)$, $L^\infty(\Omega, \mu)$ if $p = 1$ such that

$$\Lambda f = \int hf d\mu.$$

Also $||h|| = ||\Lambda||$. ($||h|| = ||h||_q$ if $p > 1$, $||h||_\infty$ if $p = 1$). Here

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Proof: We present the proof in the case where $p > 1$ and leave the case $p = 1$ to the reader. It involves routine modifications of the case presented. Define a new measure $\widetilde{\mu}$, according to the rule

$$\widetilde{\mu}(E) \equiv \int_E r d\mu. \tag{18.10}$$

Thus $\widetilde{\mu}$ is a finite measure on $\mathcal{S}$. Now define a mapping, $\eta : L^p(\Omega, \widetilde{\mu}) \to L^p(\Omega, \mu)$ by

$$\eta f = r^{\frac{1}{p}} f.$$

Then η is one to one and onto. Also it is routine to show

$$||\eta f||_{L^p(\mu)} = ||f||_{L^p(\widetilde{\mu})}. \tag{18.11}$$

Consider the diagram below which is descriptive of the situation.

$$\begin{array}{cccc} & & \eta^* & \\ L^q(\widetilde{\mu}) & L^p(\widetilde{\mu})' & \leftarrow & L^p(\mu)', \Lambda \\ & & \eta & \\ & L^p(\widetilde{\mu}) & \rightarrow & L^p(\mu) \end{array}$$

By the Riesz representation theorem for finite measures, there exists a unique $h \in L^q(\Omega, \widetilde{\mu})$ such that for all $f \in L^p(\widetilde{\mu})$,

$$\begin{aligned} \Lambda(\eta f) &= \eta^* \Lambda(f) \equiv \int_\Omega f h d\widetilde{\mu}, \\ ||h||_{L^q(\widetilde{\mu})} &= ||\eta^* \Lambda|| = ||\Lambda||, \end{aligned} \tag{18.12}$$

the last equation holding because of (18.11). But from (18.10),

$$\begin{aligned} \int_\Omega f h d\widetilde{\mu} &= \int_\Omega \left(r^{\frac{1}{p}} f\right)\left(h r^{\frac{1}{q}}\right) d\mu \\ &= \int_\Omega (\eta f)\left(h r^{\frac{1}{q}}\right) d\mu \end{aligned}$$

and so we see that

$$\Lambda(\eta f) = \int_\Omega (\eta f)\left(h r^{\frac{1}{q}}\right) d\mu$$

for all $f \in L^p(\widetilde{\mu})$. Since η is onto, this shows $hr^{\frac{1}{q}}$ represents Λ as claimed. It only remains to verify $||\Lambda|| = \left|\left|hr^{\frac{1}{q}}\right|\right|_q$. However, this equation comes immediately form (18.10) and (18.12). This proves the lemma.

A situation in which the conditions of the lemma are satisfied is the case where the measure space is σ finite. This allows us to state the following theorem.

Theorem 18.17 *(Riesz representation theorem) Let $(\Omega, \mathcal{S}, \mu)$ be σ finite and let*

$$\Lambda \in (L^p(\Omega, \mu))',\ p \geq 1.$$

Then there exists a unique $h \in L^q(\Omega, \mu),\ L^\infty(\Omega, \mu)$ if $p = 1$ such that

$$\Lambda f = \int h f d\mu.$$

Also $||h|| = ||\Lambda||$. ($||h|| = ||h||_q$ if $p > 1$, $||h||_\infty$ if $p = 1$). Here

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Proof: Let $\{\Omega_n\}$ be a sequence of disjoint elements of $\mathcal{S}$ having the property that

$$0 < \mu(\Omega_n) < \infty,\ \cup_{n=1}^\infty \Omega_n = \Omega.$$

Define

$$r(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \mathcal{X}_{\Omega_n}(x)\ \mu(\Omega_n)^{-1},\ \ \widetilde{\mu}(E) = \int_E r d\mu.$$

Thus

$$\int_\Omega r d\mu = \widetilde{\mu}(\Omega) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

so $\widetilde{\mu}$ is a finite measure. By the above lemma we obtain the existence part of the conclusion of the theorem. Uniqueness is done as before.

With the Riesz representation theorem, it is easy to show that

$$L^p(\Omega),\ p > 1$$

is a reflexive Banach space.

Theorem 18.18 *For $(\Omega, \mathcal{S}, \mu)$ a σ finite measure space and $p > 1$, $L^p(\Omega)$ is reflexive.*

Proof: Let $\delta_r : (L^r(\Omega))' \to L^{r'}(\Omega)$ be defined for $\frac{1}{r} + \frac{1}{r'} = 1$ by

$$\int (\delta_r \Lambda) g\ d\mu = \Lambda g$$

for all $g \in L^r(\Omega)$. From Theorem 18.17 δ_r is 1-1, onto, continuous and linear. By the Open Map theorem, δ_r^{-1} is also 1-1, onto, and continuous ($\delta_r \Lambda$ equals the representor of Λ). Thus δ_r^* is also 1-1, onto, and continuous by Corollary 14.28. Now observe that $J = \delta_p^* \circ \delta_q^{-1}$. To see this, let $z^* \in (L^q)'$, $y^* \in (L^p)'$,

$$\begin{aligned} \delta_p^* \circ \delta_q^{-1}(\delta_q z^*)(y^*) &= (\delta_p^* z^*)(y^*) \\ &= z^*(\delta_p y^*) \\ &= \int (\delta_q z^*)(\delta_p y^*) d\mu, \end{aligned}$$

$$\begin{aligned} J(\delta_q z^*)(y^*) &= y^*(\delta_q z^*) \\ &= \int (\delta_p y^*)(\delta_p z^*) d\mu. \end{aligned}$$

Therefore $\delta_p^* \circ \delta_q^{-1} = J$ on $\delta_q (L^q)' = L^p$. But the two δ maps are onto and so J is also onto.

18.4 Riesz Representation theorem for non σ finite measure spaces

It is not necessary to assume μ is either finite or σ finite to establish the Riesz representation theorem for $1 < p < \infty$. This involves the notion of uniform convexity. First we recall Clarkson's inequality for $p \geq 2$. This was Problem 24 in Chapter 12.

Lemma 18.19 *(Clarkson inequality $p \geq 2$) For $p \geq 2$,*

$$||\frac{f+g}{2}||_p^p + ||\frac{f-g}{2}||_p^p \leq \frac{1}{2}||f||_p^p + \frac{1}{2}||g||_p^p.$$

Definition 18.20 *A Banach space, X, is said to be uniformly convex if whenever $||x_n|| \leq 1$ and $||\frac{x_n+x_m}{2}|| \to 1$ as $n, m \to \infty$, then $\{x_n\}$ is a Cauchy sequence and $x_n \to x$ where $||x|| = 1$.*

Observe that Clarkson's inequality implies L^p is uniformly convex for all $p \geq 2$. Uniformly convex spaces have a very nice property which is described in the following lemma. Roughly, this property is that any element of the dual space achieves its norm at some point of the closed unit ball.

Lemma 18.21 *Let X be uniformly convex and let $L \in X'$. Then there exists $x \in X$ such that*

$$||x|| = 1,\ Lx = ||L||.$$

Proof: Let $||\widetilde{x}_n|| \leq 1$ and $|L\widetilde{x}_n| \to ||L||$. Let $x_n = w_n \widetilde{x}_n$ where $|w_n| = 1$ and

$$w_n L\widetilde{x}_n = |L\widetilde{x}_n|.$$

Thus $Lx_n = |Lx_n| = |L\widetilde{x}_n| \to ||L||$.

$$Lx_n \to ||L||,\ ||x_n|| \leq 1.$$

We can assume, without loss of generality, that

$$Lx_n = |Lx_n| \geq \frac{||L||}{2}$$

and $L \neq 0$.

Claim $||\frac{x_n+x_m}{2}|| \to 1$ as $n, m \to \infty$.

Proof of Claim: Let n, m be large enough that $Lx_n,\ Lx_m \geq ||L|| - \frac{\varepsilon}{2}$ where $0 < \varepsilon$. Then $||x_n + x_m|| \neq 0$ because if it equals 0, then $x_n = -x_m$ so $-Lx_n = Lx_m$ but both Lx_n and Lx_m are positive. Therefore we can consider $\frac{x_n+x_m}{||x_n+x_m||}$ a vector of norm 1. Thus,

$$||L|| \geq |L\frac{(x_n + x_m)}{||x_n + x_m||}| \geq \frac{2||L|| - \varepsilon}{||x_n + x_m||}.$$

Hence

$$||x_n + x_m||\ ||L|| \geq 2||L|| - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\lim_{n,m\to\infty} ||x_n + x_m|| = 2$. This proves the claim.

By uniform convexity, $\{x_n\}$ is Cauchy and $x_n \to x,\ ||x|| = 1$. Thus $Lx = \lim_{n\to\infty} Lx_n = ||L||$. This proves Lemma 18.21.

The proof of the Riesz representation theorem will be based on the following lemma which says that if you can show certain things, then you can represent a linear functional.

Lemma 18.22 *(McShane) Let X be a complex normed linear space and let $L \in X'$. Suppose there exists $x \in X,\ ||x|| = 1$ with $Lx = ||L|| \neq 0$. Let $y \in X$ and let $\psi_y(t) = ||x + ty||$ for $t \in \mathbb{R}$. Suppose $\psi_y'(0)$ exists for each $y \in X$. Then for all $y \in X$,*

$$\psi_y'(0) + i\psi_{-iy}'(0) = ||L||^{-1} Ly.$$

Proof: Suppose first that $||L|| = 1$. Then

$$L(x + t(y - L(y)x)) = Lx = 1 = ||L||.$$

Therefore, $||x + t(y - L(y)x)|| \geq 1$. Also for small $t, |L(y)t| < 1$, and so

$$1 \leq ||x + t(y - L(y)x)|| = ||(1 - L(y)t)x + ty||$$

$$\leq |1 - L(y)\,t| \left|\left| x + \frac{t}{1 - L(y)\,t} y \right|\right|.$$

This implies

$$\frac{1}{|1 - tL(y)|} \leq ||x + \frac{t}{1 - L(y)t} y||. \tag{18.13}$$

Using the formula for the sum of a geometric series,

$$\frac{1}{1 - tLy} = 1 + tLy + o(t)$$

where $\lim_{t \to 0} o(t)(t^{-1}) = 0$. Using this in (18.13), we obtain

$$|1 + tL(y) + o(t)| \leq ||x + ty + o(t)||$$

Now if $t > 0$, since $||x|| = 1$, we have

$$\begin{aligned} (\psi_y(t) - \psi_y(0))t^{-1} &= (||x + ty|| - ||x||)t^{-1} \\ &\geq (|1 + t\,L(y)| - 1)t^{-1} + \frac{o(t)}{t} \\ &\geq \operatorname{Re} L(y) + \frac{o(t)}{t}. \end{aligned}$$

If $t < 0$,

$$(\psi_y(t) - \psi_y(0))t^{-1} \leq \operatorname{Re} L(y) + \frac{o(t)}{t}.$$

Since $\psi_y'(0)$ is assumed to exist, this shows

$$\psi_y'(0) = \operatorname{Re} L(y). \tag{18.14}$$

Now

$$Ly = \operatorname{Re} L(y) + i \ \operatorname{Im} L(y)$$

so

$$L(-iy) = -i(Ly) = -i \ \operatorname{Re} L(y) + \operatorname{Im} L(y)$$

and

$$L(-iy) = \operatorname{Re} L(-iy) + i \operatorname{Im} L(-iy).$$

Hence

$$\operatorname{Re} L(-iy) = \operatorname{Im} L(y).$$

Consequently, by (18.14)

$$Ly = \operatorname{Re} L(y) + i \ \operatorname{Im} L(y) = \operatorname{Re} L(y) + i \operatorname{Re} L(-iy)$$

$$= \psi_y'(0) + i\, \psi_{-iy}'(0).$$

This proves the lemma when $||L|| = 1$. For arbitrary $L \neq 0$, let $Lx = ||L||, ||x|| = 1$. Then from above, if $L_1 y \equiv ||L||^{-1} L(y)$, $||L_1|| = 1$ and so from what was just shown,

$$L_1(y) = \frac{L(y)}{||L||} = \psi_y'(0) + i\psi_{-iy}(0)$$

and this proves McShane's lemma.

We will use the uniform convexity and this lemma to prove the Riesz representation theorem next. Let $p \geq 2$ and let $\eta : L^q \to (L^p)'$ be defined by

$$\eta(g)(f) = \int_\Omega gf \, d\mu. \tag{18.15}$$

Theorem 18.23 *(Riesz representation theorem $p \geq 2$) The map η is 1-1, onto, continuous, and*

$$||\eta g|| = ||g||, \; ||\eta|| = 1.$$

Proof: Obviously η is linear. Suppose $\eta g = 0$. Then $0 = \int gf \, d\mu$ for all $f \in L^p$. Let

$$f = |g|^{q-2}\overline{g}.$$

Then $f \in L^p$and so $0 = \int |g|^q d\mu$. Hence $g = 0$ and η is 1-1. That $\eta g \in (L^p)'$ is obvious from the Holder inequality. In fact,

$$|\eta(g)(f)| \leq ||g||_q ||f||_p,$$

and so $||\eta(g)|| \leq ||g||_q$. To see that equality holds, let

$$f = |g|^{q-2}\overline{g} \; ||g||_q^{1-q}.$$

Then $||f||_p = 1$ and

$$\eta(g)(f) = \int_\Omega |g|^q d\mu ||g||_q^{1-q} = ||g||_q.$$

Thus $||\eta|| = 1$. It remains to show η is onto. Let $L \in (L^p)'$. We need show $L = \eta g$ for some $g \in L^q$. Without loss of generality, we may assume $L \neq 0$. Let

$$Lg = ||L||, \; g \in L^p, \; ||g|| = 1.$$

We can assert the existence of such a g by Lemma 18.21. For $f \in L^p$,

$$\psi_f(t) \equiv ||g + tf||_p \equiv \phi_f(t)^{\frac{1}{p}}.$$

We show $\phi_f'(0)$ exists. Let $[g = 0]$ denote the set $\{x : g(x) = 0\}$.

$$\frac{\phi_f(t) - \phi_f(0)}{t} =$$

$$\frac{1}{t}\int (|g + tf|^p - |g|^p) d\mu = \frac{1}{t}\int_{[g=0]} |t|^p \, |f|^p d\mu$$

$$+ \int_{[g\neq 0]} p|g(x) + s(x)f(x)|^{p-2} \operatorname{Re}[(g(x) + s(x)f(x))\bar{f}(x)] d\mu \tag{18.16}$$

where the Mean Value theorem is being used on the function $t \to |g(x) + tf(x)|^p$ and $s(x)$ is between 0 and t, the integrand in the second integral of (18.16) equaling

$$\frac{1}{t}(|g(x) + tf(x)|^p - |g(x)|^p).$$

Now if $|t| < 1$, the integrand in the last integral of (18.16) is bounded by

$$p\left[\frac{(|g(x)| + |f(x)|)^p}{q} + \frac{|f(x)|^p}{p}\right]$$

which is a function in L^1 since f, g are in L^p(we used the inequality $ab \leq \frac{a^q}{q} + \frac{b^p}{p}$). Because of this, we can apply the Dominated Convergence theorem and obtain

$$\phi_f'(0) = p\int |g(x)|^{p-2}\,\mathrm{Re}(g(x)\bar{f}(x))d\mu.$$

Hence

$$\psi_f'(0) = ||g||^{\frac{-p}{q}}\int |g(x)|^{p-2}\,\mathrm{Re}(g(x)\bar{f}(x))d\mu.$$

Note $\frac{1}{p} - 1 = -\frac{1}{q}$. Therefore,

$$\psi_{-if}'(0) = ||g||^{\frac{-p}{q}}\int |g(x)|^{p-2}\,\mathrm{Re}(ig(x)\bar{f}(x))d\mu.$$

But $\mathrm{Re}(ig\bar{f}) = \mathrm{Im}(-g\bar{f})$ and so by the McShane lemma,

$$\begin{aligned}
Lf &= ||L||\;||g||^{\frac{-p}{q}}\int |g(x)|^{p-2}[\mathrm{Re}(g(x)\bar{f}(x)) + i\;\mathrm{Re}(ig(x)\bar{f}(x))]d\mu \\
&= ||L||\;||g||^{\frac{-p}{q}}\int |g(x)|^{p-2}[\mathrm{Re}(g(x)\bar{f}(x)) + i\;\mathrm{Im}(-g(x)\bar{f}(x))]d\mu \\
&= ||L||\;||g||^{\frac{-p}{q}}\int |g(x)|^{p-2}\overline{g}(x)f(x)d\mu.
\end{aligned}$$

This shows that

$$L = \eta(||L||\;||g||^{\frac{-p}{q}}|g|^{p-2}\overline{g})$$

and verifies η is onto. This proves the theorem.

To prove the Riesz representation theorem for $1 < p < 2$, one can verify that L^p is uniformly convex and then repeat the above argument. Note that no reference to $p \geq 2$ was used in the proof. Unfortunately, this requires Clarkson's Inequalities for $p \in (1, 2)$ which are more technical than the case where $p \geq 2$. To see this done see Hewitt & Stromberg [15] or Ray [22]. Here we take a different approach using the Milman theorem which states that uniform convexity implies the space is Reflexive.

Theorem 18.24 *(Riesz representation theorem) Let $1 < p < \infty$ and let $\eta : L^q \to (L^p)'$ be given by (18.15). Then η is 1-1, onto, and $||\eta g|| = ||g||$.*

Proof: Everything is the same as the proof of Theorem 18.23 except for the assertion that η is onto. Suppose $1 < p < 2$. (The case $p \geq 2$ was done in Theorem 18.23.) Then $q > 2$ and so we know from Theorem 18.23 that $\overline{\eta} : L^p \to (L^q)'$ defined as

$$\overline{\eta}f(g) \equiv \int_\Omega fgd\mu$$

is onto and $||\overline{\eta}f|| = ||f||$. Then $\overline{\eta}^* : (L^q)'' \to (L^p)'$ is also 1-1, onto, and $||\overline{\eta}^* L|| = ||L||$. By Milman's theorem, J is onto from $L^q \to (L^q)''$. This occurs because of the uniform convexity of L^q which follows from Clarkson's inequality. Thus both maps in the following diagram are 1-1 and onto.

$$L^q \xrightarrow{J} (L^q)'' \xrightarrow{\overline{\eta}^*} (L^p)'.$$

Now if $g \in L^q,\ f \in L^p$, then

$$\overline{\eta}^* J(g)(f) = Jg(\overline{\eta}f) = (\overline{\eta}f)(g) = \int_\Omega fg\ d\mu.$$

Thus if $\eta : L^q \to (L^p)'$ is the mapping of (18.15), this shows $\eta = \overline{\eta}^* J$. Also

$$||\eta g|| = ||\eta^* Jg|| = ||Jg|| = ||g||.$$

This proves the theorem.

In the case where $p = 1$, it is also possible to give the Riesz representation theorem in a more general context than σ finite spaces. To see this done, see Hewitt and Stromberg [15]. The dual space of L^∞ has also been studied. See Dunford and Schwartz [9].

18.5 The dual space of $C(X)$

Next we represent the dual space of $C(X)$ where X is a compact Hausdorff space. It will turn out to be a space of measures. The theorems we will present hold for X a compact or locally compact Hausdorff space but we will only give the proof in the case where X is also a metric space. This is because the proof we use depends on the Riesz representation theorem for positive linear functionals and we only gave such a proof in the special case where X was a metric space. With the more general theorem in hand, the arguments give here will apply unchanged to the more general setting. Thus X will be a compact metric space in what follows.

Let $L \in C(X)'$. Also denote by $C^+(X)$ the set of nonnegative continuous functions defined on X. Define for $f \in C^+(X)$

$$\lambda(f) = \sup\{|Lg| : |g| \le f\}.$$

Note that $\lambda(f) < \infty$ because $|Lg| \le ||L||\ ||g|| \le ||L||\ ||f||$ for $|g| \le f$.

Lemma 18.25 *If $c \ge 0,\ \lambda(cf) = c\lambda(f),\ f_1 \le f_2$ implies $\lambda f_1 \le \lambda f_2$, and*

$$\lambda(f_1 + f_2) = \lambda(f_1) + \lambda(f_2).$$

Proof: The first two assertions are easy to see so we consider the third. Let $|g_j| \le f_j$ and let $\widetilde{g}_j = e^{i\theta_j} g_j$ where θ_j is chosen such that $e^{i\theta_j} Lg_j = |Lg_j|$. Thus $L\widetilde{g}_j = |Lg_j|$. Then

$$|\widetilde{g}_1 + \widetilde{g}_2| \le f_1 + f_2.$$

Hence

$$|Lg_1| + |Lg_2| = L\widetilde{g}_1 + L\widetilde{g}_2 =$$

$$L(\widetilde{g}_1 + \widetilde{g}_2) = |L(\widetilde{g}_1 + \widetilde{g}_2)| \le \lambda(f_1 + f_2). \tag{18.17}$$

Choose g_1 and g_2 such that $|Lg_i| + \varepsilon > \lambda(f_i)$. Then (18.17) shows

$$\lambda(f_1) + \lambda(f_2) - 2\varepsilon \le \lambda(f_1 + f_2).$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$\lambda(f_1) + \lambda(f_2) \leq \lambda(f_1 + f_2). \tag{18.18}$$

Now let $|g| \leq f_1 + f_2$, $|Lg| \geq \lambda(f_1 + f_2) - \varepsilon$. Let

$$h_i(x) = \begin{cases} \frac{f_i(x)g(x)}{f_1(x)+f_2(x)} \text{ if } f_1(x) + f_2(x) > 0, \\ 0 \text{ if } f_1(x) + f_2(x) = 0. \end{cases}$$

Then h_i is continuous and $h_1(x) + h_2(x) = g(x)$, $|h_i| \leq f_i$. Therefore,

$$\begin{aligned} -\varepsilon + \lambda(f_1 + f_2) &\leq |Lg| \leq |Lh_1 + Lh_2| \leq |Lh_1| + |Lh_2| \\ &\leq \lambda(f_1) + \lambda(f_2). \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this shows with (18.18) that

$$\lambda(f_1 + f_2) \leq \lambda(f_1) + \lambda(f_2) \leq \lambda(f_1 + f_2)$$

which proves the lemma.

Let $C(X;\mathbb{R})$ be the real-valued functions in $C(X)$ and define

$$\Lambda_R(f) = \lambda f^+ - \lambda f^-$$

for $f \in C(X;\mathbb{R})$. Using Lemma 18.25 and the identity

$$(f_1 + f_2)^+ + f_1^- + f_2^- = f_1^+ + f_2^+ + (f_1 + f_2)^-$$

to write

$$\lambda(f_1 + f_2)^+ - \lambda(f_1 + f_2)^- = \lambda f_1^+ - \lambda f_1^- + \lambda f_2^+ - \lambda f_2^-,$$

we see that $\Lambda_R(f_1+f_2) = \Lambda_R(f_1)+\Lambda_R(f_2)$. To show that Λ_R is linear, we need to verify that $\Lambda_R(cf) = c\Lambda_R(f)$ for all $c \in \mathbb{R}$. But

$$(cf)^{\pm} = cf^{\pm},$$

if $c \geq 0$ while

$$(cf)^+ = -c(f)^-,$$

if $c < 0$ and

$$(cf)^- = (-c)f^+,$$

if $c < 0$. Thus, if $c < 0$,

$$\Lambda_R(cf) = \lambda(cf)^+ - \lambda(cf)^- = \lambda\left((-c)\, f^-\right) - \lambda\left((-c)f^+\right)$$

$$= -c\lambda(f^-) + c\lambda(f^+) = c(\lambda(f^+) - \lambda(f^-)).$$

(If this looks familiar it may be because we used this approach earlier in defining the integral of a real-valued function.) Now let

$$\Lambda f = \Lambda_R(\operatorname{Re} f) + i\Lambda_R(\operatorname{Im} f)$$

for arbitrary $f \in C(X)$. It is easy to see that Λ is a positive linear functional on $C(X)$ $(= C_c(X)$ since X is compact). By the Riesz representation theorem for positive linear functionals, there exists a unique Radon measure μ such that

$$\Lambda f = \int_X f\, d\mu$$

for all $f \in C(X)$. Thus $\Lambda(1) = \mu(X)$. Now we present the Riesz representation theorem for $C(X)'$.

Theorem 18.26 *Let $L \in (C(X))'$. Then there exists a Radon measure μ and a function $\sigma \in L^{\infty}(X,\mu)$ such that*

$$L(f) = \int_X f\ \sigma\ d\mu.$$

Proof: Let $f \in C(X)$. Then there exists a unique Radon measure μ such that

$$|Lf| \leq \Lambda(|f|) = \int_X |f| d\mu = ||f||_1.$$

Since μ is a Radon measure, we know $C(X)$ is dense in $L^1(X,\mu)$. Therefore L extends uniquely to an element of $(L^1(X,\mu))'$. By the Riesz representation theorem for L^1, there exists a unique $\sigma \in L^{\infty}(X,\mu)$ such that

$$Lf = \int_X f\ \sigma\ d\mu$$

for all $f \in C(X)$.

It is possible to give a simple generalization of the above theorem to locally compact Hausdorff spaces. We will do this in the special case where $X = \mathbb{R}^n$. Define

$$\widetilde{X} \equiv \prod_{i=1}^{n} [-\infty, \infty]$$

With the product topology where a subbasis for a topology on $[-\infty,\infty]$ will consist of sets of the form $[-\infty, a), (a,b)$ or $(a,\infty]$. We can also make $\widetilde{X}$ into a metric space by using the metric,

$$\rho(x,y) \equiv \sum_{i=1}^{n} |\arctan x_i - \arctan y_i|$$

We also define by $C_0(X)$ the space of continuous functions, f, defined on X such that

$$\lim_{dist\left(x,\widetilde{X}\setminus X\right)\to 0} f(x) = 0.$$

For this space of functions, $||f||_0 \equiv \sup\{|f(x)| : x \in X\}$ is a norm which makes this into a Banach space. Then the generalization is the following corollary.

Corollary 18.27 *Let $L \in (C_0(X))'$ where $X = \mathbb{R}^n$. Then there exists $\sigma \in L^{\infty}(X,\mu)$ for μ a Radon measure such that for all $f \in C_0(X)$,*

$$L(f) = \int_X f\sigma d\mu.$$

Proof: Let

$$\widetilde{D} \equiv \left\{ f \in C\left(\widetilde{X}\right) : f(z) = 0 \text{ if } z \in \widetilde{X} \setminus X \right\}.$$

Thus $\widetilde{D}$ is a closed subspace of the Banach space $C\left(\widetilde{X}\right)$. Let $\theta : C_0(X) \to \widetilde{D}$ be defined by

$$\theta f(x) = \begin{cases} f(x) \text{ if } x \in X, \\ 0 \text{ if } x \in \widetilde{X} \setminus X \end{cases}$$

Then θ is an isometry of $C_0(X)$ and $\widetilde{D}$. $(||\theta u|| = ||u||\,.)$It follows we have the following diagram.

$$\begin{array}{ccccc} C_0(X)' & \overset{\theta^*}{\leftarrow} & \left(\widetilde{D}\right)' & \overset{i^*}{\leftarrow} & C\left(\widetilde{X}\right)' \\ & & & & \\ C_0(X) & \underset{\theta}{\rightarrow} & \widetilde{D} & \underset{i}{\rightarrow} & C\left(\widetilde{X}\right) \end{array}$$

By the Hahn Banach theorem, there exists $L_1 \in C\left(\widetilde{X}\right)'$ such that $\theta^* i^* L_1 = L$. Now we apply Theorem 18.26 to get the existence of a Radon measure, μ_1, on $\widetilde{X}$ and a function $\sigma \in L^\infty\left(\widetilde{X}, \mu_1\right)$, such that

$$L_1 g = \int_{\widetilde{X}} g\sigma d\mu_1.$$

Letting the σ algebra of μ_1 measurable sets be denoted by $\mathcal{S}_1$, we define

$$\mathcal{S} \equiv \left\{E \setminus \left\{\widetilde{X} \setminus X\right\} : E \in \mathcal{S}_1\right\}$$

and let μ be the restriction of μ_1 to $\mathcal{S}$. If $f \in C_0(X)$,

$$Lf = \theta^* i^* L_1 f \equiv L_1 i\theta f = L_1 \theta f = \int_{\widetilde{X}} \theta f \sigma d\mu_1 = \int_X f\sigma d\mu.$$

This proves the corollary.

18.6 Weak $*$ convergence

A very important sort of convergence in applications of functional analysis is the concept of weak or weak $*$ convergence. It is important because it allows us to extract a convergent subsequence of a given bounded sequence. The only problem is the convergence is very weak so it does not tell us as much as we would like. Nevertheless, it is a very useful concept. The big theorems in the subject are the Eberlein Smulian theorem and the Banach Alaoglu theorem about the weak or weak $*$ compactness of the closed unit balls in either a Banach space or its dual space. These theorems are proved in Yosida [29]. Here we will consider a special case which turns out to be by far the most important in applications and it is not hard to get from the results of this chapter. First we define what we mean by weak and weak $*$ convergence.

Definition 18.28 *Let X' be the dual of a Banach space X and let $\{x_n^*\}$ be a sequence of elements of X'. Then we say x_n^* converges weak $*$ to x^* if and only if for all $x \in X$,*

$$\lim_{n\to\infty} x_n^*(x) = x^*(x)\,.$$

We say a sequence in $X, \{x_n\}$ converges weakly to $x \in X$ if and only if for all $x^* \in X'$

$$\lim_{n\to\infty} x^*(x_n) = x^*(x)\,.$$

The main result is contained in the following lemma.

Lemma 18.29 *Let X' be the dual of a Banach space, X and suppose X is separable. Then if $\{x_n^*\}$ is a bounded sequence in X', there exists a weak $*$ convergent subsequence.*

Proof: Let D be a dense countable set in X. Then the sequence, $\{x_n^*(x)\}$ is bounded for all x and in particular for all $x \in D$. Use the Cantor diagonal process to obtain a subsequence, still denoted by n such that $x_n^*(d)$ converges for each $d \in D$. Now let $x \in X$ be completely arbitrary. We will show $\{x_n^*(x)\}$ is a Cauchy sequence. Let $\varepsilon > 0$ be given and pick $d \in D$ such that for all n

$$|x_n^*(x) - x_n^*(d)| < \frac{\varepsilon}{3}.$$

We can do this because D is dense. By the first part of the proof, there exists N_ε such that for all $m, n > N_\varepsilon$,

$$|x_n^*(d) - x_m^*(d)| < \frac{\varepsilon}{3}.$$

Then for such m, n,

$$\begin{aligned} |x_n^*(x) - x_m^*(x)| &\leq |x_n^*(x) - x_n^*(d)| + |x_n^*(d) - x_m^*(d)| \\ + |x_m^*(d) - x_m^*(x)| &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Since ε is arbitrary, this shows $\{x_n^*(x)\}$ is a Cauchy sequence for all $x \in X$.

Now we define $f(x) \equiv \lim_{n\to\infty} x_n^*(x)$. Since each x_n^* is linear, it follows f is also linear. In addition to this,

$$|f(x)| = \lim_{n\to\infty} |x_n^*(x)| \leq K\, ||x||$$

where K is some constant which is larger than all the norms of the x_n^*. We know such a constant exists because we assumed that the sequence, $\{x_n^*\}$ was bounded. This proves the lemma.

The lemma implies the following important theorem.

Theorem 18.30 *Let Ω be a measurable subset of $\mathbb{R}^n$ and let $\{f_k\}$ be a bounded sequence in $L^p(\Omega)$ where $1 < p \leq \infty$. Then there exists a weak $*$ convergent subsequence.*

Proof: We know from Corollary 12.15 that $L^{p'}(\Omega)$ is separable. From the Riesz representation theorem, we obtain the desired result.

Note that from the Riesz representation theorem, it follows that if $p < \infty$, then we also have the susequence converges weakly.

18.7 Exercises

1. Suppose $\lambda(E) = \int_E f d\mu$ where λ and μ are two measures and $f \in L^1(\mu)$. Show $\lambda << \mu$.

2. Show that $\lambda << \mu$ for μ and λ two finite measures, if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $\mu(E) < \delta$, it follows $\lambda(E) < \varepsilon$.

3. Suppose λ, μ are two finite measures defined on a σ algebra $\mathcal{S}$. Show $\lambda = \lambda_S + \lambda_A$ where λ_S and $\lambda_A$are finite measures satisfying

$$\lambda_S(E) = \lambda_S(E \cap S)\ ,\ \mu(S) = 0 \text{ for some } S \subseteq \mathcal{S},$$

$$\lambda_A << \mu.$$

This is called the Lebesgue decomposition. **Hint**: This is just a generalization of the Radon Nikodym theorem. In the proof of this theorem, let

$$S = \{x : h(x) = 1\},\ \lambda_S(E) = \lambda(E \cap S),$$

$$\lambda_A(E) = \lambda(E \cap S^c).$$

We write $\mu \perp \lambda_S$ and $\lambda_A << \mu$ in this situation.

4. ↑ Generalize the result of Problem 3 to the case where μ is σ finite and λ is finite.

5. Let F be a nondecreasing right continuous, bounded function,

$$\lim_{x\to-\infty} F(x) = 0.$$

Let $Lf = \int f dF$ where $f \in C_c(\mathbb{R})$ and this is just the Riemann Stieltjes integral. Let λ be the Radon measure representing L. Show

$$\lambda((a,b]) = F(b) - F(a)\ ,\ \lambda((a,b)) = F(b-) - F(a).$$

6. ↑ Using Problems 3, 4, and 5, show there is a bounded nondecreasing function $G(x)$ such that $G(x) \leq F(x)$ and $G(x) = \int_{-\infty}^{x} \ell(t)dt$ for some $\ell \geq 0$, $\ell \in L^1(m)$. Also, if $F(x) - G(x) = S(x)$, then $S(x)$ is non decreasing and if λ_S is the measure representing $\int f dS$, then $\lambda_S \perp m$. **Hint:** Consider $G(x) = \lambda_A((-\infty, x])$.

7. Let λ and μ be two measures defined on $\mathcal{S}$, a σ algebra of subsets of Ω. Suppose μ is σ finite and $g \geq 0$ with g measurable. Show that

$$g = \frac{d\lambda}{d\mu},\ (\lambda(E) = \int_E g d\mu)$$

if and only if for all $A \in \mathcal{S}$,and $\alpha, \beta, \geq 0$,

$$\lambda(A \cap \{x : g(x) \geq \alpha\}) \geq \alpha\mu(A \cap \{x : g(x) \geq \alpha\}),$$

$$\lambda(A \cap \{x : g(x) < \beta\}) \leq \beta\mu(A \cap \{x : g(x) < \beta\}).$$

Hint: To show $g = \frac{d\lambda}{d\mu}$ from the two conditions, use the conditions to argue that for $\mu(A) < \infty$,

$$\beta\mu(A \cap \{x : g(x) \in [\alpha,\beta)\}) \geq \lambda(A \cap \{x : g(x) \in [\alpha,\beta)\})$$

$$\geq \alpha\mu(A \cap \{x : g(x) \in [\alpha,\beta)\}).$$

8. Let $r, p, q \in (1,\infty)$ satisfy

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$$

and let $f \in L^p(\mathbb{R}^n)\,,\ g \in L^q(\mathbb{R}^n)\,,\ f, g \geq 0$. Young's inequality says that

$$||f * g||_r \leq ||g||_q\, ||f||_p\,.$$

Prove Young's inequality by justifying or modifying the steps in the following argument using Problem 27 of Chapter 12. Let

$$h \in L^{r'}(\mathbb{R}^n)\,.\left(\frac{1}{r} + \frac{1}{r'} = 1\right)$$

$$\int f * g(\mathbf{x})\, |h(\mathbf{x})|\, dx = \int \int f(\mathbf{y})\, g(\mathbf{x} - \mathbf{y})\, |h(\mathbf{x})|\, dxdy.$$

Let $r\theta = p$ so $\theta \in (0,1)$, $p'(1-\theta) = q,\ p' \geq r'$. Then the above

$$\leq \int \int |f(\mathbf{y})| \left|g(\mathbf{x}-\mathbf{y})\right|^{\theta} \left|g(\mathbf{x}-\mathbf{y})\right|^{1-\theta} |h(\mathbf{x})|\, dydx$$

$$\leq \int \left(\int \left(\left|g(\mathbf{x}-\mathbf{y})\right|^{1-\theta} |h(\mathbf{x})| \right)^{r'} dx \right)^{1/r'} \cdot$$

$$\left(\int \left(|f(\mathbf{y})| \left|g(\mathbf{x}-\mathbf{y})\right|^{\theta} \right)^{r} dx \right)^{1/r} dy$$

$$\leq \left[\int \left(\int \left(\left|g(\mathbf{x}-\mathbf{y})\right|^{1-\theta} |h(\mathbf{x})| \right)^{r'} dx \right)^{p'/r'} dy \right]^{1/p'} \cdot$$

$$\left[\int \left(\int \left(|f(\mathbf{y})| \left|g(\mathbf{x}-\mathbf{y})\right|^{\theta} \right)^{r} dx \right)^{p/r} dy \right]^{1/p}$$

$$\leq \left[\int \left(\int \left(\left|g(\mathbf{x}-\mathbf{y})\right|^{1-\theta} |h(\mathbf{x})| \right)^{p'} dy \right)^{r'/p'} dx \right]^{1/r'} \cdot$$

$$\left[\int |f(\mathbf{y})|^{p} \left(\int \left|g(\mathbf{x}-\mathbf{y})\right|^{\theta r} dx \right)^{p/r} dy \right]^{1/p}$$

$$= \left[\int |h(\mathbf{x})|^{r'} \left(\int \left|g(\mathbf{x}-\mathbf{y})\right|^{(1-\theta)p'} dy \right)^{r'/p'} dx \right]^{1/r'} ||g||_q^{q/r} \, ||f||_p$$

$$= ||g||_q^{q/r} \, ||g||_q^{q/p'} \, ||f||_p \, ||h||_{r'} = ||g||_q \, ||f||_p \, ||h||_{r'} \, .$$

Therefore $||f * g||_r \leq ||g||_q \, ||f||_p$. Does this inequality continue to hold if r, p, q are only assumed to be in $[1, \infty]$? Explain.

9. Let $X = [0, \infty]$ with a subbasis for the topology given by sets of the form $[0, b)$ and $(a, \infty]$. Show that X is a compact set. Consider all functions of the form

$$\sum_{k=0}^{n} a_k e^{-xkt}$$

where $t > 0$. Show that this collection of functions is an algebra of functions in $C(X)$ if we define

$$e^{-t\infty} \equiv 0,$$

and that it separates the points and annihilates no point. Conclude that this algebra is dense in $C(X)$.

10. $\uparrow$ Suppose $f \in C(X)$ and for all t large enough, say $t \geq \delta$,

$$\int_0^\infty e^{-tx} f(x)\, dx = 0.$$

Show that $f(x) = 0$ for all $x \in (0, \infty)$. **Hint:** Show the measure given by $d\mu = e^{-\delta x} dx$ is a regular measure and so $C(X)$ is dense in $L^2(X, \mu)$. Now use Problem 9 to conclude that $f(x) = 0$ a.e.

11. $\uparrow$ Suppose f is continuous on $(0, \infty)$, and for some $\gamma > 0$,

$$\left| f(x)\, e^{-\gamma x} \right| \leq C$$

for all $x \in (0, \infty)$, and suppose also that for all t large enough,

$$\int_0^\infty f(x)\, e^{-tx} dx = 0.$$

Show $f(x) = 0$ for all $x \in (0, \infty)$. A common procedure in elementary differential equations classes is to obtain the Laplace transform of an unknown function and then, using a table of Laplace transforms, find what the unknown function is. Can the result of this problem be used to justify this procedure?

12. $\uparrow$ The following theorem is called the mean ergodic theorem.

Theorem 18.31 *Let $(\Omega, \mathcal{S}, \mu)$ be a finite measure space and let $\psi : \Omega \to \Omega$ satisfy $\psi^{-1}(E) \in \mathcal{S}$ for all $E \in \mathcal{S}$. Also suppose for all positive integers, n, that*

$$\mu\left(\psi^{-n}(E)\right) \leq K\mu(E).$$

For $f \in L^p(\Omega)$, and $p \geq 1$, let

$$Tf \equiv f \circ \psi. \tag{18.19}$$

Then $T \in \mathcal{L}(L^p(\Omega), L^p(\Omega))$ with

$$||T^n|| \leq K^{1/p}. \tag{18.20}$$

Defining $A_n \in \mathcal{L}(L^p(\Omega), L^p(\Omega))$ by

$$A_n \equiv \frac{1}{n} \sum_{k=0}^{n-1} T^k,$$

there exists $A \in \mathcal{L}(L^p(\Omega), L^p(\Omega))$ such that for all $f \in L^p(\Omega)$,

$$A_n f \to Af \text{ weakly} \tag{18.21}$$

and A is a projection, $A^2 = A$, onto the space of all $f \in L^p(\Omega)$ such that $Tf = f$. The norm of A satisfies

$$||A|| \leq K^{1/p}. \tag{18.22}$$

To prove this theorem first show (18.20) by verifying it for simple functions and then using the density of simple functions in $L^p(\Omega)$. Next let

$$M \equiv \left\{ g \in L^p(\Omega) : ||A_n g||_p \to 0. \right\}$$

Show M is a closed subspace of $L^p(\Omega)$ containing $(I-T)(L^p(\Omega))$. Next show that if

$$A_{n_k}f \to h \text{ weakly,}$$

then $Th = h$. If $\xi \in L^{p'}(\Omega)$ is such that $\int \xi g d\mu = 0$ for all $g \in M$, then using the definition of A_n, it follows that for all $k \in L^p(\Omega)$,

$$\int \xi k d\mu = \int (\xi T^n k + \xi (I - T^n) k) d\mu = \int \xi T^n k d\mu$$

and so

$$\int \xi k d\mu = \int \xi A_n k d\mu. \tag{18.23}$$

Using (18.23) show that if $A_{n_k}f \to g$ weakly and $A_{m_k}f \to h$ weakly, then

$$\int \xi g d\mu = \lim_{k\to\infty} \int \xi A_{n_k} f d\mu = \int \xi f d\mu = \lim_{k\to\infty} \int \xi A_{m_k} f d\mu = \int \xi h d\mu. \tag{18.24}$$

Now argue that

$$T^n g = \text{ weak } \lim_{k\to\infty} T^n A_{n_k} f = \text{ weak } \lim_{k\to\infty} A_{n_k} T^n f = g.$$

Conclude that $A_n(g-h) = g-h$ so if $g \neq h$, then $g-h \notin M$ because

$$A_n(g-h) \to g-h \neq 0.$$

Now show there exists $\xi \in L^{p'}(\Omega)$ such that $\int \xi (g-h) d\mu \neq 0$ but $\int \xi k d\mu = 0$ for all $k \in M$, contradicting (18.24). Use this to conclude that if $p > 1$ then $A_n f$ converges weakly for each $f \in L^p(\Omega)$. Let Af denote this weak limit and verify that the conclusions of the theorem hold in the case where $p > 1$. To get the case where $p = 1$ use the theorem just proved for $p > 1$ on $L^2(\Omega) \cap L^1(\Omega)$ a dense subset of $L^1(\Omega)$ along with the observation that $L^\infty(\Omega) \subseteq L^2(\Omega)$ due to the assumption that we are in a finite measure space.

13. Generalize Corollary 18.27 to $X = \mathbb{C}^n$.

14. Show that in a reflexive Banach space, weak and weak $*$ convergence are the same.

Weak Derivatives

19.1 Test functions and weak derivatives

In elementary courses in mathematics, functions are often thought of as things which have a formula associated with them and it is the formula which receives the most attention. For example, in beginning calculus courses the derivative of a function is defined as the limit of a difference quotient. We start with one function which we tend to identify with a formula and, by taking a limit, we get another formula for the derivative. A jump in abstraction occurs as soon as we encounter the derivative of a function of n variables where the derivative is defined as a certain linear transformation which is determined not by a formula but by what it does to vectors. When this is understood, we see that it reduces to the usual idea in one dimension. The idea of weak partial derivatives goes further in the direction of defining something in terms of what it does rather than by a formula, and extra generality is obtained when it is used. In particular, it is possible to differentiate almost anything if we use a weak enough notion of what we mean by the derivative. This has the advantage of letting us talk about a weak partial derivative of a function without having to agonize over the important question of existence but it has the disadvantage of not allowing us to say very much about this weak partial derivative. Nevertheless, it is the idea of weak partial derivatives which makes it possible to use functional analytic techniques in the study of partial differential equations and we will show in this chapter that the concept of weak derivative is useful for unifying the discussion of some very important theorems. We will also show that certain things we wish were true, such as the equality of mixed partial derivatives, are true within the context of weak derivatives.

Let $\Omega \subseteq \mathbb{R}^n$. A distribution on Ω is defined to be a linear functional on $C_c^{\infty}(\Omega)$, called the space of test functions. The space of all such linear functionals will be denoted by $\mathcal{D}^*(\Omega)$. Actually, more is sometimes done here. One imposes a topology on $C_c^{\infty}(\Omega)$ making it into something called a topological vector space, and when this has been done, $\mathcal{D}'(\Omega)$ is defined as the dual space of this topological vector space. To see this, consult the book by Yosida [29] or the book by Rudin [25].

Example: The space $L_{loc}^1(\Omega)$ may be considered as a subset of $\mathcal{D}^*(\Omega)$ as follows.

$$f(\phi) \equiv \int_{\Omega} f(\mathbf{x})\, \phi(\mathbf{x})\, dx$$

for all $\phi \in C_c^{\infty}(\Omega)$. Recall that $f \in L_{loc}^1(\Omega)$ if $f\mathcal{X}_K \in L^1(\Omega)$ whenever K is compact.

Example: $\delta_x \in \mathcal{D}^*(\Omega)$ where $\delta_{\mathbf{x}}(\phi) \equiv \phi(\mathbf{x})$.

It will be observed from the above two examples and a little thought that $\mathcal{D}^*(\Omega)$ is truly enormous. We shall define the derivative of a distribution in such a way that it agrees with the usual notion of a derivative on those distributions which are also continuously differentiable functions. With this in mind, let f be the restriction to Ω of a smooth function defined on $\mathbb{R}^n$. Then $D_{x_i} f$ makes sense and for $\phi \in C_c^{\infty}(\Omega)$

$$D_{x_i} f(\phi) \equiv \int_{\Omega} D_{x_i} f(\mathbf{x})\, \phi(\mathbf{x})\, dx = -\int_{\Omega} f D_{x_i} \phi dx = -f(D_{x_i} \phi).$$

Motivated by this we make the following definition.

Definition 19.1 *For $T \in \mathcal{D}^* (\Omega)$*

$$D_{x_i} T (\phi) \equiv -T (D_{x_i} \phi).$$

Of course one can continue taking derivatives indefinitely. Thus,

$$D_{x_i x_j} T \equiv D_{x_i} \left(D_{x_j} T \right)$$

and it is clear that all mixed partial derivatives are equal because this holds for the functions in $C_c^\infty (\Omega)$. Thus we can differentiate virtually anything, even functions that may be discontinuous everywhere. However the notion of "derivative" is very weak, hence the name, "weak derivatives".

Example: Let $\Omega = \mathbb{R}$ and let

$$H (x) \equiv \begin{cases} 1 \text{ if } x \geq 0, \\ 0 \text{ if } x < 0. \end{cases}$$

Then

$$DH (\phi) = -\int H (x) \phi' (x) \, dx = \phi (0) = \delta_0(\phi).$$

Note that in this example, DH is not a function.

What happens when Df is a function?

Theorem 19.2 *Let $\Omega = (a, b)$ and suppose that f and Df are both in $L^1 (a, b)$. Then f is equal to a continuous function a.e., still denoted by f and*

$$f (x) = f (a) + \int_a^x Df (t) \, dt.$$

In proving Theorem 19.2 we shall use the following lemma.

Lemma 19.3 *Let $T \in \mathcal{D}^* (a, b)$ and suppose $DT = 0$. Then there exists a constant C such that*

$$T (\phi) = \int_a^b C \phi dx.$$

Proof: $T (D\phi) = 0$ for all $\phi \in C_c^\infty (a, b)$ from the definition of $DT = 0$. Let

$$\phi_0 \in C_c^\infty (a, b), \int_a^b \phi_0 (x) \, dx = 1,$$

and let

$$\psi_\phi (x) = \int_a^x [\phi (t) - \left(\int_a^b \phi (y) \, dy \right) \phi_0 (t)] dt$$

for $\phi \in C_c^\infty (a, b)$. Thus $\psi_\phi \in C_c^\infty (a, b)$ and

$$D\psi_\phi = \phi - \left(\int_a^b \phi (y) \, dy \right) \phi_0.$$

Therefore,

$$\phi = D\psi_\phi + \left(\int_a^b \phi (y) \, dy \right) \phi_0$$

and so

$$T(\phi) = T(D\psi_\phi) + \left(\int_a^b \phi(y)\,dy\right) T(\phi_0) = \int_a^b T(\phi_0)\,\phi(y)\,dy.$$

Let $C = T\phi_0$. This proves the lemma.

Proof of Theorem 19.2 Since f and Df are both in $L^1(a,b)$,

$$Df(\phi) - \int_a^b Df(x)\,\phi(x)\,dx = 0.$$

Consider

$$f(\cdot) - \int_a^{(\cdot)} Df(t)\,dt$$

and let $\phi \in C_c^\infty(a,b)$.

$$D\left(f(\cdot) - \int_a^{(\cdot)} Df(t)\,dt\right)(\phi)$$

$$\equiv -\int_a^b f(x)\,\phi'(x)\,dx + \int_a^b \left(\int_a^x Df(t)\,dt\right)\phi'(x)\,dx$$

$$\begin{aligned} &= Df(\phi) + \int_a^b \int_t^b Df(t)\,\phi'(x)\,dxdt \\ &= Df(\phi) - \int_a^b Df(t)\,\phi(t)\,dt = 0. \end{aligned}$$

By Lemma 19.3, there exists a constant, C, such that

$$\left(f(\cdot) - \int_a^{(\cdot)} Df(t)\,dt\right)(\phi) = \int_a^b C\phi(x)\,dx$$

for all $\phi \in C_c^\infty(a,b)$. Thus

$$\int_a^b \{\left(f(x) - \int_a^x Df(t)\,dt\right) - C\}\phi(x)\,dx = 0$$

for all $\phi \in C_c^\infty(a,b)$. It follows from Lemma 19.6 in the next section that

$$f(x) - \int_a^x Df(t)\,dt - C = 0 \text{ a.e. } x.$$

Thus we let $f(a) = C$ and write

$$f(x) = f(a) + \int_a^x Df(t)\,dt.$$

This proves Theorem 19.2.

Theorem 19.2 says that

$$f(x) = f(a) + \int_a^x Df(t)\,dt$$

whenever it makes sense to write $\int_a^x Df(t)\,dt$, if Df is interpreted as a weak derivative. Somehow, this is the way it ought to be. It follows from the fundamental theorem of calculus in Chapter 20 that $f'(x)$ exists for a.e. x where the derivative is taken in the sense of a limit of difference quotients and $f'(x) = Df(x)$. This raises an interesting question. Suppose f is continuous on $[a,b]$ and $f'(x)$ exists in the classical sense for a.e. x. Does it follow that

$$f(x) = f(a) + \int_a^x f'(t)\,dt?$$

The answer is no. To see an example, consider Problem 3 of Chapter 7 which gives an example of a function which is continuous on $[0,1]$, has a zero derivative for a.e. x but climbs from 0 to 1 on $[0,1]$. Thus this function is not recovered from integrating its classical derivative.

In summary, if the notion of weak derivative is used, one can at least give meaning to the derivative of almost anything, the mixed partial derivatives are always equal, and, in one dimension, one can recover the function from integrating its derivative. None of these claims are true for the classical derivative. Thus weak derivatives are convenient and rule out pathologies.

19.2 Weak derivatives in L^p_{loc}

Definition 19.4 *We say $f \in L^p_{loc}(\mathbb{R}^n)$ if $f\mathcal{X}_K \in L^p$ whenever K is compact.*

Definition 19.5 *For $\alpha = (k_1, \cdots, k_n)$ where the k_i are nonnegative integers, we define*

$$|\alpha| \equiv \sum_{i=1}^n |k_{x_i}|,\ D^\alpha f(\mathbf{x}) \equiv \frac{\partial^{|\alpha|} f(\mathbf{x})}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}}.$$

We want to consider the case where u and $D^\alpha u$ for $|\alpha| = 1$ are each in $L^p_{loc}(\mathbb{R}^n)$. The next lemma is the one alluded to in the proof of Theorem 19.2.

Lemma 19.6 *Suppose $f \in L^1_{loc}(\mathbb{R}^n)$ and suppose*

$$\int f\phi dx = 0$$

for all $\phi \in C_c^\infty(\mathbb{R}^n)$. Then $f(\mathbf{x}) = 0$ a.e. $\mathbf{x}$.

Proof: Without loss of generality f is real-valued. Let

$$E \equiv \{\mathbf{x} : f(\mathbf{x}) > \epsilon\}$$

and let

$$E_m \equiv E \cap B(0,m).$$

We show that $m(E_m) = 0$. If not, there exists an open set, V, and a compact set K satisfying

$$K \subseteq E_m \subseteq V \subseteq B(0,m),\ m(V \setminus K) < 4^{-1} m(E_m),$$

$$\int_{V\setminus K} |f|\, dx < \epsilon 4^{-1} m(E_m).$$

Let H and W be open sets satisfying

$$K \subseteq H \subseteq \overline{H} \subseteq W \subseteq \overline{W} \subseteq V$$

and let

$$\overline{H} \prec g \prec W$$

where the symbol, $\prec$, has the same meaning as it does in Chapter 6. Then let ϕ_δ be a mollifier and let $h \equiv g * \phi_\delta$ for δ small enough that

$$K \prec h \prec V.$$

Thus

$$\begin{aligned} 0 &= \int f h dx = \int_K f dx + \int_{V\setminus K} f h dx \\ &\geq \epsilon m(K) - \epsilon 4^{-1} m(E_m) \\ &\geq \epsilon \left(m(E_m) - 4^{-1} m(E_m) \right) - \epsilon 4^{-1} m(E_m) \\ &\geq 2^{-1} \epsilon m(E_m). \end{aligned}$$

Therefore, $m(E_m) = 0$, a contradiction. Thus

$$m(E) \leq \sum_{m=1}^{\infty} m(E_m) = 0$$

and so, since $\epsilon > 0$ is arbitrary,

$$m(\{\mathbf{x} : f(\mathbf{x}) > 0\}) = 0.$$

Similarly $m(\{\mathbf{x} : f(\mathbf{x}) < 0\}) = 0$. This proves the lemma.

This lemma allows the following definition.

Definition 19.7 *We say for $u \in L^1_{loc}(\mathbb{R}^n)$ that $D^\alpha u \in L^1_{loc}(\mathbb{R}^n)$ if there exists a function $g \in L^1_{loc}(\mathbb{R}^n)$, necessarily unique by Lemma 19.6, such that for all $\phi \in C_c^\infty(\mathbb{R}^n)$,*

$$\int g\phi dx = D^\alpha u(\phi) = \int (-1)^{|\alpha|} u(D^\alpha \phi)\, dx.$$

We call g $D^\alpha u$ when this occurs.

Lemma 19.8 *Let $u \in L^1_{loc}$ and suppose $u_{,i} \in L^1_{loc}$, where the subscript on the u following the comma denotes the ith weak partial derivative. Then if ϕ_ϵ is a mollifier and $u_\epsilon \equiv u * \phi_\epsilon$, we can conclude that $u_{\epsilon,i} \equiv u_{,i} * \phi_\epsilon$.*

Proof: If $\psi \in C_c^\infty(\mathbb{R}^n)$, then

$$\begin{aligned} \int u(\mathbf{x} - \mathbf{y})\, \psi_{,i}(\mathbf{x})\, dx &= \int u(\mathbf{z})\, \psi_{,i}(\mathbf{z} + \mathbf{y})\, dz \\ &= -\int u_{,i}(\mathbf{z})\, \psi(\mathbf{z} + \mathbf{y})\, dz \\ &= -\int u_{,i}(\mathbf{x} - \mathbf{y})\, \psi(\mathbf{x})\, dx. \end{aligned}$$

Therefore,

$$\begin{aligned} u_{\epsilon,i}(\psi) &= -\int u_\epsilon \psi_{,i} = -\int\int u(\mathbf{x}-\mathbf{y})\,\phi_\epsilon(\mathbf{y})\,\psi_{,i}(\mathbf{x})\,d\,ydx \\ &= -\int\int u(\mathbf{x}-\mathbf{y})\,\psi_{,i}(\mathbf{x})\,\phi_\epsilon(\mathbf{y})\,dxdy \\ &= \int\int u_{,i}(\mathbf{x}-\mathbf{y})\,\psi(\mathbf{x})\,\phi_\epsilon(\mathbf{y})\,dxdy \\ &= \int u_{,i} * \phi_\epsilon(\mathbf{x})\,\psi(\mathbf{x})\,dx. \end{aligned}$$

The technical questions about product measurability in the use of Fubini's theorem may be resolved by picking a Borel measurable representative for u. This proves the lemma.

Next we discuss a form of the product rule.

Lemma 19.9 *Let $\psi \in C^\infty(\mathbb{R}^n)$ and suppose $u, u_{,i} \in L^p_{loc}(\mathbb{R}^n)$. Then $(u\psi)_{,i}$ and $u\psi$ are in L^p_{loc} and*

$$(u\psi)_{,i} = u_{,i}\psi + u\psi_{,i}.$$

Proof: Let $\psi \in C_c^\infty(\mathbb{R}^n)$ then

$$\begin{aligned} (u\psi)_{,i}(\phi) &= -\int u\psi\phi_{,i}dx \\ &= -\int u[(\psi\phi)_{,i} - \phi\psi_{,i}]dx \\ &= \int \left(u_{,i}\psi\phi + u\psi_{,i}\phi\right)dx. \end{aligned}$$

This proves the lemma. We recall the notation for the gradient of a function.

$$\nabla u(\mathbf{x}) \equiv (u_{,1}(\mathbf{x}) \cdots u_{,n}(\mathbf{x}))^T$$

thus

$$Du(\mathbf{x})\,\mathbf{v} = \nabla u(\mathbf{x}) \cdot \mathbf{v}.$$

19.3 Morrey's inequality

The following inequality will be called Morrey's inequality. It relates an expression which is given pointwise to an integral of the *pth* power of the derivative.

Lemma 19.10 *Let $u \in C^1(\mathbb{R}^n)$ and $p > n$. Then there exists a constant, C, depending only on n such that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,*

$$|u(\mathbf{x}) - u(\mathbf{y})|$$

$$\leq C\left(\int_{B(\mathbf{x},2|\mathbf{x}-\mathbf{y}|)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} \left(|\mathbf{x}-\mathbf{y}|^{(1-n/p)}\right).$$

Proof: In the argument C will be a generic constant which depends on n.

$$\int_{B(\mathbf{x},r)} |u(\mathbf{x}) - u(\mathbf{y})| \, dy = \int_0^r \int_{S^{n-1}} |u(\mathbf{x} + \rho\omega) - u(\mathbf{x})| \, \rho^{n-1} d\sigma d\rho$$

$$\begin{aligned}
&\leq \int_0^r \int_{S^{n-1}} \int_0^\rho |\nabla u(\mathbf{x} + t\omega)| \rho^{n-1} dt d\sigma d\rho \\
&\leq \int_{S^{n-1}} \int_0^r \int_0^r |\nabla u(\mathbf{x} + t\omega)| \rho^{n-1} d\rho dt d\sigma \\
&\leq Cr^n \int_{S^{n-1}} \int_0^r |\nabla u(\mathbf{x} + t\omega)| dt d\sigma
\end{aligned}$$

$$= Cr^n \int_{S^{n-1}} \int_0^r \frac{|\nabla u(\mathbf{x} + t\omega)|}{t^{n-1}} t^{n-1} dt d\sigma$$

$$= Cr^n \int_{B(\mathbf{x},r)} \frac{|\nabla u(\mathbf{z})|}{|\mathbf{z} - \mathbf{x}|^{n-1}} dz.$$

Thus if we define

$$⨍_E f dx = \frac{1}{m(E)} \int_E f dx,$$

then

$$⨍_{B(\mathbf{x},r)} |u(\mathbf{y}) - u(\mathbf{x})| \, dy \leq C \int_{B(\mathbf{x},r)} |\nabla u(\mathbf{z})| |\mathbf{z} - \mathbf{x}|^{1-n} dz. \tag{19.1}$$

Now let $r = |\mathbf{x} - \mathbf{y}|$ and

$$U = B(\mathbf{x}, r), \ V = B(\mathbf{y}, r), \ W = U \cap V.$$

Thus W equals the intersection of two balls of radius r with the center of one on the boundary of the other. It is clear there exists a constant, C, depending only on n such that

$$\frac{m(W)}{m(U)} = \frac{m(W)}{m(V)} = C.$$

Then from (19.1),

$$|u(\mathbf{x}) - u(\mathbf{y})| = ⨍_W |u(\mathbf{x}) - u(\mathbf{y})| \, dz$$

$$\leq ⨍_W |u(\mathbf{x}) - u(\mathbf{z})| \, dz + ⨍_W |u(\mathbf{z}) - u(\mathbf{y})| \, dz$$

$$= \frac{C}{m(U)} \left[\int_W |u(\mathbf{x}) - u(\mathbf{z})| \, dz + \int_W |u(\mathbf{z}) - u(\mathbf{y})| \, dz \right]$$

$$\leq C \left[⨍_U |u(\mathbf{x}) - u(\mathbf{z})| \, dz + ⨍_V |u(\mathbf{y}) - u(\mathbf{z})| \, dz \right]$$

$$\leq C\left[\int_U |\nabla u(\mathbf{z})| |\mathbf{z}-\mathbf{x}|^{1-n} dz + \int_V |\nabla u(\mathbf{z})| |\mathbf{z}-\mathbf{y}|^{1-n} dz\right]. \tag{19.2}$$

Consider the first of these two integrals. This is no smaller than

$$\leq C\left(\int_{B(\mathbf{x},r)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} \left(\int_{B(\mathbf{x},r)} (|\mathbf{z}-\mathbf{x}|^{1-n})^{p/(p-1)} dz\right)^{(p-1)/p}$$

$$= C\left(\int_{B(\mathbf{x},r)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} \left(\int_0^r \int_{S^{n-1}} \rho^{p(1-n)/(p-1)} \rho^{n-1} d\sigma d\rho\right)^{(p-1)/p}$$

$$= C\left(\int_{B(\mathbf{x},r)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} \left(\int_0^r \rho^{(1-n)/(p-1)} d\rho\right)^{(p-1)/p}$$

$$= C\left(\int_{B(\mathbf{x},r)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} r^{(1-n/p)} \tag{19.3}$$

$$\leq C\left(\int_{B(\mathbf{x},2|\mathbf{x}-\mathbf{y}|)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} r^{(1-n/p)}. \tag{19.4}$$

The second integral in (19.2) is dominated by the same expression found in (19.3) except the ball over which the integral is taken is centered at $\mathbf{y}$ not $\mathbf{x}$. Thus this integral is also dominated by the expression in (19.4) and so,

$$|u(\mathbf{x}) - u(\mathbf{y})| \leq C\left(\int_{B(\mathbf{x},2|\mathbf{x}-\mathbf{y}|)} |\nabla u(\mathbf{z})|^p dz\right)^{1/p} |\mathbf{x}-\mathbf{y}|^{(1-n/p)} \tag{19.5}$$

which proves the lemma.

19.4 Rademacher's theorem

Next we extend this inequality to the case where we only have u and $u_{,i}$ in L^p_{loc} for $p > n$. This leads to an elegant proof of the differentiability a.e. of a Lipschitz continuous function. Let $\psi_k \in C_c^\infty(\mathbb{R}^n)$, $\psi_k \geq 0$, and $\psi_k(\mathbf{z}) = 1$ for all $\mathbf{z} \in B(\mathbf{0}, k)$. Then

$$u\psi_k,\ (u\psi_k)_{,i} \in L^p(\mathbb{R}^n).$$

Let ϕ_ϵ be a mollifier and consider

$$(u\psi_k)_\epsilon \equiv u\psi_k * \phi_\epsilon.$$

By Lemma 19.8,

$$(u\psi_k)_{\epsilon,i} = (u\psi_k)_{,i} * \phi_\epsilon.$$

Therefore

$$(u\psi_k)_{\epsilon,i} \to (u\psi_k)_{,i} \text{ in } L^p(\mathbb{R}^n) \tag{19.6}$$

and

$$(u\psi_k)_\epsilon \to u\psi_k \text{ in } L^p(\mathbb{R}^n) \tag{19.7}$$

as $\epsilon \to 0$. By (19.7), there exists a subsequence $\epsilon \to 0$ such that

$$(u\psi_k)_\epsilon(\mathbf{z}) \to u\psi_k(\mathbf{z}) \text{ a.e.} \tag{19.8}$$

Since $\psi_k(\mathbf{z}) = 1$ for $|\mathbf{z}| < k$, this shows

$$(u\psi_k)_\epsilon(\mathbf{z}) \to u(\mathbf{z}) \tag{19.9}$$

and for a.e. $\mathbf{z}$ with $|\mathbf{z}| < k$. Denoting the exceptional set of (19.9) by E_k, let

$$\mathbf{x}, \mathbf{y} \notin \cup_{k=1}^{\infty} E_k \equiv E.$$

Also let k be so large that

$$B(\mathbf{0},k) \supseteq B(\mathbf{x},2|\mathbf{x}-\mathbf{y}|).$$

Then by (19.5),

$$|(u\psi_k)_\epsilon(\mathbf{x}) - (u\psi_k)_\epsilon(\mathbf{y})|$$

$$\leq C\left(\int_{B(\mathbf{x},2|\mathbf{y}-\mathbf{x}|)} |\nabla(u\psi_k)_\epsilon|^p dz\right)^{1/p} |\mathbf{x}-\mathbf{y}|^{(1-n/p)}$$

where C depends only on n. Now by (19.8), there exists a subsequence, $\epsilon \to 0$, such that (19.9), holds for $\mathbf{z} = \mathbf{x}, \mathbf{y}$. Thus, from (19.6),

$$|u(\mathbf{x}) - u(\mathbf{y})| \leq C\left(\int_{B(\mathbf{x},2|\mathbf{y}-\mathbf{x}|)} |\nabla u|^p dz\right)^{1/p} |\mathbf{x}-\mathbf{y}|^{(1-n/p)}. \tag{19.10}$$

Redefining u on E, in the case where $p > n$, we can obtain (19.10) for all $\mathbf{x}, \mathbf{y}$. This has proved the following theorem.

Theorem 19.11 *Suppose $u, u_{,i} \in L^p_{loc}(\mathbb{R}^n)$ for $i = 1, \cdots, n$ and $p > n$. Then u has a representative, still denoted by u, such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,*

$$|u(\mathbf{x}) - u(\mathbf{y})| \leq C\left(\int_{B(\mathbf{x},2|\mathbf{y}-\mathbf{x}|)} |\nabla u|^p dz\right)^{1/p} |\mathbf{x}-\mathbf{y}|^{(1-n/p)}.$$

The next corollary is a very remarkable result. It says that not only is u continuous by virtue of having weak partial derivatives in L^p for large p, but also it is differentiable a.e.

Corollary 19.12 *Let $u, u_{,i} \in L^p_{loc}(\mathbb{R}^n)$ for $i = 1, \cdots, n$ and $p > n$. Then the representative of u described in Theorem 19.11 is differentiable a.e.*

Proof: Consider

$$|u(\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x}) \cdot (\mathbf{y}-\mathbf{x})|$$

where $u_{,i}$ is a representative of $u_{,i}$, an element of L^p. Define

$$g(\mathbf{z}) \equiv u(\mathbf{z}) + \nabla u(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{z}).$$

Then the above expression is of the form

$$|g(\mathbf{y}) - g(\mathbf{x})|$$

and

$$\nabla g(\mathbf{z}) = \nabla u(\mathbf{z}) - \nabla u(\mathbf{x}).$$

Therefore $g \in L^p_{loc}(\mathbb{R}^n)$ and $g_{,i} \in L^p_{loc}(\mathbb{R}^n)$. It follows from Theorem 19.11 that

$$|g(\mathbf{y}) - g(\mathbf{x})| = |u(\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})|$$

$$\leq C \left(\int_{B(\mathbf{x}, 2|\mathbf{y}-\mathbf{x}|)} |\nabla g(\mathbf{z})|^p dz \right)^{1/p} |\mathbf{x} - \mathbf{y}|^{(1-n/p)}$$

$$= C \left(\int_{B(\mathbf{x}, 2|\mathbf{y}-\mathbf{x}|)} |\nabla u(\mathbf{z}) - \nabla u(\mathbf{x})|^p dz \right)^{1/p} |\mathbf{x} - \mathbf{y}|^{(1-n/p)}$$

$$= C \left(\fint_{B(\mathbf{x}, 2|\mathbf{y}-\mathbf{x}|)} |\nabla u(\mathbf{z}) - \nabla u(\mathbf{x})|^p dz \right)^{1/p} |\mathbf{x} - \mathbf{y}|.$$

This last expression is $o(|\mathbf{y} - \mathbf{x}|)$ at every Lebesgue point, $\mathbf{x}$, of ∇u. This proves the corollary and shows ∇u is the gradient a.e.

Now suppose u is Lipschitz on $\mathbb{R}^n$,

$$|u(\mathbf{x}) - u(\mathbf{y})| \leq K |\mathbf{x} - \mathbf{y}|$$

for some constant K. We define $\operatorname{Lip}(u)$ as the smallest value of K that works in this inequality. The following corollary is known as Rademacher's theorem. It states that every Lipschitz function is differentiable a.e.

Corollary 19.13 *If u is Lipschitz continuous then u is differentiable a.e. and* $||u_{,i}||_\infty \leq \operatorname{Lip}(u)$.

Proof: We do this by showing that Lipschitz continuous functions have weak derivatives in $L^\infty(\mathbb{R}^n)$ and then using the previous results. Let

$$D^h_{\mathbf{e}_i} u(x) \equiv h^{-1} [u(x + h\mathbf{e}_i) - u(x)].$$

Then $D^h_{\mathbf{e}_i} u$ is bounded in $L^\infty(\mathbb{R}^n)$ and

$$||D^h_{\mathbf{e}_i} u||_\infty \leq \operatorname{Lip}(u).$$

It follows that $D^h_{\mathbf{e}_i} u$ is contained in a ball in $L^\infty(\mathbb{R}^n)$, the dual space of $L^1(\mathbb{R}^n)$. By Theorem 18.30 there is a subsequence $h \to 0$ such that

$$D^h_{\mathbf{e}_i} u \rightharpoonup w, \ ||w||_\infty \leq \operatorname{Lip}(u) \tag{19.11}$$

where the convergence takes place in the weak $*$ topology of $L^\infty(\mathbb{R}^n)$. Let $\phi \in C_c^\infty(\mathbb{R}^n)$. Then

$$\int w\phi dx = \lim_{h\to 0} \int D_{\mathbf{e}_i}^h u\phi dx$$

$$= \lim_{h\to 0} \int u(\mathbf{x}) \frac{(\phi(\mathbf{x} - h\mathbf{e}_i) - \phi(\mathbf{x}))}{h} dx$$

$$= -\int u(\mathbf{x}) \phi_{,i}(\mathbf{x}) dx.$$

Thus $w = u_{,i}$ and we see that $u_{,i} \in L^\infty(\mathbb{R}^n)$ for each i. Hence $u, u_{,i} \in L_{loc}^p(\mathbb{R}^n)$ for all $p > n$ and so u is differentiable a.e. and ∇u is given in terms of the weak derivatives of u by Corollary 19.12. This proves the corollary.

19.5 Exercises

1. Let K be a bounded subset of $L^p(\mathbb{R}^n)$ and suppose that for all $\epsilon > 0$, there exist a $\delta > 0$ and G such that $\overline{G}$ is compact such that if $|\mathbf{h}| < \delta$, then

   $$\int |u(\mathbf{x} + \mathbf{h}) - u(\mathbf{x})|^p dx < \epsilon^p$$

   for all $u \in K$ and

   $$\int_{\mathbb{R}^n \setminus \overline{G}} |u(\mathbf{x})|^p dx < \epsilon^p$$

   for all $u \in K$. Show that K is precompact in $L^p(\mathbb{R}^n)$. **Hint:** Let ϕ_k be a mollifier and consider

   $$K_k \equiv \{u * \phi_k : u \in K\}.$$

   Verify the conditions of the Ascoli Arzela theorem for these functions defined on $\overline{G}$ and show there is an ϵ net for each $\epsilon > 0$. Can you modify this to let an arbitrary open set take the place of $\mathbb{R}^n$?

2. In (19.11), why is $||w||_\infty \leq \text{Lip}(u)$?

3. Suppose $D_{\mathbf{e}_i}^h u$ is bounded in $L^p(\mathbb{R}^n)$ for $p > n$. Can we conclude as in Corollary 19.13 that u is differentiable a.e.?

4. Show that a closed subspace of a reflexive Banach space is reflexive. **Hint:** The proof of this is an exercise in the use of the Hahn Banach theorem. Let Y be the closed subspace of the reflexive space X and let $y^{**} \in Y''$. Then $i^{**}y^{**} \in X''$ and so $i^{**}y^{**} = Jx$ for some $x \in X$ because X is reflexive. Now argue that $x \in Y$ as follows. If $x \notin Y$, then there exists x^* such that $x^*(Y) = 0$ but $x^*(x) \neq 0$. Thus, $i^*x^* = 0$. Use this to get a contradiction. When you know that $x = y \in Y$, the Hahn Banach theorem implies i^* is onto Y' and for all $x^* \in X'$,

   $$y^{**}(i^*x^*) = i^{**}y^{**}(x^*) = Jx(x^*) = x^*(x) = x^*(iy) = i^*x^*(y).$$

5. For an arbitrary open set $U \subseteq \mathbb{R}^n$, we define $X^{1p}(U)$ as the set of all functions in $L^p(U)$ whose weak partial derivatives are also in $L^p(U)$. Here we say a function in $L^p(U)$, g equals $u_{,i}$ if and only if

   $$\int_U g\phi dx = -\int_U u\phi_{,i} dx$$

for all $\phi \in C_c^\infty (U)$. The norm in this space is given by

$$||u||_{1p} \equiv \left(\int_U |u|^p + |\nabla u|^p \, dx \right)^{1/p}.$$

Then we define the Sobolev space $W^{1p}(U)$ to be the closure of $C^\infty\left(\overline{U}\right)$ in $X^{1p}(U)$ where $C^\infty\left(\overline{U}\right)$ is defined to be restrictions of all functions in $C_c^\infty(\mathbb{R}^n)$ to U. Show that this definition of weak derivative is well defined and that $X^{1p}(U)$ is a reflexive Banach space. **Hint:** To do this, show the operator $u \to u_{,i}$ is a closed operator and that $X^{1p}(U)$ can be considered as a closed subspace of $L^p(U)^{n+1}$. Show that in general the product of reflexive spaces is reflexive and then use Problem 4 above which states that a closed subspace of a reflexive Banach space is reflexive. Thus, conclude that $W^{1p}(U)$ is also a reflexive Banach space.

6. Theorem 19.11 shows that if the weak derivatives of a function $u \in L^p(\mathbb{R}^n)$ are in $L^p(\mathbb{R}^n)$, for $p > n$, then the function has a continuous representative. (In fact, one can conclude more than continuity from this theorem.) It is also important to consider the case when $p < n$. To aid in the study of this case which will be carried out in the next few problems, show the following inequality for $n \geq 2$.

$$\int_{\mathbb{R}^n} \prod_{j=1}^n |w_j(\mathbf{x})| \, dm_n \leq \prod_{i=1}^n \left(\int_{\mathbb{R}^{n-1}} |w_j(\mathbf{x})|^{n-1} \, dm_{n-1} \right)^{1/n-1}$$

where w_j does not depend on the jth component of $\mathbf{x}$, x_j. **Hint:** First show it is true for $n = 2$ and then use Holder's inequality and induction. You might benefit from first trying the case $n = 3$ to get the idea.

7. $\uparrow$ Show that if $\phi \in C_c^\infty(\mathbb{R}^n)$, then

$$||\phi||_{n/(n-1)} \leq \frac{1}{\sqrt[n]{n}} \sum_{j=1}^n \left|\left| \frac{\partial \phi}{\partial x_j} \right|\right|_1.$$

Hint: First show that if $a_i \geq 0$, then

$$\prod_{i=1}^n a_i^{1/n} \leq \frac{1}{\sqrt[n]{n}} \sum_{j=1}^n a_i.$$

Then observe that

$$|\phi(\mathbf{x})| \leq \int_{-\infty}^{\infty} \left|\phi_{,j}(\mathbf{x})\right| dx_j$$

so

$$||\phi||_{n/(n-1)}^{n/(n-1)} = \int |\phi|^{n/(n-1)} \, dm_n$$

$$\leq \int \prod_{j=1}^n \left(\int_{-\infty}^{\infty} \left|\phi_{,j}(\mathbf{x})\right| dx_j \right)^{1/(n-1)} dm_n$$

$$\leq \prod_{j=1}^n \left(\int \left|\phi_{,j}(\mathbf{x})\right| dm_n \right)^{1/(n-1)}.$$

Hence

$$||\phi||_{n/(n-1)} \leq \prod_{j=1}^n \left(\int \left|\phi_{,j}(\mathbf{x})\right| dm_n \right)^{1/n}.$$

8. $\uparrow$ Show that if $\phi \in C_c^\infty(\mathbb{R}^n)$, then if $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$, where $p < n$, then

$$||\phi||_q \leq \frac{1}{\sqrt[n]{n}} \frac{(n-1)\,p}{n-p} \sum_{j=1}^n ||\phi_{,i}||_p.$$

Also show that if $u \in W^{1p}(\mathbb{R}^n)$, then $u \in L^q(\mathbb{R}^n)$ and the inclusion map is continuous. This is part of the Sobolev embedding theorem. For more on Sobolev spaces see Adams [1]. **Hint:** Let $r > 1$. Then $|\phi|^r \in C_c^\infty(\mathbb{R}^n)$ and

$$\left| |\phi|^r_{,i} \right| = r\,|\phi|^{r-1} \left|\phi_{,i}\right|.$$

Now apply the result of Problem 7 to write

$$\left(\int |\phi|^{\frac{rn}{n-1}} \, dm_n \right)^{(n-1)/n} \leq \frac{r}{\sqrt[n]{n}} \sum_{i=1}^n \int |\phi|^{r-1} \left|\phi_{,i}\right| dm_n$$

$$\leq \frac{r}{\sqrt[n]{n}} \sum_{i=1}^n \left(\int \left|\phi_{,i}\right|^p \right)^{1/p} \left(\int \left(|\phi|^{r-1} \right)^{p/(p-1)} dm_n \right)^{(p-1)/p}.$$

Now choose r such that

$$\frac{(r-1)\,p}{p-1} = \frac{rn}{n-1}$$

so that the last term on the right can be cancelled with the first term on the left and simplify.

Fundamental Theorem of Calculus

One of the most remarkable theorems in Lebesgue integration is the Lebesgue fundamental theorem of calculus which says that if f is a function in L^1, then the indefinite integral,

$$x \to \int_a^x f(t)\, dt,$$

can be differentiated for a.e. x and gives $f(x)$ a.e. This is a very significant generalization of the usual fundamental theorem of calculus found in calculus. To prove this theorem, we use a covering theorem due to Vitali and theorems on maximal functions. This approach leads to very general results without very many painful technicalities. We will be in the context of $(\mathbb{R}^n, \mathcal{S}, m)$ where m is n-dimensional Lebesgue measure. When this important theorem is established, it will be used to prove the very useful theorem about change of variables in multiple integrals.

By Lemma 6.6 of Chapter 6 and the completeness of m, we know that the Lebesgue measurable sets are exactly those measurable in the sense of Caratheodory. Also, we can regard m as an outer measure defined on all of $\mathcal{P}(\mathbb{R}^n)$. We will use the following notation.

$$B(\mathbf{p}, r) = \{\mathbf{x} : |\mathbf{x} - \mathbf{p}| < r\}. \tag{20.1}$$

If

$$B = B(\mathbf{p}, r), \text{ then } \hat{B} = B(\mathbf{p}, 5r).$$

20.1 The Vitali covering theorem

Lemma 20.1 *Let $\mathcal{F}$ be a collection of balls as in (20.1). Suppose*

$$\infty > M \equiv \sup\{r : B(\mathbf{p}, r) \in \mathcal{F}\} > 0.$$

Then there exists $\mathcal{G} \subseteq \mathcal{F}$ such that

$$\textit{if } B(\mathbf{p}, r) \in \mathcal{G} \textit{ then } r > \frac{M}{2}, \tag{20.2}$$

$$\textit{if } B_1, B_2 \in \mathcal{G} \textit{ then } B_1 \cap B_2 = \emptyset, \tag{20.3}$$

$\mathcal{G}$ is maximal with respect to Formulas (20.2) and (20.3).

Proof: Let $\mathcal{H} = \{\mathcal{B} \subseteq \mathcal{F}$ such that (20.2) and (20.3) hold$\}$. Obviously $\mathcal{H} \neq \emptyset$ because there exists $B(\mathbf{p}, r) \in \mathcal{F}$ with $r > \frac{M}{2}$. Partially order $\mathcal{H}$ by set inclusion and use the Hausdorff maximal theorem (see the appendix on set theory) to let $\mathcal{C}$ be a maximal chain in $\mathcal{H}$. Clearly $\cup\mathcal{C}$ satisfies (20.2) and (20.3). If $\cup\mathcal{C}$ is not maximal with respect to these two properties, then $\mathcal{C}$ was not a maximal chain. Let $\mathcal{G} = \cup\mathcal{C}$.

Theorem 20.2 *(Vitali) Let $\mathcal{F}$ be a collection of balls and let*

$$A \equiv \cup\{B : B \in \mathcal{F}\}.$$

Suppose

$$\infty > M \equiv \sup\{r : B(\mathbf{p}, r) \in \mathcal{F}\} > 0.$$

Then there exists $\mathcal{G} \subseteq \mathcal{F}$ such that $\mathcal{G}$ consists of disjoint balls and

$$A \subseteq \cup\{\widehat{B} : B \in \mathcal{G}\}.$$

Proof: Let $\mathcal{G}_1 \subseteq \mathcal{F}$ satisfy

$$B(\mathbf{p}, r) \in \mathcal{G}_1 \text{ implies } r > \frac{M}{2}, \tag{20.4}$$

$$B_1, B_2 \in \mathcal{G}_1 \text{ implies } B_1 \cap B_2 = \emptyset, \tag{20.5}$$

$$\mathcal{G}_1 \text{ is maximal with respect to Formulas (20.4), and (20.5).}$$

Suppose $\mathcal{G}_1, \cdots, \mathcal{G}_{m-1}$ have been chosen, $m \geq 2$. Let

$$\mathcal{F}_m = \{B \in \mathcal{F} : B \subseteq \mathbb{R}^n \setminus \cup\{\mathcal{G}_1 \cup \cdots \cup \mathcal{G}_{m-1}\}\}.$$

Let $\mathcal{G}_m \subseteq \mathcal{F}_m$ satisfy the following.

$$B(\mathbf{p}, r) \in \mathcal{G}_m \text{ implies } r > \frac{M}{2^m}, \tag{20.6}$$

$$B_1, B_2 \in \mathcal{G}_m \text{ implies } B_1 \cap B_2 = \emptyset, \tag{20.7}$$

$$\mathcal{G}_m \text{ is a maximal subset of } \mathcal{F}_m \text{ with respect to Formulas (20.6) and (20.7).}$$

If $\mathcal{F}_m = \emptyset$, $\mathcal{G}_m = \emptyset$. Define

$$\mathcal{G} \equiv \cup_{k=1}^{\infty} \mathcal{G}_k.$$

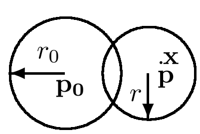
Thus $\mathcal{G}$ is a collection of disjoint balls in $\mathcal{F}$. We need to show $\{\widehat{B} : B \in \mathcal{G}\}$ covers A. Let $\mathbf{x} \in A$. Then $\mathbf{x} \in B(\mathbf{p}, r) \in \mathcal{F}$. Pick m such that

$$\frac{M}{2^m} < r \leq \frac{M}{2^{m-1}}.$$

We claim $\mathbf{x} \in \widehat{B}$ for some $B \in \mathcal{G}_1 \cup \cdots \cup \mathcal{G}_m$. To see this, note that $B(\mathbf{p}, r)$ must intersect some set of $\mathcal{G}_1 \cup \cdots \cup \mathcal{G}_m$ because if it didn't, then

$$\mathcal{G}_m \cup \{B(\mathbf{p}, r)\} = \mathcal{G}'_m$$

would satisfy Formulas (20.6) and (20.7), and $\mathcal{G}'_m \supsetneq \mathcal{G}_m$ contradicting the maximality of $\mathcal{G}_m$. Let the set intersected be $B(\mathbf{p}_0, r_0)$. Thus $r_0 > M2^{-m}$.



Then if $\mathbf{x} \in B(\mathbf{p}, r)$,

$$\begin{aligned} |\mathbf{x} - \mathbf{p}_0| &\leq |\mathbf{x} - \mathbf{p}| + |\mathbf{p} - \mathbf{p}_0| < r + r_0 + r \\ &\leq \frac{2M}{2^{m-1}} + r_0 < 4r_0 + r_0 = 5r_0 \end{aligned}$$

since $r_0 > M/2^{-m}$. Hence $B(\mathbf{p}, r) \subseteq B(\mathbf{p}_0, 5r_0)$ and this proves the theorem.

20.2 Differentiation with respect to Lebesgue measure

The covering theorem just presented will now be used to establish the fundamental theorem of calculus. In discussing this, we introduce the space of functions which is locally integrable in the following definition. This space of functions is the most general one for which the maximal function defined below makes sense.

Definition 20.3 *$f \in L^1_{loc}(\mathbb{R}^n)$ means $f\mathcal{X}_{B(0,R)} \in L^1(\mathbb{R}^n)$ for all $R > 0$. For $f \in L^1_{loc}(\mathbb{R}^n)$, the Hardy Littlewood Maximal Function, Mf, is defined by*

$$Mf(\mathbf{x}) \equiv \sup_{r>0} \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y})|dy.$$

Theorem 20.4 *If $f \in L^1(\mathbb{R}^n)$, then for $\alpha > 0$,*

$$\overline{m}([Mf > \alpha]) \leq \frac{5^n}{\alpha} ||f||_1.$$

(Here and elsewhere, $[Mf > \alpha] \equiv \{\mathbf{x} \in \mathbb{R}^n : Mf(\mathbf{x}) > \alpha\}$ with other occurrences of $[\quad]$ being defined similarly.)

Proof: Let $S \equiv [Mf > \alpha]$. For $\mathbf{x} \in S$, choose $r_{\mathbf{x}} > 0$ with

$$\frac{1}{m(B(\mathbf{x}, r_{\mathbf{x}}))} \int_{B(\mathbf{x},r_{\mathbf{x}})} |f| \, dm > \alpha.$$

The $r_{\mathbf{x}}$ are all bounded because

$$m(B(\mathbf{x}, r_{\mathbf{x}})) < \frac{1}{\alpha} \int_{B(\mathbf{x},r_{\mathbf{x}})} |f| \, dm < \frac{1}{\alpha} ||f||_1.$$

By the Vitali covering theorem, there are disjoint balls $B(\mathbf{x}_i, r_i)$ such that

$$S \subseteq \cup_{\mathbf{x} \in S} B(\mathbf{x}, r_{\mathbf{x}}) \subseteq \cup_{i=1}^{\infty} B(\mathbf{x}_i, 5r_i)$$

and

$$\frac{1}{m(B(\mathbf{x}_i, r_i))} \int_{B(\mathbf{x}_i,r_i)} |f| \, dm > \alpha.$$

Therefore

$$\begin{aligned}
\overline{m}(S) &\leq \sum_{i=1}^{\infty} m(B(\mathbf{x}_i, 5r_i)) = 5^n \sum_{i=1}^{\infty} m(B(\mathbf{x}_i, r_i)) \\
&\leq \frac{5^n}{\alpha} \sum_{i=1}^{\infty} \int_{B(\mathbf{x}_i,r_i)} |f| \, dm \\
&\leq \frac{5^n}{\alpha} \int_{\mathbb{R}^n} |f| \, dm,
\end{aligned}$$

the last inequality being valid because the balls $B(\mathbf{x}_i, r_i)$ are disjoint. This proves the theorem.

Lemma 20.5 *Let $f \geq 0$, and $f \in L^1$, then*

$$\lim_{r \to 0} \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} f(\mathbf{y})dy = f(\mathbf{x}) \quad \text{a.e. } \mathbf{x}.$$

Proof: Let $\alpha > 0$ and let

$$B_\alpha = \left[\limsup_{r\to 0}\left|\frac{1}{m(B(\mathbf{x},r))}\int_{B(\mathbf{x},r)} f(\mathbf{y})dy - f(\mathbf{x})\right| > \alpha\right].$$

Then for any $g \in C_c(\mathbb{R}^n)$, B_α equals

$$\left[\limsup_{r\to 0}\left|\frac{1}{m(B(\mathbf{x},r))}\int_{B(\mathbf{x},r)} f(\mathbf{y}) - g(\mathbf{y})dy - (f(\mathbf{x}) - g(\mathbf{x}))\right| > \alpha\right]$$

because for any $g \in C_c(\mathbb{R}^n)$,

$$\lim_{r\to 0}\frac{1}{m(B(\mathbf{x},r))}\int_{B(\mathbf{x},r)} g(\mathbf{y})dy = g(\mathbf{x}).$$

Thus

$$B_\alpha \subseteq [M(|f-g|) + |f-g| > \alpha],$$

and so

$$B_\alpha \subseteq \left[M(|f-g|) > \frac{\alpha}{2}\right] \cup \left[|f-g| > \frac{\alpha}{2}\right].$$

Now

$$\frac{\alpha}{2} m\left(\left[|f-g| > \frac{\alpha}{2}\right]\right) = \frac{\alpha}{2}\int_{[|f-g|>\frac{\alpha}{2}]} dx$$

$$\leq \int_{[|f-g|>\frac{\alpha}{2}]} |f-g|dx \leq ||f-g||_1.$$

Therefore by Theorem 20.4,

$$\overline{m}(B_\alpha) \leq \left(\frac{2(5^n)}{\alpha} + \frac{2}{\alpha}\right)||f-g||_1.$$

Since $C_c(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$ and g is arbitrary, this estimate shows $\overline{m}(B_\alpha) = 0$. It follows by Lemma 6.6, since B_α is measurable in the sense of Caratheodory, that B_α is Lebesgue measurable and $m(B_\alpha) = 0$.

$$\left[\limsup_{r\to 0}\left|\frac{1}{m(B(\mathbf{x},r))}\int_{B(\mathbf{x},r)} f(\mathbf{y})dy - f(\mathbf{x})\right| > 0\right] \subseteq \cup_{m=1}^{\infty} B_{\frac{1}{m}} \tag{20.8}$$

and each set $B_{1/m}$ has measure 0 so the set on the left in (20.8) is also Lebesgue measurable and has measure 0. Thus, if $\mathbf{x}$ is not in this set,

$$\begin{aligned} 0 &= \limsup_{r\to 0}\left|\frac{1}{m(B(\mathbf{x},r))}\int_{B(\mathbf{x},r)} f(\mathbf{y})dy - f(\mathbf{x})\right| \geq \\ &\geq \liminf_{r\to 0}\left|\frac{1}{m(B(\mathbf{x},r))}\int_{B(\mathbf{x},r)} f(\mathbf{y})dy - f(\mathbf{x})\right| \geq 0. \end{aligned}$$

This proves the lemma.

Corollary 20.6 *If $f \geq 0$ and $f \in L^1_{loc}(\mathbb{R}^n)$, then*

$$\lim_{r \to 0} \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} f(\mathbf{y}) dy = f(\mathbf{x}) \quad a.e.\ \mathbf{x}. \tag{20.9}$$

Proof: Apply Lemma 20.5 to $f\mathcal{X}_{B(0,R)}$ for $R = 1, 2, 3, \cdots$. Thus (20.9) holds for a.e. $\mathbf{x} \in B(0, R)$ for each $R = 1, 2, \cdots$.

Theorem 20.7 *(Fundamental Theorem of Calculus) Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then there exists a set of measure $0, B$, such that if $\mathbf{x} \notin B$, then*

$$\lim_{r \to 0} \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - f(\mathbf{x})| dy = 0.$$

Proof: Let $\{d_i\}_{i=1}^{\infty}$ be a countable dense subset of $\mathbb{C}$. By Corollary 20.6, there exists a set of measure 0, B_i, such that if $\mathbf{x} \notin B_i$

$$\lim_{r \to 0} \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - d_i| dy = |f(\mathbf{x}) - d_i|. \tag{20.10}$$

Let $B = \cup_{i=1}^{\infty} B_i$ and let $\mathbf{x} \notin B$. Pick d_i such that $|f(\mathbf{x}) - d_i| < \frac{\varepsilon}{2}$. Then

$$\frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - f(\mathbf{x})| dy \leq \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - d_i| dy$$

$$+ \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{x}) - d_i| dy$$

$$\leq \frac{\varepsilon}{2} + \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - d_i| dy.$$

By (20.10)

$$\frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} |f(\mathbf{y}) - f(\mathbf{x})| dy \leq \varepsilon$$

whenever r is small enough. This proves the theorem.

Definition 20.8 *For B the set of Theorem 20.7, B^C is called the Lebesgue set or the set of Lebesgue points.*

Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then

$$\lim_{r \to 0} \frac{1}{m(B(\mathbf{x}, r))} \int_{B(\mathbf{x},r)} f(\mathbf{y}) dy = f(\mathbf{x}) \quad \text{a.e. } \mathbf{x}.$$

The next corollary is a one dimensional version of what was just presented.

Corollary 20.9 *Let $f \in L^1(\mathbb{R})$ and let*

$$F(x) = \int_{-\infty}^{x} f(t) dt.$$

Then for a.e. x, $F'(x) = f(x)$.

Proof: For $h > 0$

$$\frac{1}{h}\int_x^{x+h} |f(y) - f(x)| dy \leq 2(\frac{1}{2h})\int_{x-h}^{x+h} |f(y) - f(x)| dy$$

By Theorem 20.7, this converges to 0 a.e. Similarly

$$\frac{1}{h}\int_{x-h}^{x} |f(y) - f(x)| dy$$

converges to 0 a.e. x.

$$\left|\frac{F(x+h) - F(x)}{h} - f(x)\right| \leq \frac{1}{h}\int_x^{x+h} |f(y) - f(x)| dy$$

and

$$\left|\frac{F(x) - F(x-h)}{h} - f(x)\right| \leq \frac{1}{h}\int_{x-h}^{x} |f(y) - f(x)| dy.$$

Therefore,

$$\lim_{h\to 0}\frac{F(x+h) - F(x)}{h} = f(x) \text{ a.e. } x$$

This proves the corollary.

20.3 The change of variables formula for Lipschitz maps

This section is on a generalization of the change of variables formula for multiple integrals presented in Chapter 11. In this section, Ω will be a Lebesgue measurable set in $\mathbb{R}^n$ and $\mathbf{h} : \Omega \to \mathbb{R}^n$ will be Lipschitz. We recall Rademacher's theorem a proof of which was given in Chapter 19.

Theorem 20.10 *Let* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ *be Lipschitz. Then* $D\mathbf{f}(\mathbf{x})$ *exists a.e.and* $||f_{i,j}||_\infty \leq \text{Lip}(f)$.

It turns out that a Lipschitz function defined on some subset of $\mathbb{R}^n$ always has a Lipschitz extension to all of $\mathbb{R}^n$. The next theorem gives a proof of this. For more on this sort of theorem we refer to [12].

Theorem 20.11 *If* $\mathbf{h} : \Omega \to \mathbb{R}^m$ *is Lipschitz, then there exists* $\overline{\mathbf{h}} : \mathbb{R}^n \to \mathbb{R}^m$ *which extends* $\mathbf{h}$ *and is also Lipschitz.*

Proof: It suffices to assume $m = 1$ because if this is shown, it may be applied to the components of $\mathbf{h}$ to get the desired result. Suppose

$$|h(\mathbf{x}) - h(\mathbf{y})| \leq K |\mathbf{x} - \mathbf{y}|. \tag{20.11}$$

Define

$$\bar{h}(\mathbf{x}) \equiv \inf\{h(\mathbf{w}) + K|\mathbf{x} - \mathbf{w}| : \mathbf{w} \in \Omega\}. \tag{20.12}$$

If $\mathbf{x} \in \Omega$, then for all $\mathbf{w} \in \Omega$,

$$h(\mathbf{w}) + K|\mathbf{x} - \mathbf{w}| \geq h(\mathbf{x})$$

by (20.11). This shows $h(\mathbf{x}) \leq \bar{h}(\mathbf{x})$. But also we can take $\mathbf{w} = \mathbf{x}$ in (20.12) which yields $\bar{h}(\mathbf{x}) \leq h(\mathbf{x})$. Therefore $\bar{h}(\mathbf{x}) = h(\mathbf{x})$ if $\mathbf{x} \in \Omega$.

Now suppose $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ and consider $\left|\bar{h}(\mathbf{x})-\bar{h}(\mathbf{y})\right|$. Without loss of generality we may assume $\bar{h}(\mathbf{x})\geq\bar{h}(\mathbf{y})$. (If not, repeat the following argument with $\mathbf{x}$ and $\mathbf{y}$ interchanged.) Pick $\mathbf{w}\in\Omega$ such that

$$h(\mathbf{w})+K|\mathbf{y}-\mathbf{w}|-\epsilon<\bar{h}(\mathbf{y}).$$

Then

$$\left|\bar{h}(\mathbf{x})-\bar{h}(\mathbf{y})\right|=\bar{h}(\mathbf{x})-\bar{h}(\mathbf{y})\leq h(\mathbf{w})+K|\mathbf{x}-\mathbf{w}|-$$

$$[h(\mathbf{w})+K|\mathbf{y}-\mathbf{w}|-\epsilon]\leq K|\mathbf{x}-\mathbf{y}|+\epsilon.$$

Since ϵ is arbitrary,

$$\left|\bar{h}(\mathbf{x})-\bar{h}(\mathbf{y})\right|\leq K|\mathbf{x}-\mathbf{y}|$$

and this proves the theorem.

We will use $\overline{\mathbf{h}}$ to denote a Lipschitz extension of the Lipschitz function $\mathbf{h}$. From now on $\mathbf{h}$ will denote a Lipschitz map from a measurable set in $\mathbb{R}^n$ to $\mathbb{R}^n$. The next lemma is an application of the Vitali covering theorem. It states that every open set can be filled with disjoint balls except for a set of measure zero.

Lemma 20.12 *Let V be an open set in $\mathbb{R}^r, m_r(V)<\infty$. Then there exists a sequence of disjoint open balls $\{B_i\}$ having radii less than δ and a set of measure 0, T, such that*

$$V=(\cup_{i=1}^{\infty}B_i)\cup T.$$

Proof: This is left as a problem. See Problem 8 in this chapter.

We wish to show that $\mathbf{h}$ maps Lebesgue measurable sets to Lebesgue measurable sets. In showing this the key result is the next lemma which states that $\mathbf{h}$ maps sets of measure zero to sets of measure zero.

Lemma 20.13 *If $m_n(T)=0$ then $m_n\left(\overline{\mathbf{h}}(T)\right)=0$.*

Proof: Let V be an open set containing T whose measure is less than ϵ. Now using the Vitali covering theorem, there exists a sequence of disjoint balls $\{B_i\}$, $B_i=B(\mathbf{x}_i,r_i)$, which are contained in V such that the sequence of enlarged balls, $\left\{\widehat{B}_i\right\}$, having the same center but 5 times the radius, covers T. Then

$$m_n\left(\overline{\mathbf{h}}(T)\right)\leq m_n\left(\overline{\mathbf{h}}\left(\cup_{i=1}^{\infty}\widehat{B}_i\right)\right)$$

$$\leq\sum_{i=1}^{\infty}m_n\left(\overline{\mathbf{h}}\left(\widehat{B}_i\right)\right)$$

$$\begin{aligned}\leq&\quad\sum_{i=1}^{\infty}\alpha(n)\left(\mathrm{Lip}\left(\overline{\mathbf{h}}\right)\right)^n5^nr_i^n=5^n\left(\mathrm{Lip}\left(\overline{\mathbf{h}}\right)\right)^n\sum_{i=1}^{\infty}m_n(B_i)\\ \leq&\quad\left(\mathrm{Lip}\left(\overline{\mathbf{h}}\right)\right)^n5^nm_n(V)\leq\epsilon\left(\mathrm{Lip}\left(\overline{\mathbf{h}}\right)\right)^n5^n.\end{aligned}$$

Since ϵ is arbitrary, this proves the lemma.

Actually, the argument in this lemma holds in other contexts which do not imply $\mathbf{h}$ is Lipschitz continuous. For one such example, see Problem 23.

With the conclusion of this lemma, the next lemma is fairly easy to obtain.

Lemma 20.14 *If A is Lebesgue measurable, then $\bar{\mathbf{h}}(A)$ is Lebesgue measurable. Furthermore,*

$$m_n\left(\bar{\mathbf{h}}(A)\right) \leq \left(\operatorname{Lip}\left(\bar{\mathbf{h}}\right)\right)^n m_n(A). \tag{20.13}$$

Proof: Let $A_k = A \cap B(\mathbf{0}, k), k \in \mathbb{N}$. We establish (20.13) for A_k in place of A and then let $k \to \infty$ to obtain (20.13). Let $V \supseteq A_k$ and let $m_n(V) < \infty$. By Lemma 20.12, there is a sequence of disjoint balls $\{B_i\}$, and a set of measure 0, T, such that

$$V = \cup_{i=1}^{\infty} B_i \cup T,\ B_i = B(x_i, r_i).$$

Then by Lemma 20.13,

$$m_n\left(\bar{\mathbf{h}}(A_k)\right) \leq m_n\left(\bar{\mathbf{h}}(V)\right)$$

$$\leq m_n\left(\bar{\mathbf{h}}\left(\cup_{i=1}^{\infty} B_i\right)\right) + m_n\left(\bar{\mathbf{h}}(T)\right) = m_n\left(\bar{\mathbf{h}}\left(\cup_{i=1}^{\infty} B_i\right)\right)$$

$$\leq \sum_{i=1}^{\infty} m_n\left(\bar{\mathbf{h}}(B_i)\right) \leq \sum_{i=1}^{\infty} m_n\left(B\left(\bar{\mathbf{h}}(x_i), \operatorname{Lip}\left(\bar{\mathbf{h}}\right) r_i\right)\right)$$

$$\leq \sum_{i=1}^{\infty} \alpha(n)\left(\operatorname{Lip}\left(\bar{\mathbf{h}}\right) r_i\right)^n = \operatorname{Lip}\left(\bar{\mathbf{h}}\right)^n \sum_{i=1}^{\infty} m_n(B_i) = \operatorname{Lip}\left(\bar{\mathbf{h}}\right)^n m_n(V).$$

Since V is an arbitrary open set containing A_k, it follows from regularity of Lebesgue measure that

$$m_n\left(\bar{\mathbf{h}}(A_k)\right) \leq \operatorname{Lip}\left(\bar{\mathbf{h}}\right)^n m_n(A_k). \tag{20.14}$$

Now let $k \to \infty$ to obtain (20.13). This proves the formula. It remains to show $\bar{\mathbf{h}}(A)$ is Lebesgue measurable.

By inner regularity of Lebesgue measure, there exists a set, F, which is the countable union of compact sets and a set T with $m_n(T) = 0$ such that

$$F \cup T = A_k.$$

Then $\bar{\mathbf{h}}(F) \subseteq \bar{\mathbf{h}}(A_k) \subseteq \bar{\mathbf{h}}(F) \cup \bar{\mathbf{h}}(T)$. By continuity of $\bar{\mathbf{h}}$, $\bar{\mathbf{h}}(F)$ is a countable union of compact sets and so it is Borel. By (20.14) with T in place of A_k,

$$m_n\left(\bar{\mathbf{h}}(T)\right) = 0$$

and so $\bar{\mathbf{h}}(T)$ is Lebesgue measurable. Therefore, $\bar{\mathbf{h}}(A_k)$ is Lebesgue measurable because m_n is a complete measure and we have exhibited $\bar{\mathbf{h}}(A_k)$ between two Lebesgue measurable sets whose difference has measure 0. Now

$$\bar{\mathbf{h}}(A) = \cup_{k=1}^{\infty} \bar{\mathbf{h}}(A_k)$$

so $\bar{\mathbf{h}}(A)$ is also Lebesgue measurable and this proves the lemma.

The following lemma, found in Rudin [25], is interesting for its own sake and will serve as the basis for many of the theorems and lemmas which follow. Its proof is based on the Brouwer fixed point theorem, a short proof of which is given in the chapter on the Brouwer degree. The idea is that if a continuous function mapping a ball in $\mathbb{R}^k$ to $\mathbb{R}^k$ doesn't move any point very much, then the image of the ball must contain a slightly smaller ball.

Lemma 20.15 *Let $B = B(\mathbf{0}, r)$, a ball in $\mathbb{R}^k$ and let $\mathbf{F} : \overline{B} \to \mathbb{R}^k$ be continuous and suppose for some $\epsilon < 1$,*

$$|\mathbf{F}(\mathbf{v}) - \mathbf{v}| < \epsilon r$$

for all $\mathbf{v} \in \overline{B}$. Then

$$\mathbf{F}\left(\overline{B}\right) \supseteq B(\mathbf{0}, r(1-\epsilon)).$$

Proof: Suppose $\mathbf{a} \in B(\mathbf{0}, r(1-\epsilon)) \setminus \mathbf{F}\left(\overline{B}\right)$ and let

$$\mathbf{G}(\mathbf{v}) \equiv \frac{r(\mathbf{a} - \mathbf{F}(\mathbf{v}))}{|\mathbf{a} - \mathbf{F}(\mathbf{v})|}.$$

If $|\mathbf{v}| = r$,

$$\mathbf{v} \cdot (\mathbf{a} - \mathbf{F}(\mathbf{v})) = \mathbf{v} \cdot \mathbf{a} - \mathbf{v} \cdot \mathbf{F}(\mathbf{v})$$

$$= \mathbf{v} \cdot \mathbf{a} - \mathbf{v} \cdot (\mathbf{F}(\mathbf{v}) - \mathbf{v}) - r^2$$

$$< r^2(1-\epsilon) + \epsilon r^2 - r^2 = 0.$$

Then for $|\mathbf{v}| = r$, $\mathbf{G}(\mathbf{v}) \neq \mathbf{v}$ because we just showed that $\mathbf{v} \cdot \mathbf{G}(\mathbf{v}) < 0$ but $\mathbf{v} \cdot \mathbf{v} = r^2 > 0$. If $|\mathbf{v}| < r$, it follows that $\mathbf{G}(\mathbf{v}) \neq \mathbf{v}$ because $|\mathbf{G}(\mathbf{v})| = r$ but $|\mathbf{v}| < r$. This lack of a fixed point contradicts the Brouwer fixed point theorem and this proves the lemma.

We are interested in generalizing the change of variables formula. Since $\mathbf{h}$ is only Lipschitz, $D\mathbf{h}(\mathbf{x})$ may not exist for all $\mathbf{x}$ but from the theorem of Rademacher $D\bar{\mathbf{h}}(\mathbf{x})$ exists a.e. $\mathbf{x}$.

In the arguments below, we will define a measure and use the Radon Nikodym theorem to obtain a function which is of interest to us. Then we will identify this function. In order to do this, we need some technical lemmas.

Lemma 20.16 *Let $\mathbf{x} \in \Omega$ be a point where $D\bar{\mathbf{h}}(\mathbf{x})^{-1}$ and $D\bar{\mathbf{h}}(\mathbf{x})$ exists. Then if $\epsilon \in (0,1)$ the following hold for all r small enough.*

$$m_n\left(\bar{\mathbf{h}}\left(\overline{B(\mathbf{x},r)}\right)\right) = m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right) \geq m_n\left(D\bar{\mathbf{h}}(\mathbf{x}) B(\mathbf{0}, r(1-\epsilon))\right), \tag{20.15}$$

$$\bar{\mathbf{h}}(B(\mathbf{x}, r)) \subseteq \bar{\mathbf{h}}(\mathbf{x}) + D\bar{\mathbf{h}}(\mathbf{x}) B(\mathbf{0}, r(1+\epsilon)), \tag{20.16}$$

$$m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right) \leq m_n\left(D\bar{\mathbf{h}}(\mathbf{x}) B(\mathbf{0}, r(1+\epsilon))\right) \tag{20.17}$$

If $\mathbf{x}$ is also a point of density of Ω, then

$$\lim_{r \to 0} \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x}, r) \cap \Omega)\right)}{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x}, r))\right)} = 1. \tag{20.18}$$

Proof: Since $D\bar{\mathbf{h}}(\mathbf{x})$ exists,

$$\begin{aligned} \bar{\mathbf{h}}(\mathbf{x} + \mathbf{v}) &= \bar{\mathbf{h}}(\mathbf{x}) + D\bar{\mathbf{h}}(\mathbf{x}) \mathbf{v} + o(|\mathbf{v}|) & (20.19) \\ &= \bar{\mathbf{h}}(\mathbf{x}) + D\bar{\mathbf{h}}(\mathbf{x}) \left(\mathbf{v} + D\bar{\mathbf{h}}(\mathbf{x})^{-1} o(|\mathbf{v}|)\right) & (20.20) \end{aligned}$$

Consequently, when r is small enough, (20.16) holds. Therefore, (20.17) holds. From (20.20),

$$\bar{\mathbf{h}}(\mathbf{x}+\mathbf{v}) = \bar{\mathbf{h}}(\mathbf{x}) + D\bar{\mathbf{h}}(\mathbf{x})(\mathbf{v}+o(|\mathbf{v}|)).$$

Thus, from the assumption that $D\bar{\mathbf{h}}(\mathbf{x})^{-1}$ exists,

$$D\bar{\mathbf{h}}(\mathbf{x})^{-1}\bar{\mathbf{h}}(\mathbf{x}+\mathbf{v}) - D\bar{\mathbf{h}}(\mathbf{x})^{-1}\bar{\mathbf{h}}(\mathbf{x}) - \mathbf{v} = o(|\mathbf{v}|). \tag{20.21}$$

Letting

$$\mathbf{F}(\mathbf{v}) = D\bar{\mathbf{h}}(\mathbf{x})^{-1}\bar{\mathbf{h}}(\mathbf{x}+\mathbf{v}) - D\bar{\mathbf{h}}(\mathbf{x})^{-1}\bar{\mathbf{h}}(\mathbf{x}),$$

we can apply Lemma 20.15 in (20.21) to conclude that for r small enough,

$$D\bar{\mathbf{h}}(\mathbf{x})^{-1}\bar{\mathbf{h}}(\mathbf{x}+\mathbf{v}) - D\bar{\mathbf{h}}(\mathbf{x})^{-1}\bar{\mathbf{h}}(\mathbf{x}) \supseteq B(\mathbf{0},(1-\epsilon)r).$$

Therefore,

$$\bar{\mathbf{h}}\left(\overline{B(\mathbf{x},r)}\right) \supseteq \bar{\mathbf{h}}(\mathbf{x}) + D\bar{\mathbf{h}}(\mathbf{x})B(\mathbf{0},(1-\epsilon)r)$$

which implies

$$m_n\left(\bar{\mathbf{h}}\left(\overline{B(\mathbf{x},r)}\right)\right) \geq m_n\left(D\bar{\mathbf{h}}(\mathbf{x})B(\mathbf{0},r(1-\epsilon))\right)$$

which shows (20.15).

Now suppose that $\mathbf{x}$ is also a point of density of Ω. Then whenever r is small enough,

$$m_n(B(\mathbf{x},r)\setminus\Omega) < \epsilon\alpha(n)r^n. \tag{20.22}$$

Then for such r we write

$$1 \geq \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r)\cap\Omega)\right)}{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right)}$$

$$\geq \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right) - m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r)\setminus\Omega)\right)}{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right)}.$$

From Lemma 20.14, and (20.15), this is no larger than

$$1 - \frac{\mathrm{Lip}\left(\bar{\mathbf{h}}\right)^n \epsilon\alpha(n)r^n}{m_n\left(D\bar{\mathbf{h}}(\mathbf{x})B(\mathbf{0},r(1-\epsilon))\right)}.$$

By the theorem on the change of variables for a linear map, this expression equals

$$1 - \frac{\mathrm{Lip}\left(\bar{\mathbf{h}}\right)^n \epsilon\alpha(n)r^n}{\left|\det\left(D\bar{\mathbf{h}}(\mathbf{x})\right)\right| r^n\alpha(n)(1-\epsilon)^n} \equiv 1 - g(\epsilon)$$

where $\lim_{\epsilon\to 0} g(\epsilon) = 0$. Then for all r small enough,

$$1 \geq \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r)\cap\Omega)\right)}{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right)} \geq 1 - g(\epsilon)$$

which proves the lemma since ϵ is arbitrary.

For simplicity in notation, we write $J(\mathbf{x})$ for the expression $\left|\det\left(D\bar{\mathbf{h}}(\mathbf{x})\right)\right|$.

Theorem 20.17 *Let $N \equiv \{\mathbf{x} \in \Omega : D\bar{\mathbf{h}}(\mathbf{x})$ does not exist $\}$. Then N has measure zero and if $\mathbf{x} \notin N$ then*

$$J(\mathbf{x}) = \lim_{r \to 0} \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x}, r))\right)}{m_n(B(\mathbf{x},r))}. \tag{20.23}$$

Proof: Suppose first that $D\bar{\mathbf{h}}(\mathbf{x})^{-1}$ exists. Using (20.15), (20.17) and the change of variables formula for linear maps,

$$\begin{aligned} J(\mathbf{x})(1-\epsilon)^n &= \frac{m_n\left(D\bar{\mathbf{h}}(\mathbf{x}) B(\mathbf{0},r(1-\epsilon))\right)}{m_n(B(\mathbf{x},r))} \leq \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right)}{m_n(B(\mathbf{x},r))} \\ &\leq \frac{m_n\left(D\bar{\mathbf{h}}(\mathbf{x}) B(\mathbf{0},r(1+\epsilon))\right)}{m_n(B(\mathbf{x},r))} = J(\mathbf{x})(1+\epsilon)^n \end{aligned}$$

whenever r is small enough. It follows that since $\epsilon > 0$ is arbitrary, (20.23) holds.

Now suppose $D\bar{\mathbf{h}}(\mathbf{x})^{-1}$ does not exist. Then from the definition of the derivative,

$$\bar{\mathbf{h}}(\mathbf{x}+\mathbf{v}) = \bar{\mathbf{h}}(\mathbf{x}) + D\bar{\mathbf{h}}(\mathbf{x})\mathbf{v} + o(\mathbf{v}),$$

and so for all r small enough, $\bar{\mathbf{h}}(B(\mathbf{x}, r))$ lies in a cylinder having height $r\varepsilon$ and diameter no more than

$$\left|\left|D\bar{\mathbf{h}}(\mathbf{x})\right|\right| 2r(1+\varepsilon).$$

Therefore, for such r,

$$\frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right)}{m_n(B(\mathbf{x},r))} \leq \frac{\left(\left|\left|D\bar{\mathbf{h}}(\mathbf{x})\right|\right| r(1+\varepsilon)\right)^{n-1} r\varepsilon}{r^n} \leq C\varepsilon$$

Since ε is arbitrary,

$$J(\mathbf{x}) = 0 = \lim_{r \to 0} \frac{m_n\left(\bar{\mathbf{h}}(B(\mathbf{x},r))\right)}{m_n(B(\mathbf{x},r))}.$$

This proves the theorem.

We define the following two sets for future reference

$$S \equiv \{\mathbf{x} \in \Omega : D\bar{\mathbf{h}}(\mathbf{x}) \text{ exists but } D\bar{\mathbf{h}}(\mathbf{x})^{-1} \text{ does not exist}\} \tag{20.24}$$

$$N \equiv \{\mathbf{x} \in \Omega : D\bar{\mathbf{h}}(\mathbf{x}) \text{ does not exist}\}, \tag{20.25}$$

and we assume for now that $\mathbf{h} : \Omega \to \mathbb{R}^n$ is one to one and Lipschitz. Since $\mathbf{h}$ is one to one, Lemma 20.14 implies we can define a measure, ν, on the $\sigma-$ algebra of Lebesgue measurable sets as follows.

$$\nu(E) \equiv m_n(\mathbf{h}(E \cap \Omega)).$$

By Lemma 20.14, we see this is a measure and $\nu << m_n$. Therefore by the corollary to the Radon Nikodym theorem, Corollary 18.3, there exists $f \in L^1_{loc}(\mathbb{R}^n)$, $f \geq 0$, $f(\mathbf{x}) = 0$ if $\mathbf{x} \notin \Omega$, and

$$\nu(E) = \int_E f dm = \int_{\Omega \cap E} f dm.$$

We want to identify f. Define

$$Q \equiv \{\mathbf{x} \in \Omega : \mathbf{x} \text{ is not a point of density of } \Omega\} \cup N \cup$$

$$\{\mathbf{x} \in \Omega : \mathbf{x} \text{ is not a Lebesgue point of } f\}.$$

Then E is a set of measure zero and if $\mathbf{x} \in (\Omega \setminus Q) \cap S^C$, Lemma 20.16 and Theorem 20.17 imply

$$\begin{aligned} f(\mathbf{x}) &= \lim_{r\to 0} \frac{1}{m_n(B(\mathbf{x},r))} \int_{B(\mathbf{x},r)} f(\mathbf{y})\,dm = \lim_{r\to 0} \frac{m_n\left(\overline{\mathbf{h}}(B(\mathbf{x},r)\cap\Omega)\right)}{m_n(B(\mathbf{x},r))} \\ &= \lim_{r\to 0} \frac{m_n\left(\overline{\mathbf{h}}(B(\mathbf{x},r)\cap\Omega)\right)}{m_n\left(\overline{\mathbf{h}}(B(\mathbf{x},r))\right)} \frac{m_n\left(\overline{\mathbf{h}}(B(\mathbf{x},r))\right)}{m_n(B(\mathbf{x},r))} = J(\mathbf{x}). \end{aligned}$$

On the other hand, if $\mathbf{x} \in (\Omega \setminus Q) \cap S$, then by Theorem 20.17,

$$\begin{aligned} f(\mathbf{x}) &= \lim_{r\to 0} \frac{1}{m_n(B(\mathbf{x},r))} \int_{B(\mathbf{x},r)} f(\mathbf{y})\,dm = \lim_{r\to 0} \frac{m_n\left(\overline{\mathbf{h}}(B(\mathbf{x},r)\cap\Omega)\right)}{m_n(B(\mathbf{x},r))} \\ &\leq \lim_{r\to 0} \frac{m_n\left(\overline{\mathbf{h}}(B(\mathbf{x},r))\right)}{m_n(B(\mathbf{x},r))} = J(\mathbf{x}) = 0. \end{aligned}$$

Therefore, $f(\mathbf{x}) = J(\mathbf{x})$ a.e., whenever $\mathbf{x} \in \Omega \setminus Q$.

Now let F be a Borel measurable set in $\mathbb{R}^n$. Recall this implies F is Lebesgue measurable. Then

$$\int_{\mathbf{h}(\Omega)} \mathcal{X}_F(\mathbf{y})\,dm_n = \int \mathcal{X}_{F\cap\mathbf{h}(\Omega)}(\mathbf{y})\,dm_n = m_n\left(\mathbf{h}\left(\mathbf{h}^{-1}(F)\cap\Omega\right)\right)$$

$$= \nu\left(\mathbf{h}^{-1}(F)\right) = \int \mathcal{X}_{\Omega\cap\mathbf{h}^{-1}(F)}(\mathbf{x})\,J(\mathbf{x})\,dm_n = \int_\Omega \mathcal{X}_F(\mathbf{h}(\mathbf{x}))\,J(\mathbf{x})\,dm_n. \tag{20.26}$$

Can we write a similar formula for F only m_n measurable? Note that there are no measurability questions in the above formula because $\mathbf{h}^{-1}(F)$ is a Borel set due to the continuity of $\mathbf{h}$ but it is not clear $\mathbf{h}^{-1}(F)$ is measurable for F only Lebesgue measurable.

First consider the case where E is only Lebesgue measurable but

$$m_n(E\cap\mathbf{h}(\Omega)) = 0.$$

By regularity of Lebesgue measure, there exists a Borel set $F \supseteq E\cap\mathbf{h}(\Omega)$ such that

$$m_n(F) = m_n(E\cap\mathbf{h}(\Omega)) = 0.$$

Then from (20.26),

$$\mathcal{X}_{\Omega\cap\mathbf{h}^{-1}(F)}(\mathbf{x})\,J(\mathbf{x}) = 0 \text{ a.e.}$$

But

$$0 \leq \mathcal{X}_{\Omega\cap\mathbf{h}^{-1}(E)}(\mathbf{x})\,J(\mathbf{x}) \leq \mathcal{X}_{\Omega\cap\mathbf{h}^{-1}(F)}(\mathbf{x})\,J(\mathbf{x}) \tag{20.27}$$

which shows the two functions in (20.27) are equal a.e. Therefore

$$\mathcal{X}_{\Omega\cap\mathbf{h}^{-1}(E)}(\mathbf{x})\,J(\mathbf{x})$$

is Lebesgue measurable and so from (20.26),

$$0 = \int \mathcal{X}_{E\cap\mathbf{h}(\Omega)}(\mathbf{y})\,dm_n = \int \mathcal{X}_{F\cap\mathbf{h}(\Omega)}(\mathbf{y})\,dm_n$$

$$= \int \mathcal{X}_{\Omega \cap \mathbf{h}^{-1}(F)}(\mathbf{x}) J(\mathbf{x})\, dm = \int \mathcal{X}_{\Omega \cap \mathbf{h}^{-1}(E)}(\mathbf{x}) J(\mathbf{x})\, dm, \tag{20.28}$$

which shows (20.26) holds in this case where

$$m_n(E \cap \mathbf{h}(\Omega)) = 0.$$

Now let $\Omega_R \equiv \Omega \cap B(\mathbf{0}, R)$ where R is large enough that $\Omega_R \neq \emptyset$ and let E be m_n measurable. By regularity of Lebesgue measure, there exists $F \supseteq E \cap \mathbf{h}(\Omega_R)$ such that F is Borel and

$$m_n(F \setminus (E \cap \mathbf{h}(\Omega_R))) = 0. \tag{20.29}$$

Now

$$(E \cap \mathbf{h}(\Omega_R)) \cup (F \setminus (E \cap \mathbf{h}(\Omega_R)) \cap \mathbf{h}(\Omega_R)) = F \cap \mathbf{h}(\Omega_R)$$

and so

$$\mathcal{X}_{\Omega_R \cap \mathbf{h}^{-1}(F)} J = \mathcal{X}_{\Omega_R \cap \mathbf{h}^{-1}(E)} J + \mathcal{X}_{\Omega_R \cap \mathbf{h}^{-1}(F \setminus (E \cap \mathbf{h}(\Omega_R)))} J$$

where from (20.29) and (20.28), the second function on the right of the equal sign is Lebesgue measurable and equals zero a.e. Therefore, by completeness of Lebesgue measure, the first function on the right of the equal sign is also Lebesgue measurable and equals the function on the left a.e. Thus,

$$\int \mathcal{X}_{E \cap \mathbf{h}(\Omega_R)}(\mathbf{y})\, dm_n = \int \mathcal{X}_{F \cap \mathbf{h}(\Omega_R)}(\mathbf{y})\, dm_n$$

$$= \int \mathcal{X}_{\Omega_R \cap \mathbf{h}^{-1}(F)}(\mathbf{x}) J(\mathbf{x})\, dm_n = \int \mathcal{X}_{\Omega_R \cap \mathbf{h}^{-1}(E)}(\mathbf{x}) J(\mathbf{x})\, dm_n. \tag{20.30}$$

Letting $R \to \infty$ we obtain (20.30) with Ω replacing Ω_R and the function

$$\mathbf{x} \to \mathcal{X}_{\Omega \cap \mathbf{h}^{-1}(E)}(\mathbf{x}) J(\mathbf{x})$$

is Lebesgue measurable. Writing this in a more familiar form yields

$$\int_{\mathbf{h}(\Omega)} \mathcal{X}_E(\mathbf{y})\, dm_n = \int_\Omega \mathcal{X}_E(\mathbf{h}(\mathbf{x})) J(\mathbf{x})\, dm. \tag{20.31}$$

From this, it follows that if s is a nonnegative m_n measurable simple function, (20.31) continues to be valid with s in place of $\mathcal{X}_E$. Then approximating an arbitrary nonnegative m_n measurable function, g, by an increasing sequence of simple functions, it follows that (20.31) holds with g in place of $\mathcal{X}_E$ and there are no measurability problems because $\mathbf{x} \to g(\mathbf{h}(\mathbf{x})) J(\mathbf{x})$ is Lebesgue measurable. This proves the following change of variables theorem.

Theorem 20.18 *Let $g : \mathbf{h}(\Omega) \to [0, \infty]$ be Lebesgue measurable where $\mathbf{h}$ is one to one and Lipschitz on Ω, and Ω is a Lebesgue measurable set. Then if $J(\mathbf{x})$ is defined to equal 0 for $\mathbf{x} \in N$,*

$$\mathbf{x} \to (g \circ \mathbf{h})(\mathbf{x}) J(\mathbf{x})$$

is Lebesgue measurable and

$$\int_{\mathbf{h}(\Omega)} g(\mathbf{y})\, dm_n = \int_\Omega g(\mathbf{h}(\mathbf{x})) J(\mathbf{x})\, dm_n.$$

For another version of this theorem based on the same arguments given here, in which the function $\mathbf{h}$ is not assumed to be Lipschitz, see Problems 23 - 26.

20.4 Mappings that are not one to one

In this section, $\mathbf{h}:\Omega\to\mathbb{R}^n$ will only be Lipschitz. We drop the requirement that $\mathbf{h}$ be one to one. Let S and N be given in (20.24) and (20.25). The following lemma is a version of Sard's theorem.

Lemma 20.19 *For S defined above, $m_n(\mathbf{h}(S)) = 0$.*

Proof: From Theorem 20.17, whenever $\mathbf{x} \in S$ and r is small enough,

$$\frac{m_n\left(\overline{\mathbf{h}}\left(B\left(\mathbf{x},r\right)\right)\right)}{m_n\left(B\left(\mathbf{x},r\right)\right)} < \epsilon.$$

Therefore, whenever $\mathbf{x} \in S$ and r small enough,

$$m_n\left(\overline{\mathbf{h}}\left(B\left(\mathbf{x},r\right)\right)\right) \leq \epsilon\alpha\left(n\right)r^n. \tag{20.32}$$

Let $S_k = S \cap B(\mathbf{0},k)$ and for each $\mathbf{x} \in S_k$, let $r_{\mathbf{x}}$ be such that (20.32) holds with r replaced by $5r_{\mathbf{x}}$ and

$$B\left(\mathbf{x},r_{\mathbf{x}}\right) \subseteq B\left(\mathbf{0},k\right).$$

By the Vitali covering theorem, there is a disjoint subsequence of these balls, $\{B(\mathbf{x}_i, r_i)\}$, with the property that $\{B(\mathbf{x}_i, 5r_i)\} \equiv \left\{\widehat{B}_i\right\}$ covers S_k. Then by the way these balls were defined, with (20.32) holding for $r = 5r_i$,

$$m_n\left(\overline{\mathbf{h}}\left(S_k\right)\right) \leq \sum_{i=1}^{\infty} m_n\left(\overline{\mathbf{h}}\left(\widehat{B}_i\right)\right) \leq 5^n\epsilon\sum_{i=1}^{\infty}\alpha\left(n\right)r_i^n$$

$$= 5^n\epsilon\sum_{i=1}^{\infty} m_n\left(B\left(\mathbf{x}_i,r_i\right)\right) \leq 5^n\epsilon m_n\left(B\left(\mathbf{0},k\right)\right).$$

Since ϵ is arbitrary, this shows $m_n\left(\overline{\mathbf{h}}\left(S_k\right)\right) = 0$. Now letting $k \to \infty$, this shows $m_n\left(\overline{\mathbf{h}}\left(S\right)\right) = 0$ which proves the lemma.

Thus $m_n(N) = 0$ and $m_n(\mathbf{h}(S)) = 0$ and so by Lemma 20.14

$$m_n\left(\mathbf{h}\left(S \cup N\right)\right) \leq m_n\left(\mathbf{h}\left(S\right)\right) + m_n\left(\mathbf{h}\left(N\right)\right) = 0. \tag{20.33}$$

Let $B \equiv \Omega \setminus (S \cup N)$.

A similar lemma to the following was proved in the section on the change of variables formula for a C^1 map. There the proof was based on the inverse function theorem. However, this is no longer possible; so, a slightly more technical argument is required.

Lemma 20.20 *There exists a sequence of disjoint measurable sets, $\{F_i\}$, such that*

$$\cup_{i=1}^{\infty} F_i = B$$

and $\mathbf{h}$ is one to one on F_i.

Proof: $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a finite dimensional normed linear space. In fact,

$$\{\mathbf{e}_i \otimes \mathbf{e}_j : i, j \in \{1, \cdots, n\}\}$$

is easily seen to be a basis. Let $\mathcal{I}$ be the elements of $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ which are invertible and let $\mathcal{F}$ be a countable dense subset of $\mathcal{I}$. Also let C be a countable dense subset of B. For $\mathbf{c} \in C$ and $T \in \mathcal{F}$,

$$E\left(\mathbf{c}, T, i\right) \equiv \left\{\mathbf{b} \in B\left(\mathbf{c}, i^{-1}\right) \cap B \text{ such that } (a.) \text{ and } (b.) \text{ hold}\right\}$$

where the conditions *(a.)* and *(b.)* are as follows.

$$\frac{1}{1+\epsilon}|T\mathbf{v}| \leq |D\mathbf{h}(\mathbf{b})\mathbf{v}| \text{ for all } \mathbf{v} \tag{a.}$$

$$|\mathbf{h}(\mathbf{a}) - \mathbf{h}(\mathbf{b}) - D\mathbf{h}(\mathbf{b})(\mathbf{a}-\mathbf{b})| \leq \epsilon|T(\mathbf{a}-\mathbf{b})| \tag{b.}$$

for all $\mathbf{a} \in B\left(\mathbf{b}, 2i^{-1}\right)$. Here $0 < \epsilon < 1/2$.

Obviously, there are countably many $E(\mathbf{c}, T, i)$. Now suppose $\mathbf{a}, \mathbf{b} \in E(\mathbf{c}, T, i)$ and $\mathbf{h}(\mathbf{a}) = \mathbf{h}(\mathbf{b})$. Then

$$|\mathbf{a}-\mathbf{b}| \leq |\mathbf{a}-\mathbf{c}| + |\mathbf{c}-\mathbf{b}| < \frac{2}{i}.$$

Therefore, from $(a.)$ and $(b.)$,

$$\begin{aligned} \frac{1}{1+\epsilon}|T(\mathbf{a}-\mathbf{b})| &\leq |D\mathbf{h}(\mathbf{b})(\mathbf{a}-\mathbf{b})| \\ &= |\mathbf{h}(\mathbf{a}) - \mathbf{h}(\mathbf{b}) - D\mathbf{h}(\mathbf{b})(\mathbf{a}-\mathbf{b})| \leq \epsilon|T(\mathbf{a}-\mathbf{b})|. \end{aligned}$$

Since T is one to one, this shows that $\mathbf{a} = \mathbf{b}$. Thus $\mathbf{h}$ is one to one on $E(\mathbf{c}, T, i)$.

Now let $\mathbf{b} \in B$. Choose $T \in \mathcal{F}$ such that

$$||D\mathbf{h}(\mathbf{b}) - T|| < \epsilon \left|\left|D\mathbf{h}(\mathbf{b})^{-1}\right|\right|^{-1}.$$

Then for all $\mathbf{v} \in \mathbb{R}^n$,

$$|T\mathbf{v} - \mathbf{Dh}(\mathbf{b})\mathbf{v}| \leq \epsilon \left|\left|D\mathbf{h}(\mathbf{b})^{-1}\right|\right|^{-1}|\mathbf{v}| \leq \epsilon|D\mathbf{h}(\mathbf{b})\mathbf{v}|$$

and so

$$|T\mathbf{v}| \leq (1+\epsilon)|D\mathbf{h}(\mathbf{b})\mathbf{v}|$$

which yields $(a.)$. Now choose i large enough that for $|\mathbf{a}-\mathbf{b}| < 2i^{-1}$,

$$\begin{aligned} |\mathbf{h}(\mathbf{a}) - \mathbf{h}(\mathbf{b}) - D\mathbf{h}(\mathbf{b})(\mathbf{a}-\mathbf{b})| &< \frac{\epsilon}{||T^{-1}||}|\mathbf{a}-\mathbf{b}| \\ &\leq \epsilon|T(\mathbf{a}-\mathbf{b})| \end{aligned}$$

and pick $\mathbf{c} \in C \cap B\left(\mathbf{b}, i^{-1}\right)$. Then $\mathbf{b} \in E(\mathbf{c}, T, i)$ and this shows that B equals the union of these sets.

Let $\{E_i\}$ be an enumeration of these sets and define $F_1 \equiv E_1$, and if $F_1, \cdots, F_n$ have been chosen, $F_{n+1} \equiv E_{n+1} \setminus \cup_{i=1}^n F_i$. Then $\{F_i\}$ satisfies the conditions of the lemma and this proves the lemma.

The following corollary is also of interest.

Corollary 20.21 *For each E_i in Lemma 20.20, $\mathbf{h}^{-1}$ is Lipschitz on $\mathbf{h}(E_i)$.*

Proof: Pick $\mathbf{a}, \mathbf{b} \in E_i$. Then by condition $a.$ and $b.$,

$$|\mathbf{h}(\mathbf{a}) - \mathbf{h}(\mathbf{b})| \geq |D\mathbf{h}(\mathbf{b})(\mathbf{a}-\mathbf{b})| - \epsilon|T(\mathbf{a}-\mathbf{b})|$$

$$\geq \left(\frac{1}{1+\epsilon} - \epsilon\right)|T(\mathbf{a}-\mathbf{b})| \geq r|\mathbf{a}-\mathbf{b}|$$

for some $r > 0$ by the equivalence of all norms on a finite dimensional space. Therefore,

$$\left|\mathbf{h}^{-1}(\mathbf{h}(\mathbf{a})) - \mathbf{h}^{-1}(\mathbf{h}(\mathbf{b}))\right| \leq \frac{1}{r}\left|\mathbf{h}(\mathbf{a}) - \mathbf{h}(\mathbf{b})\right|$$

and this proves the corollary.

Now let $g : \mathbf{h}(\Omega) \to [0, \infty]$ be m_n measurable. By Theorem 20.18,

$$\int_{\mathbf{h}(\Omega)} \mathcal{X}_{\mathbf{h}(F_i)}(\mathbf{y})\, g(\mathbf{y})\, dm_n = \int_{F_i} g(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm. \tag{20.34}$$

Now define

$$\mathfrak{n}(\mathbf{y}) = \sum_{i=1}^{\infty} \mathcal{X}_{\mathbf{h}(F_i)}(\mathbf{y}).$$

By Lemma 20.14, $\mathbf{h}(F_i)$ is m_n measurable and so $\mathfrak{n}$ is a m_n measurable function. For each $\mathbf{y} \in B$, $\mathfrak{n}(\mathbf{y})$ gives the number of elements in $\mathbf{h}^{-1}(\mathbf{y}) \cap B$. From (20.34),

$$\int_{\mathbf{h}(\Omega)} \mathfrak{n}(\mathbf{y})\, g(\mathbf{y})\, dm_n = \int_{B} g(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm. \tag{20.35}$$

Now define

$$\#(\mathbf{y}) \equiv \text{ number of elements in } \mathbf{h}^{-1}(\mathbf{y}).$$

Theorem 20.22 *The function* $\mathbf{y} \to \#(\mathbf{y})$ *is* m_n *measurable and if*

$$g : \mathbf{h}(\Omega) \to [0, \infty]$$

is m_n *measurable, then*

$$\int_{\mathbf{h}(\Omega)} g(\mathbf{y})\, \#(\mathbf{y})\, dm_n = \int_{\Omega} g(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm.$$

Proof: If $\mathbf{y} \notin \mathbf{h}(S \cup N)$, then $\mathfrak{n}(\mathbf{y}) = \#(\mathbf{y})$. By (20.33)

$$m_n(\mathbf{h}(S \cup N)) = 0$$

and so $\mathfrak{n}(\mathbf{y}) = \#(\mathbf{y})$ a.e. Since m_n is a complete measure, $\#(\cdot)$ is m_n measurable. Letting

$$G \equiv \mathbf{h}(\Omega) \setminus \mathbf{h}(S \cup N),$$

(20.35) implies

$$\begin{aligned}
\int_{\mathbf{h}(\Omega)} g(\mathbf{y})\, \#(\mathbf{y})\, dm_n &= \int_{G} g(\mathbf{y})\, \mathfrak{n}(\mathbf{y})\, dm_n \\
&= \int_{B} g(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm \\
&= \int_{\Omega} g(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm.
\end{aligned}$$

This proves the theorem.

This theorem is a special case of the area formula proved in [19] and [11].

20.5 Differential forms on Lipschitz manifolds

With the change of variables formula for Lipschitz maps, we can generalize much of what was done for C^k manifolds to Lipschitz manifolds. To begin with we define what these are and then we will discuss the integration of differential forms on Lipschitz manifolds.

In order to show the integration of differential forms is well defined, we need a version of the chain rule valid in the context of Lipschitz maps.

Theorem 20.23 *Let* $\mathbf{f}$ *,*$\mathbf{g}$ *and* $\mathbf{h}$ *be Lipschitz mappings from* $\mathbb{R}^n$ *to* $\mathbb{R}^n$ *with* $\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{h}(\mathbf{x})$ *on* A*, a measurable set. Also suppose that* $\mathbf{f}^{-1}$ *is Lipschitz on* $\mathbf{f}(A)$*. Then for a.e.* $\mathbf{x} \in A$*,* $D\mathbf{g}(\mathbf{f}(\mathbf{x}))$*,* $D\mathbf{f}(\mathbf{x})$*,* $D\mathbf{h}(\mathbf{x})$ *,and* $D(\mathbf{f} \circ \mathbf{g})(\mathbf{x})$ *all exist and*

$$D\mathbf{h}(\mathbf{x}) = D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x}))\, D\mathbf{f}(\mathbf{x}).$$

The proof of this theorem is based on the following lemma.

Lemma 20.24 *Let* $\mathbf{k} : \mathbb{R}^n \to \mathbb{R}^n$ *be Lipschitz, then if* $\mathbf{k}(\mathbf{x}) = \mathbf{0}$ *for all* $\mathbf{x} \in A$*, then* $\det(D\mathbf{k}(\mathbf{x})) = 0$ *a.e.*$\mathbf{x} \in A$*.*

Proof: By the change of variables formula, Theorem 20.22, $0 = \int_{\{\mathbf{0}\}} \#(\mathbf{y})\, dy = \int_A |\det(D\mathbf{k}(\mathbf{x}))|\, dx$ and so $\det(D\mathbf{k}(\mathbf{x})) = 0$ a.e.

Proof of the theorem: On A, $\mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{h}(\mathbf{x}) = \mathbf{0}$ and so by the lemma, there exists a set of measure zero, N_1 such that if $\mathbf{x} \notin N_1$, $D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) - D\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Let M be the set of points in $\mathbf{f}(A)$ where $\mathbf{g}$ fails to be differentiable and let $N_2 \equiv \mathbf{f}^{-1}(M) \cap A$, also a set of measure zero because by Lemma 20.13, Lipschitz functions map sets of measure zero to sets of measure zero. Finally let N_3 be the set of points where $\mathbf{f}$ fails to be differentiable. Then if $\mathbf{x} \in A \setminus (N_1 \cup N_2 \cup N_3)$, the chain rule implies $D\mathbf{h}(\mathbf{x}) = D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x}))\, D\mathbf{f}(\mathbf{x})$. This proves the theorem.

Definition 20.25 *We say* Ω *is a Lipschitz manifold with boundary if it is a manifold with boundary for which the maps,* $\mathbf{R}_i$ *and* $\mathbf{R}_i^{-1}$ *are Lipschitz. It will be called orientable if there is an atlas,* $(U_i, \mathbf{R}_i)$ *for which* $\det\left(D\left(\mathbf{R}_i \circ \mathbf{R}_j^{-1}\right)(\mathbf{u})\right) > 0$ *for a.e.* $\mathbf{u} \in \mathbf{R}_j(U_i \cap U_j)$*. We will call such an atlas a Lipschitz atlas.*

Note we had to say the determinant of the derivative of the overlap map is greater than zero a.e. because all we know is that this map is Lipschitz and so it only has a derivative a.e.

Lemma 20.26 *Suppose* $(V, \mathbf{S})$ *and* $(U, \mathbf{R})$ *are two charts for a Lipschitz n manifold,* $\Omega \subseteq \mathbb{R}^m$*, that* $\det(D\mathbf{u}(\mathbf{v}))$ *exists for a.e.* $\mathbf{v} \in \mathbf{S}(V \cap U)$*, and* $\det(D\mathbf{v}(\mathbf{u}))$ *exists for a.e.* $\mathbf{u} \in \mathbf{R}(U \cap V)$*. Here we are writing* $\mathbf{u}(\mathbf{v})$ *for* $\mathbf{R} \circ \mathbf{S}^{-1}(\mathbf{v})$ *with a similar definition for* $\mathbf{v}(\mathbf{u})$*. Then letting* $I = (i_1, \cdots, i_n)$ *be a sequence of indices, we have the following formula for a.e.* $\mathbf{u} \in \mathbf{R}(U \cap V)$*.*

$$\frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(u^1 \cdots u^n\right)} = \frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(v^1 \cdots v^n\right)} \frac{\partial\left(v^1 \cdots v^n\right)}{\partial\left(u^1 \cdots u^n\right)}. \tag{20.36}$$

Proof: Let $\mathbf{x}^I \in \mathbb{R}^n$ denote $\left(x^{i_1}, \cdots, x^{i_n}\right)$. Then using Theorem 20.23, we can say that there is a set of Lebesgue measure zero, $N \subseteq \mathbf{R}(U \cap V)$, such that if $\mathbf{u} \notin N$, then

$$D\mathbf{x}^I(\mathbf{u}) = D\mathbf{x}^I(\mathbf{v})\, D\mathbf{v}(\mathbf{u}).$$

Taking determinants, we obtain (20.36) for $\mathbf{u} \notin N$. This proves the lemma.

Similarly we have the following formula which holds for a.e. $\mathbf{v} \in \mathbf{S}(U \cap V)$.

$$\frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(v^1 \cdots v^n\right)} = \frac{\partial\left(x^{i_1} \cdots x^{i_n}\right)}{\partial\left(u^1 \cdots u^n\right)} \frac{\partial\left(u^1 \cdots u^n\right)}{\partial\left(v^1 \cdots v^n\right)}.$$

20.6 Some examples of orientable Lipschitz manifolds

The following simple proposition will give many examples of Lipschitz manifolds.

Proposition 20.27 *Suppose Ω is a bounded open subset of $\mathbb{R}^n$ with $n \geq 2$ having the property that for all $\mathbf{p} \in \partial\Omega \equiv \overline{\Omega} \setminus \Omega$, there exists an open set, U_1, containing $\mathbf{p}$, and an orthogonal transformation, $P : \mathbb{R}^n \to \mathbb{R}^n$ whose determinant equals 1, having the following properties. The set, $P(U_1)$ is of the form*

$$B \times (a, b)$$

where B is a bounded open subset of $\mathbb{R}^{n-1}$. Letting

$$U \equiv \widetilde{U} \cap \Omega,$$

it follows that

$$P(U) = \left\{\mathbf{u} \in \mathbb{R}^n : \widehat{\mathbf{u}} \in B \text{ and } u^1 \in (a, g(\widehat{\mathbf{u}}))\right\} \tag{20.37}$$

while

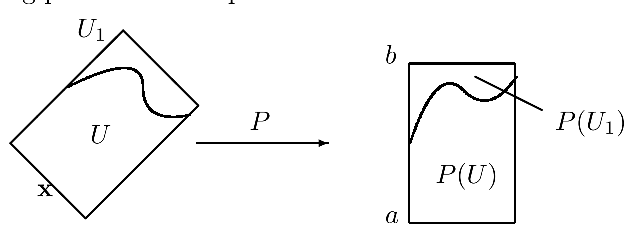
$$P(\partial\Omega \cap U_1) = \left\{\mathbf{u} \in \mathbb{R}^n : \widehat{\mathbf{u}} \in B \text{ and } u^1 = g(\widehat{\mathbf{u}})\right\} \tag{20.38}$$

where g is a Lipschitz map and

$$\widehat{\mathbf{u}} \equiv \left(u^2, \cdots, u^n\right)^T \text{ where } \mathbf{u} = \left(u^1, u^2 \cdots, u^n\right)^T.$$

Then Ω is an oriented Lipschitz manifold.

The following picture is descriptive of the situation described in the proposition.



Proof: Let U_1, U, and g be as described above. For $\mathbf{x} \in U$ define

$$\mathbf{R}(\mathbf{x}) \equiv \alpha \circ P(\mathbf{x})$$

where

$$\alpha(\mathbf{u}) \equiv \begin{pmatrix} u^1 - g(\widehat{\mathbf{u}}) & u^2 & \cdots & u^n \end{pmatrix}^T$$

The pair, $(U, \mathbf{R})$ is a chart in an atlas for Ω. It is clear that $\mathbf{R}$ is Lipschitz continuous. Furthermore, it is clear that the inverse map, $\mathbf{R}^{-1}$ is given by the formula

$$\mathbf{R}^{-1}(\mathbf{u}) = P^* \circ \alpha^{-1}(\mathbf{u})$$

where

$$\alpha^{-1}(\mathbf{u}) = \begin{pmatrix} u^1 + g(\widehat{\mathbf{u}}) & u^2 & \cdots & u^n \end{pmatrix}^T$$

and that $\mathbf{R}$, $\mathbf{R}^{-1}$, α and α^{-1} are all are Lipschitz continuous. Then the version of the chain rule found in Theorem 20.23 implies $D\mathbf{R}(\mathbf{x})$, and $D\mathbf{R}^{-1}(\mathbf{x})$ exist and have postive determinant a.e. Thus if $(U,\mathbf{R})$ and $(V,\mathbf{S})$ are two of these charts, we may apply Theorem 20.23 to find that for a.e. $\mathbf{v}$,

$$\begin{aligned}\det\left(D\left(\mathbf{R}\circ\mathbf{S}^{-1}\right)(\mathbf{v})\right) &= \det\left(D\mathbf{R}\left(\mathbf{S}^{-1}(\mathbf{v})\right)D\mathbf{S}^{-1}(\mathbf{v})\right)\\ &= \det\left(D\mathbf{R}\left(\mathbf{S}^{-1}(\mathbf{v})\right)\right)\det\left(D\mathbf{S}^{-1}(\mathbf{v})\right)>0.\end{aligned}$$

The set, $\partial\Omega$ is compact and so there are p of these sets, U_{1j} covering $\partial\Omega$ along with functions $\mathbf{R}_j$ as just described. Let U_0 satisfy

$$\Omega\setminus\cup_{j=1}^{p}U_{1j}\subseteq U_0\subseteq\overline{U_0}\subseteq\Omega$$

and let $\mathbf{R}_0(\mathbf{x})\equiv\left(\begin{array}{cccc}x^1-k & x^2 & \cdots & x^n\end{array}\right)$ where k is chosen large enough that $\mathbf{R}_0$ maps U_0 into $\mathbb{R}^n_<$. Then $(U_r,\mathbf{R}_r)$ is an oriented atlas for Ω if we define $U_r\equiv U_{1r}\cap\Omega$. As above, the chain rule shows the derivatives of the overlap maps have positive determinants.

For example, a ball of radius $r>0$ is an oriented Lipschitz n manifold with boundary because it satisfies the conditions of the above proposition. So is a box, $\prod_{i=1}^{n}(a_i,b_i)$. This proposition gives examples of Lipschitz n manifolds in $\mathbb{R}^n$ but we want to have examples of n manifolds in $\mathbb{R}^m$ for $m>n$. We recall the following lemma from Section 17.3.

Lemma 20.28 *Suppose O is a bounded open subset of $\mathbb{R}^n$ and let $\mathbf{F}:O\to\mathbb{R}^m$ be a function in $C^k\left(\overline{O};\mathbb{R}^m\right)$ where $m\geq n$ with the property that for all $\mathbf{x}\in O$, $D\mathbf{F}(\mathbf{x})$ has rank n. Then if $\mathbf{y}_0=\mathbf{F}(\mathbf{x}_0)$, there exists a bounded open set in $\mathbb{R}^m, W$, which contains $\mathbf{y}_0$, a bounded open set, $U\subseteq O$ which contains $\mathbf{x}_0$ and a function $\mathbf{G}:W\to U$ such that $\mathbf{G}$ is in $C^k\left(\overline{W};\mathbb{R}^n\right)$ and for all $\mathbf{x}\in U$,*

$$\mathbf{G}(\mathbf{F}(\mathbf{x}))=\mathbf{x}.$$

Furthermore, $\mathbf{G}=\mathbf{G}_1\circ\mathbf{P}$ on W where $\mathbf{P}$ is a map of the form

$$\mathbf{P}(\mathbf{y})=\left(y^{i_1},\cdots,y^{i_n}\right)$$

for some list of indices, $i_1<\cdots<i_n$.

With this lemma we can give a theorem which will provide many other examples. We note that $\mathbf{G}$, and $\mathbf{F}$ are Lipschitz continuous in the above because of the requirement that they are restrictions of C^k functions defined on $\mathbb{R}^n$ which have compact support.

Theorem 20.29 *Let Ω be a Lipschitz n manifold with boundary in $\mathbb{R}^n$ and suppose $\Omega\subseteq O$, an open bounded subset of $\mathbb{R}^n$. Suppose $\mathbf{F}\in C^1\left(\overline{O};\mathbb{R}^m\right)$ is one to one on O and $D\mathbf{F}(\mathbf{x})$ has rank n for all $\mathbf{x}\in O$. Then $\mathbf{F}(\Omega)$ is also a Lipschitz manifold with boundary and $\partial\mathbf{F}(\Omega)=\mathbf{F}(\partial\Omega)$.*

Proof: Let $(U_r,\mathbf{R}_r)$ be an atlas for Ω and suppose $U_r=O_r\cap\Omega$ where O_r is an open subset of O. Let $\mathbf{x}_0\in U_r$. By Lemma 20.28 there exists an open set, $W_{\mathbf{x}_0}$ in $\mathbb{R}^m$ containing $\mathbf{F}(\mathbf{x}_0)$, an open set in $\mathbb{R}^n,\widetilde{U_{\mathbf{x}_0}}$ containing $\mathbf{x}_0$, and $\mathbf{G}_{\mathbf{x}_0}\in C^k\left(\overline{W_{\mathbf{x}_0}};\mathbb{R}^n\right)$ such that

$$\mathbf{G}_{\mathbf{x}_0}(\mathbf{F}(\mathbf{x}))=\mathbf{x}$$

for all $\mathbf{x}\in\widetilde{U_{\mathbf{x}_0}}$. Let $U_{\mathbf{x}_0}\equiv U_r\cap\widetilde{U_{\mathbf{x}_0}}$.

Claim: $\mathbf{F}(U_{\mathbf{x}_0})$ is open in $\mathbf{F}(\Omega)$.

Proof: Let $\mathbf{x}\in U_{\mathbf{x}_0}$. If $\mathbf{F}(\mathbf{x}_1)$ is close enough to $\mathbf{F}(\mathbf{x})$ where $\mathbf{x}_1\in\Omega$, then $\mathbf{F}(\mathbf{x}_1)\in W_{\mathbf{x}_0}$ and so

$$\begin{aligned}|\mathbf{x}-\mathbf{x}_1| &= |\mathbf{G}_{\mathbf{x}_0}(\mathbf{F}(\mathbf{x}))-\mathbf{G}_{\mathbf{x}_0}(\mathbf{F}(\mathbf{x}_1))|\\ &\leq K\,|(\mathbf{F}(\mathbf{x}))-\mathbf{F}(\mathbf{x}_1)|\end{aligned}$$

where K is some constant which depends only on

$$\max\left\{||D\mathbf{G}_{\mathbf{x}_0}(\mathbf{y})|| : \mathbf{y} \in\mathbb{R}^m\right\}.$$

Therefore, if $\mathbf{F}(\mathbf{x}_1)$ is close enough to $\mathbf{F}(\mathbf{x})$, it follows we can conclude $|\mathbf{x}-\mathbf{x}_1|$ is very small. Since $U_{\mathbf{x}_0}$ is open in Ω it follows that whenever $\mathbf{F}(\mathbf{x}_1)$ is sufficiently close to $\mathbf{F}(\mathbf{x})$, we have $\mathbf{x}_1 \in U_{\mathbf{x}_0}$. Consequently $\mathbf{F}(\mathbf{x}_1) \in \mathbf{F}(U_{\mathbf{x}_0})$. This shows $\mathbf{F}(U_{\mathbf{x}_0})$ is open in $\mathbf{F}(\Omega)$ and proves the claim.

With this claim it follows that $(\mathbf{F}(U_{\mathbf{x}_0}), \mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_0})$ is a chart and since $\mathbf{R}_r$ is given to be Lipschitz continuous, we know the map, $\mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_0}$ is Lipschitz continuous. The inverse map of $\mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_0}$ is also seen to equal $\mathbf{F} \circ \mathbf{R}_r^{-1}$, also Lipschitz continuous. Since Ω is compact there are finitely many of these sets, $\mathbf{F}(U_{\mathbf{x}_i})$ covering $\mathbf{F}(\Omega)$. This yields an atlas for $\mathbf{F}(\Omega)$ of the form $(\mathbf{F}(U_{\mathbf{x}_i}), \mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_i})$ where $\mathbf{x}_i \in U_r$ and proves $\mathbf{F}(\Omega)$ is a Lipschitz manifold.

Since the $\mathbf{R}_r^{-1}$ are Lipschitz, the overlap map for two of these charts is of the form,

$$\left(\mathbf{R}_s \circ \mathbf{G}_{\mathbf{x}_j}\right) \circ \left(\mathbf{F} \circ \mathbf{R}_r^{-1}\right) = \mathbf{R}_s \circ \mathbf{R}_r^{-1},$$

showing that if $(U_r, \mathbf{R}_r)$ is an oriented atlas for Ω, then $\mathbf{F}(\Omega)$ also has an oriented atlas since the overlap maps described above are all of the form $\mathbf{R}_s \circ \mathbf{R}_r^{-1}$.

It remains to verify the assertion about boundaries. $\mathbf{y} \in \partial\mathbf{F}(\Omega)$ if and only if for some $\mathbf{x}_i \in U_r$,

$$\mathbf{R}_r \circ \mathbf{G}_{\mathbf{x}_i}(\mathbf{y}) \in \mathbb{R}_0^n$$

if and only if

$$\mathbf{G}_{\mathbf{x}_i}(\mathbf{y}) \in \partial\Omega$$

if and only if

$$\mathbf{G}_{\mathbf{x}_i}(\mathbf{F}(\mathbf{x})) = \mathbf{x} \in \partial\Omega$$

where $\mathbf{F}(\mathbf{x}) = \mathbf{y}$ if and only if $\mathbf{y} \in \mathbf{F}(\partial\Omega)$. This proves the theorem.

20.7 Stoke's theorem on Lipschitz manifolds

In the proof of Stoke's theorem, we used that $\mathbf{R}_i^{-1}$ is C^2 but the end result is a formula which involves only the first derivatives of $\mathbf{R}_i^{-1}$. This suggests that it is not necessary to make this assumption. Here we will show that one can do with assuming the manifold in question is only a Lipschitz manifold.

Using the Theorem 20.22 and the version of the chain rule for Lipschitz maps given above, we can verify that the integral of a differential form makes sense for an orientable Lipschitz manifold using the same arguments given earlier in the chapter on differential forms. Now suppose Ω is only a Lipschitz orientable n manifold. Then in the proof of Stoke's theorem, we can say that for some sequence of $n \to \infty$,

$$\int_\Omega d\omega \equiv \lim_{n\to\infty} \sum_{j=1}^m \sum_I \sum_{r=1}^p \int_{\mathbf{R}_r U_r} \left(\psi_r \frac{\partial a_I}{\partial x^j} \circ \left(\mathbf{R}_r^{-1} * \phi_N\right)\right)(\mathbf{u}) \frac{\partial\left(x^j x^{i_1} \cdots x^{i_n-1}\right)}{\partial\left(u^1 \cdots u^n\right)} du \tag{20.39}$$

where ϕ_N is a mollifier and

$$\frac{\partial\left(x^j x^{i_1} \cdots x^{i_n-1}\right)}{\partial\left(u^1 \cdots u^n\right)}$$

is obtained from

$$\mathbf{x} = \mathbf{R}_r^{-1} * \phi_N(\mathbf{u}).$$

The reason we can write the above limit is that from Rademacher's theorem and the dominated convergence theorem we may write

$$\left(\mathbf{R}_r^{-1} * \phi_N\right)_{,i} = \left(\mathbf{R}_r^{-1}\right)_{,i} * \phi_N.$$

As in Chapter 12 we see $\left(\mathbf{R}_r^{-1} * \phi_N\right)_{,i} \to \left(\mathbf{R}_r^{-1}\right)_{,i}$ in $L^p\left(\mathbf{R}_r U_r\right)$ for every p. Taking an appropriate subsequence, we can obtain, in addition to this, almost everywhere convergence for every partial derivative and every $\mathbf{R}_r$ yielding (20.39) for that subsequence.

We may also arrange to have $\sum \psi_r = 1$ near Ω. Then for N large enough, $\mathbf{R}_r^{-1} * \phi_N\left(\mathbf{R}_r U_i\right)$ will lie in this set where $\sum \psi_r = 1$. Then we do the computations as in the proof of Stokes theorem. Using the same computations, with $\mathbf{R}_r^{-1} * \phi_N$ in place of $\mathbf{R}_r^{-1}$,

$$\int_\Omega d\omega =$$

$$\lim_{n\to\infty} \sum_{j=1}^m \sum_I \sum_{r\in B} \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \left(\psi_r a_I \circ \left(\mathbf{R}_r^{-1} * \phi_N\right)\right) \frac{\partial\left(x^{i_1}\cdots x^{i_{n-1}}\right)}{\partial\left(u^2\cdots u^n\right)}\left(0,u^2,\cdots,u^n\right) du_2\cdots du_n$$

$$= \sum_{j=1}^m \sum_I \sum_{r\in B} \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \left(\psi_r a_I \circ \mathbf{R}_r^{-1}\right) \frac{\partial\left(x^{i_1}\cdots x^{i_{n-1}}\right)}{\partial\left(u^2\cdots u^n\right)}\left(0,u^2,\cdots,u^n\right) du_2\cdots du_n \equiv \int_{\partial\Omega}\omega.$$

This yields the following significant generalization of Stoke's theorem.

Theorem 20.30 *(Stokes theorem) Let Ω be a Lipschitz orientable manifold with boundary and let $\omega \equiv \sum_I a_I(\mathbf{x})\, dx^{i_1}\wedge\cdots\wedge dx^{i_{n-1}}$ be a differential form of order $n-1$ for which a_I is C^1. Then*

$$\int_{\partial\Omega}\omega = \int_\Omega d\omega. \tag{20.40}$$

The theorem can be generalized further but this version seems particularly natural so we stop with this one. You need a change of variables formula and you need to be able to take the derivative of $\mathbf{R}_i^{-1}$ a.e. These ingredients are available for more general classes of functions than Lipschitz continuous functions. See [19] in the exercises on the area formula. You could also relax the assumptions on a_I slightly.

20.8 Surface measures on Lipschitz manifolds

Let Ω be a Lipschitz manifold in $\mathbb{R}^m$, oriented or not. Let f be a continuous function defined on Ω, and let $(U_i, \mathbf{R}_i)$ be an atlas and let $\{\psi_i\}$ be a C^∞ partition of unity subordinate to the sets, U_i as described earlier. If $\omega = \sum_I a_I(\mathbf{x})\, d\mathbf{x}^I$ is a differential form, we may always assume

$$d\mathbf{x}^I = dx^{i_1}\wedge\cdots\wedge dx^{i_n}$$

where $i_1 < i_2 < \cdots < i_n$. The reason for this is that in taking an integral used to integrate the differential form, a switch in two of the dx^j results in switching two rows in the determinant, $\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)}$, implying that any two of these differ only by a multiple of -1. Therefore, there is no loss of generality in assuming from now on that in the sum for ω, I is always a list of indices which are strictly increasing. The case where some term of ω has a repeat, $dx^{i_r} = dx^{i_s}$ can be ignored because such terms deliver zero in integrating the differential form because they involve a determinant having two equal rows. We emphasize again that from now on I will refer to an increasing list of indices.

Let

$$J_i(\mathbf{u}) \equiv \left[\sum_I \left(\frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)}\right)^2\right]^{1/2}$$

where here the sum is taken over all possible increasing lists of indices, I, from $\{1,\cdots,m\}$ and $\mathbf{x}=\mathbf{R}_i^{-1}\mathbf{u}$. Thus there are $\binom{m}{n}$ terms in the sum. We define a positive linear functional, Λ on $C(\Omega)$ as follows:

$$\Lambda f \equiv \sum_{i=1}^p \int_{\mathbf{R}_i U_i} f\psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du. \tag{20.41}$$

We will now show this is well defined.

Lemma 20.31 *The functional defined in (20.41) does not depend on the choice of atlas or the partition of unity.*

Proof: In (20.41), let $\{\psi_i\}$ be a partition of unity as described there which is associated with the atlas $(U_i,\mathbf{R}_i)$ and let $\{\eta_i\}$ be a partition of unity associated in the same manner with the atlas $(V_i,\mathbf{S}_i)$. Using the change of variables formula, Theorem 20.22 with $\mathbf{u}=\left(\mathbf{R}_i\circ\mathbf{S}_j^{-1}\right)\mathbf{v}$,

$$\sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du = \tag{20.42}$$

$$\sum_{i=1}^p\sum_{j=1}^q \int_{\mathbf{R}_i U_i} \eta_j\psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du =$$

$$\sum_{i=1}^p\sum_{j=1}^q \int_{\mathbf{S}_j(U_i\cap V_j)} \eta_j\left(\mathbf{S}_j^{-1}(\mathbf{v})\right)\psi_i\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) f\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) J_i(\mathbf{u})\left|\frac{\partial\left(u^1\cdots u^n\right)}{\partial\left(v^1\cdots v^n\right)}\right| dv$$

$$=\sum_{i=1}^p\sum_{j=1}^q \int_{\mathbf{S}_j(U_i\cap V_j)} \eta_j\left(\mathbf{S}_j^{-1}(\mathbf{v})\right)\psi_i\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) f\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) J_j(\mathbf{v})\, dv. \tag{20.43}$$

This yields

$$\text{the definition of } \Lambda f \text{ using } (U_i,\mathbf{R}_i) \equiv$$

$$\sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du =$$

$$\sum_{i=1}^p\sum_{j=1}^q \int_{\mathbf{S}_j(U_i\cap V_j)} \eta_j\left(\mathbf{S}_j^{-1}(\mathbf{v})\right)\psi_i\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) f\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) J_j(\mathbf{v})\, dv$$

$$=\sum_{j=1}^q \int_{\mathbf{S}_j(V_j)} \eta_j\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) f\left(\mathbf{S}_j^{-1}(\mathbf{v})\right) J_j(\mathbf{v})\, dv$$

$$\text{the definition of } \Lambda f \text{ using } (V_i,\mathbf{S}_i).$$

This proves the lemma.

This lemma implies the following theorem.

Theorem 20.32 *Let Ω be a Lipschitz manifold with boundary. Then there exists a unique Radon measure, μ, defined on Ω such that whenever f is a continuous function defined on Ω and $(U_i, \mathbf{R}_i)$ denotes an atlas and $\{\psi_i\}$ a partition of unity subordinate to this atlas, we have*

$$\Lambda f = \int_{\Omega} f d\mu = \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du. \tag{20.44}$$

Furthermore, for any $f \in L^1(\Omega, \mu)$,

$$\int_{\Omega} f d\mu = \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \tag{20.45}$$

and a subset, A, of Ω is μ measurable if and only if for all $r, \mathbf{R}_r(U_r \cap A)$ is $J_r(\mathbf{u})\, du$ measurable.

Proof: We begin by proving the following claim.

Claim :A set, $S \subseteq U_i$, has μ measure zero in U_i, if and only if $\mathbf{R}_i S$ has measure zero in $\mathbf{R}_i U_i$ with respect to the measure, $J_i(\mathbf{u})\, du$.

Proof of the claim: Let $\varepsilon > 0$ be given. By outer regularity, there exists a set, $V \subseteq U_i$, open in Ω such that $\mu(V) < \varepsilon$ and $S \subseteq V \subseteq U_i$. Then $\mathbf{R}_i V$ is open in $\mathbb{R}^n_{\leq}$ and contains $\mathbf{R}_i S$. Letting $h \prec O$, where $O \cap \mathbb{R}^n_{\leq} = \mathbf{R}_i V$ and $m_n(O) < m_n(\mathbf{R}_i V) + \varepsilon$, and letting $h_1(\mathbf{x}) \equiv h(\mathbf{R}_i(\mathbf{x}))$ for $\mathbf{x} \in U_i$, we see $h_1 \prec V$. By Corollary 12.24, we can also choose our partition of unity so that $spt(h_1) \subseteq \{\mathbf{x} \in \mathbb{R}^m : \psi_i(\mathbf{x}) = 1\}$. Thus $\psi_j h_1\left(\mathbf{R}_j^{-1}(u)\right) = 0$ unless $j = i$ when this reduces to $h_1\left(\mathbf{R}_i^{-1}(u)\right)$. Thus

$$\begin{aligned}
\varepsilon &\geq \mu(V) \geq \int_V h_1 d\mu = \int_{\Omega} h_1 d\mu = \sum_{j=1}^{p} \int_{\mathbf{R}_j U_j} \psi_j h_1\left(\mathbf{R}_j^{-1}(\mathbf{u})\right) J_j(\mathbf{u})\, du \\
&= \int_{\mathbf{R}_i U_i} h_1\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du = \int_{\mathbf{R}_i U_i} h(\mathbf{u}) J_i(\mathbf{u})\, du = \int_{\mathbf{R}_i V} h(\mathbf{u}) J_i(\mathbf{u})\, du \\
&\geq \int_O h(\mathbf{u}) J_i(\mathbf{u})\, du - K_i \varepsilon,
\end{aligned}$$

where $K_i \geq ||J_i||_{\infty}$. Now this holds for all $h \prec O$ and so

$$\int_{\mathbf{R}_i S} J_i(\mathbf{u})\, du \leq \int_{\mathbf{R}_i V} J_i(\mathbf{u})\, du \leq \int_O J_i(\mathbf{u})\, du \leq \varepsilon(1 + K_i).$$

Since ε is arbitrary, this shows $\mathbf{R}_i S$ has mesure zero with respect to the measure, $J_i(\mathbf{u})\, du$ as claimed.

Now we prove the converse. Suppose $\mathbf{R}_i S$ has $J_r(\mathbf{u})\, du$ measure zero. Then there exists an open set, O such that $O \supseteq \mathbf{R}_i S$ and

$$\int_O J_i(\mathbf{u})\, du < \varepsilon.$$

Thus $\mathbf{R}_i^{-1}(O \cap \mathbf{R}_i U_i)$ is open in Ω and contains S. Let $h \prec \mathbf{R}_i^{-1}(O \cap \mathbf{R}_i U_i)$ be such that

$$\int_{\Omega} h d\mu + \varepsilon > \mu\left(\mathbf{R}_i^{-1}(O \cap \mathbf{R}_i U_i)\right) \geq \mu(S)$$

As in the first part, we can choose our partition of unity such that $h(\mathbf{x}) = 0$ off the set,

$$\{\mathbf{x} \in \mathbb{R}^m : \psi_i(\mathbf{x}) = 1\}$$

and so as in this part of the argument,

$$\begin{aligned}\int_{\Omega} h d\mu &\equiv \sum_{j=1}^{p} \int_{\mathbf{R}_j U_j} \psi_j h\left(\mathbf{R}_j^{-1}(\mathbf{u})\right) J_j(\mathbf{u})\, du \\ &= \int_{\mathbf{R}_i U_i} h\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \\ &= \int_{O\cap \mathbf{R}_i U_i} h\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \\ &\leq \int_{O} J_i(\mathbf{u})\, du < \varepsilon\end{aligned}$$

and so $\mu(S) \leq 2\varepsilon$. Since ε is arbitrary, this proves the claim.

Now let $A \subseteq U_r$ be μ measurable. By the regularity of the measure, there exists an F_σ set, F and a G_δ set, G such that $U_r \supseteq G \supseteq A \supseteq F$ and $\mu(G \setminus F) = 0$.(Recall a G_δ set is a countable intersection of open sets and an F_σ set is a countable union of closed sets.) Then since Ω is compact, it follows each of the closed sets whose union equals F is a compact set. Thus if $F = \cup_{k=1}^{\infty} F_k$ we know $\mathbf{R}_r(F_k)$ is also a compact set and so $\mathbf{R}_r(F) = \cup_{k=1}^{\infty} \mathbf{R}_r(F_k)$ is a Borel set. Similarly, $\mathbf{R}_r(G)$ is also a Borel set. Now by the claim,

$$\int_{\mathbf{R}_r(G\setminus F)} J_r(\mathbf{u})\, du = 0.$$

We also see that since $\mathbf{R}_r$ is one to one,

$$\mathbf{R}_r G \setminus \mathbf{R}_r F = \mathbf{R}_r(G \setminus F)$$

and so

$$\mathbf{R}_r(F) \subseteq \mathbf{R}_r(A) \subseteq \mathbf{R}_r(G)$$

where $\mathbf{R}_r(G) \setminus \mathbf{R}_r(F)$ has measure zero. By completeness of the measure, $J_i(\mathbf{u})\, du$, we see $\mathbf{R}_r(A)$ is measurable. It follows that if $A \subseteq \Omega$ is μ measurable, then $\mathbf{R}_r(U_r \cap A)$ is $J_r(\mathbf{u})\, du$ measurable for all r. The converse is entirely similar.

Letting $f \in L^1(\Omega, \mu)$, we use the fact that μ is a Radon mesure to obtain a sequence of continuous functions, $\{f_k\}$ which converge to f in $L^1(\Omega, \mu)$ and for μ a.e. $\mathbf{x}$. Therefore, the sequence $\left\{f_k\left(\mathbf{R}_i^{-1}(\cdot)\right)\right\}$ is a Cauchy sequence in $L^1\left(\mathbf{R}_i U_i; \psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du\right)$. It follows there exists

$$g \in L^1\left(\mathbf{R}_i U_i; \psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du\right)$$

such that $f_k\left(\mathbf{R}_i^{-1}(\cdot)\right) \to g$ in $L^1\left(\mathbf{R}_i U_i; \psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du\right)$. By the pointwise convergence, $g(\mathbf{u}) = f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right)$ for μ a.e. $\mathbf{R}_i^{-1}(\mathbf{u}) \in U_i$. By the above claim, $g = f \circ \mathbf{R}_i^{-1}$ for a.e. $\mathbf{u} \in \mathbf{R}_i U_i$ and so

$$f \circ \mathbf{R}_i^{-1} \in L^1(\mathbf{R}_i U_i; J_i(\mathbf{u})\, du)$$

and we can write

$$\begin{aligned}\int_{\Omega} f d\mu &= \lim_{k\to\infty} \int_{\Omega} f_k d\mu = \lim_{k\to\infty} \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f_k\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \\ &= \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) g(\mathbf{u}) J_i(\mathbf{u})\, du \\ &= \sum_{i=1}^{p} \int_{\mathbf{R}_i U_i} \psi_i f\left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du.\end{aligned}$$

This proves the theorem.

In the case of a Lipschitz manifold, note that by Rademacher's theorem the set in $\mathbf{R}_r U_r$ on which $\mathbf{R}_r^{-1}$ has no derivative has Lebesgue measure zero and so contributes nothing to the definition of μ and can be ignored. We will do so from now on. Other sets of measure zero in the sets $\mathbf{R}_r U_r$ can also be ignored and we do so whenever convenient.

Corollary 20.33 *Let $f \in L^1(\Omega; \mu)$ and suppose $f(\mathbf{x}) = 0$ for all $x \notin U_r$ where $(U_r, \mathbf{R}_r)$ is a chart in a Lipschitz atlas for Ω. Then*

$$\int_\Omega f d\mu = \int_{U_r} f d\mu = \int_{\mathbf{R}_r U_r} f\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du. \tag{20.46}$$

Proof: Using regularity of the measures, we can pick a compact subset, K, of U_r such that

$$\left| \int_{U_r} f d\mu - \int_K f d\mu \right| < \varepsilon.$$

Now by Corollary 12.24, we can choose the partition of unity such that $K \subseteq \{\mathbf{x} \in \mathbb{R}^m : \psi_r(\mathbf{x}) = 1\}$. Then

$$\begin{aligned} \int_K f d\mu &= \sum_{i=1}^p \int_{\mathbf{R}_i U_i} \psi_i f \mathcal{X}_K \left(\mathbf{R}_i^{-1}(\mathbf{u})\right) J_i(\mathbf{u})\, du \\ &= \int_{\mathbf{R}_r U_r} f \mathcal{X}_K \left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du. \end{aligned}$$

Therefore, letting $K_l \uparrow \mathbf{R}_r U_r$ we can take a limit and conclude

$$\left| \int_{U_r} f d\mu - \int_{\mathbf{R}_r U_r} f\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du \right| \leq \varepsilon.$$

Since ε is arbitrary, this proves the corollary.

20.9 The divergence theorem for Lipschitz manifolds

What about writing the integral of a differential form in terms of this measure? Let ω be a differential form,

$$\omega(\mathbf{x}) = \sum_I a_I(\mathbf{x})\, d\mathbf{x}^I$$

where a_I is continuous or Borel measurable if you desire, and the sum is taken over all increasing lists from $\{1, \cdots, m\}$. We assume Ω is a Lipschitz or C^k manifold which is orientable and that $(U_r, \mathbf{R}_r)$ is an oriented atlas for Ω while, $\{\psi_r\}$ is a C^∞ partition of unity subordinate to the U_r.

Lemma 20.34 *Consider the set,*

$$S \equiv \left\{ \mathbf{x} \in \Omega : \text{ for some } r, \mathbf{x} = \mathbf{R}_r^{-1}(\mathbf{u}) \text{ where } \mathbf{x} \in U_r \text{ and } J_r(\mathbf{u}) = 0 \right\}.$$

Then $\mu(S) = 0$.

Proof: Let S_r denote those points, $\mathbf{x}$, of U_r for which $\mathbf{x} = \mathbf{R}_r^{-1}(\mathbf{u})$ and $J_r(\mathbf{u}) = 0$. Thus $S = \cup_{r=1}^p S_r$. From Corollary 20.33

$$\int_\Omega \mathcal{X}_{S_r} d\mu = \int_{\mathbf{R}_r U_r} \mathcal{X}_{S_r}\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du = 0$$

and so

$$\mu(S) \leq \sum_{k=1}^{p} \mu(S_k) = 0.$$

This proves the lemma.

With respect to the above atlas, we define a function of $\mathbf{x}$ in the following way. For $I = (i_1, \cdots, i_n)$ an increasing list of indices,

$$o^I(\mathbf{x}) \equiv \begin{cases} \left(\frac{\partial(x^{i_1}\cdots x^{i_n})}{\partial(u^1\cdots u^n)}\right)/J_r(\mathbf{u}), \text{ if } \mathbf{x} \in U_r \setminus S \\ 0 \text{ if } \mathbf{x} \in S \end{cases}$$

Then (20.36) implies this is well defined aside from a possible set of measure zero, N. Now it follows from Corollary 20.33 that $\mathbf{R}^{-1}(N)$ is a set of μ measure zero on Ω and that o^I is given by the top line off a set of μ measure zero. It also shows that if we had used a different atlas having the same orientation, then $\mathbf{o}(\mathbf{x})$ defined using the new atlas would change only on a set of μ measure zero. Also note

$$\sum_I o^I(\mathbf{x})^2 = 1 \ \mu \text{ a.e.}$$

Thus we may consider $o^I(\mathbf{x})$ to be well defined if we regard two functions as equal provided they differ only on a set of measure zero. Now we define a vector valued function, $\mathbf{o}$ having values in $\mathbb{R}^{\binom{m}{n}}$ by letting the I^{th} component of $\mathbf{o}(\mathbf{x})$ equal $o^I(\mathbf{x})$. Also define

$$\omega(\mathbf{x}) \cdot \mathbf{o}(\mathbf{x}) \equiv \sum_I a_I(\mathbf{x})\, o^I(\mathbf{x}),$$

From the definition of what we mean by the integral of a differential form, Definition 17.11, it follows that

$$\begin{aligned} \int_\Omega \omega &\equiv \sum_{r=1}^{p} \sum_I \int_{\mathbf{R}_r U_r} \psi_r\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) a_I\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \frac{\partial\left(x^{i_1}\cdots x^{i_n}\right)}{\partial\left(u^1\cdots u^n\right)} du \\ &= \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r} \psi_r\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \omega\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \cdot \mathbf{o}\left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du \\ &\equiv \int_\Omega \omega \cdot \mathbf{o}\, d\mu \end{aligned} \tag{20.47}$$

Note that $\omega \cdot \mathbf{o}$ is bounded and measurable so is in L^1. For convenience, we state the following lemma whose proof is essentially the same as the proof of Lemma 17.25.

Lemma 20.35 *Let Ω be a Lipschitz oriented manifold in $\mathbb{R}^n$ with an oriented atlas, $(U_r, \mathbf{R}_r)$. Letting $\mathbf{x} = \mathbf{R}_r^{-1}\mathbf{u}$ and letting $2 \leq j \leq n$, we have*

$$\sum_{i=1}^{n} \frac{\partial x^i}{\partial u^j} (-1)^{i+1} \frac{\partial\left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial\left(u^2 \cdots u^n\right)} = 0 \ \text{a.e.} \tag{20.48}$$

for each r. Here, $\widehat{x^i}$ means this is deleted. If for each r,

$$\frac{\partial\left(x^1 \cdots x^n\right)}{\partial\left(u^1 \cdots u^n\right)} \geq 0 \ \text{a.e.},$$

then for each r

$$\sum_{i=1}^{n} \frac{\partial x^i}{\partial u^1} (-1)^{i+1} \frac{\partial\left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial\left(u^2 \cdots u^n\right)} \geq 0 \ \text{a.e.} \tag{20.49}$$

Proof: (20.49) follows from the observation that

$$\frac{\partial\left(x^1\cdots x^n\right)}{\partial\left(u^1\cdots u^n\right)}=\sum_{i=1}^n\frac{\partial x^i}{\partial u^1}\left(-1\right)^{i+1}\frac{\partial\left(x^1\cdots\widehat{x^i}\cdots x^n\right)}{\partial\left(u^2\cdots u^n\right)}$$

by expanding the determinant,

$$\frac{\partial\left(x^1\cdots x^n\right)}{\partial\left(u^1\cdots u^n\right)},$$

along the first column. Formula (20.48) follows from the observation that the sum in (20.48) is just the determinant of a matrix which has two equal columns. This proves the lemma.

With this lemma, it is easy to verify a general form of the divergence theorem from Stoke's theorem. First we recall the definition of the divergence of a vector field.

Definition 20.36 *Let O be an open subset of $\mathbb{R}^n$ and let $\mathbf{F}(\mathbf{x})\equiv\sum_{k=1}^n F^k(\mathbf{x})\,\mathbf{e}_k$ be a vector field for which $F^k\in C^1(O)$. Then*

$$div\left(\mathbf{F}\right)\equiv\sum_{k=1}^n\frac{\partial F_k}{\partial x^k}.$$

Theorem 20.37 *Let Ω be an orientable Lipschitz n manifold with boundary in $\mathbb{R}^n$ having an oriented atlas, $(U_r,\mathbf{R}_r)$ which satisfies*

$$\frac{\partial\left(x^1\cdots x^n\right)}{\partial\left(u^1\cdots u^n\right)}\geq 0 \tag{20.50}$$

for all r. Then letting $\mathbf{n}(\mathbf{x})$ be the vector field whose i^{th} component taken with respect to the usual basis of $\mathbb{R}^n$ is given by

$$n^i\left(\mathbf{x}\right)\equiv\begin{cases}\left(-1\right)^{i+1}\frac{\partial\left(x^1\cdots\widehat{x^i}\cdots x^n\right)}{\partial\left(u^2\cdots u^n\right)}/J_r\left(\mathbf{u}\right) \text{ if } J_r\left(\mathbf{u}\right)\neq 0\\ 0 \text{ if } J_r\left(\mathbf{u}\right)=0\end{cases} \tag{20.51}$$

for $\mathbf{x}\in U_r\cap\partial\Omega$, it follows $\mathbf{n}(\mathbf{x})$ is independent of the choice of atlas provided the orientation remains unchanged. Also $\mathbf{n}(\mathbf{x})$ is a unit vector for a.e. $\mathbf{x}\in\partial\Omega$. Let $\mathbf{F}\in C^1\left(\overline{\Omega};\mathbb{R}^n\right)$. Then we have the following formula which is called the divergence theorem.

$$\int_\Omega div\left(\mathbf{F}\right)dx=\int_{\partial\Omega}\mathbf{F}\cdot\mathbf{n}d\mu, \tag{20.52}$$

where μ is the surface measure on $\partial\Omega$ defined above.

Proof: Recall that on $\partial\Omega$

$$J_r\left(\mathbf{u}\right)=\left[\sum_{i=1}^n\left(\frac{\partial\left(x^1\cdots\widehat{x^i}\cdots x^n\right)}{\partial\left(u^2\cdots u^n\right)}\right)^2\right]^{1/2}$$

From Lemma 20.26 and the definition of two atlass having the same orientation, we see that aside from sets of measure zero, the assertion about the independence of choice of atlas for the normal, $\mathbf{n}(\mathbf{x})$ is verified. Also, by Lemma 20.34, we know $J_r(\mathbf{u})\neq 0$ off some set of measure zero for each atlas and so $\mathbf{n}(\mathbf{x})$ is a unit vector for μ a.e. $\mathbf{x}$.

Now we define the differential form,

$$\omega \equiv \sum_{i=1}^{n} (-1)^{i+1} F_i(\mathbf{x})\, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n.$$

Then from the definition of $d\omega$,

$$d\omega = div(\mathbf{F})\, dx^1 \wedge \cdots \wedge dx^n.$$

Now let $\{\psi_r\}$ be a partition of unity subordinate to the U_r. Then using (20.50) and the change of variables formula,

$$\begin{aligned}
\int_\Omega d\omega &\equiv \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r} (\psi_r div(\mathbf{F})) \left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \frac{\partial (x^1 \cdots x^n)}{\partial (u^1 \cdots u^n)} du \\
&= \sum_{r=1}^{p} \int_{U_r} (\psi_r div(\mathbf{F}))(\mathbf{x})\, dx = \int_\Omega div(\mathbf{F})\, dx.
\end{aligned}$$

Now

$$\begin{aligned}
\int_{\partial\Omega} \omega &\equiv \sum_{i=1}^{n} (-1)^{i+1} \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \psi_r F_i \left(\mathbf{R}_r^{-1}(\mathbf{u})\right) \frac{\partial \left(x^1 \cdots \widehat{x^i} \cdots x^n\right)}{\partial (u^2 \cdots u^n)} du^2 \cdots du^n \\
&= \sum_{r=1}^{p} \int_{\mathbf{R}_r U_r \cap \mathbb{R}_0^n} \psi_r \left(\sum_{i=1}^{n} F_i n^i\right) \left(\mathbf{R}_r^{-1}(\mathbf{u})\right) J_r(\mathbf{u})\, du^2 \cdots du^n \\
&\equiv \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n}\, d\mu.
\end{aligned}$$

By Stoke's theorem,

$$\int_\Omega div(\mathbf{F})\, dx = \int_\Omega d\omega = \int_{\partial\Omega} \omega = \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n}\, d\mu$$

and this proves the theorem.

Definition 20.38 *In the context of the divergence theorem, the vector,* $\mathbf{n}$ *is called the unit outer normal.*

20.10 Exercises

1. Let E be a Lebesgue measurable set. We say $\mathbf{x} \in E$ is a point of density if

$$\lim_{r \to 0} \frac{m(E \cap B(\mathbf{x}, r))}{m(B(\mathbf{x}, r))} = 1.$$

   Show that a.e. point of E is a point of density.

2. Let $(\Omega, \mathcal{S}, \mu)$ be any σ finite measure space, $f \geq 0$, f real-valued, and measurable. Let ϕ be an increasing C^1 function with $\phi(0) = 0$. Show

$$\int_\Omega \phi \circ f d\mu = \int_0^\infty \phi'(t) \mu([f(x) > t]) dt.$$

   Hint:

$$\int_\Omega \phi(f(x)) d\mu = \int_\Omega \int_0^{f(x)} \phi'(t) dt d\mu = \int_\Omega \int_0^\infty \mathcal{X}_{[0, f(x))}(t) \phi'(t) dt d\mu.$$

   Argue $\phi'(t)\mathcal{X}_{[0,f(x))}(t)$ is product measurable and use Fubini's theorem. The function $t \to \mu([f(x) > t])$ is called the distribution function.

3. Let f be in $L^1_{loc}(\mathbb{R}^n)$. Show Mf is Borel measurable.

4. If $f \in L^p, 1 < p < \infty$, show $Mf \in L^p$ and

$$||Mf||_p \leq A(p,n)||f||_p.$$

Hint: Let

$$f_1(\mathbf{x}) \equiv \begin{cases} f(\mathbf{x}) \text{ if } |f(\mathbf{x})| > \alpha/2, \\ 0 \text{ if } |f(\mathbf{x})| \leq \alpha/2. \end{cases}$$

Argue $[Mf(\mathbf{x}) > \alpha] \subseteq [Mf_1(\mathbf{x}) > \alpha/2]$. Then by Problem 2,

$$\int (Mf)^p dx = \int_0^\infty p\alpha^{p-1} m([Mf > \alpha]) d\alpha$$

$$\leq \int_0^\infty p\alpha^{p-1} m([Mf_1 > \alpha/2]) d\alpha.$$

Now use Theorem 20.4 and Fubini's Theorem as needed.

5. Show $|f(\mathbf{x})| \leq Mf(\mathbf{x})$ at every Lebesgue point of f whenever $f \in L^1(\mathbb{R}^n)$.

6. In the proof of the Vitali covering theorem there is nothing sacred about the constant $\frac{1}{2}$. Go through the proof replacing this constant with λ where $\lambda \in (0,1)$. Show that it follows that for every $\delta > 0$, the conclusion of the Vitali covering theorem follows with 5 replaced by $(3+\delta)$ in the definition of $\widehat{B}$.

7. Suppose A is covered by a finite collection of Balls, $\mathcal{F}$. Show that then there exists a disjoint collection of these balls, $\{B_i\}_{i=1}^p$, such that $A \subseteq \cup_{i=1}^p \widehat{B}_i$ where 5 can be replaced with 3 in the definition of $\widehat{B}$. **Hint:** Since the collection of balls is finite, we can arrange them in order of decreasing radius.

8. The result of this Problem is sometimes called the Vitali Covering Theorem. It is very important in some applications. It has to do with covering sets in except for a set of measure zero with disjoint balls. Let $E \subseteq \mathbb{R}^n$ be Lebesgue measurable, $m(E) < \infty$, and let $\mathcal{F}$ be a collection of balls that cover E in the sense of Vitali. This means that if $\mathbf{x} \in E$ and $\varepsilon > 0$, then there exists $B \in \mathcal{F}$, diameter of $B < \varepsilon$ and $\mathbf{x} \in B$. Show there exists a countable sequence of disjoint balls of $\mathcal{F}$, $\{B_j\}$, such that $m(E \setminus \cup_{j=1}^\infty B_j) = 0$. **Hint:** Let $E \subseteq U$, U open and

$$m(E) > (1 - 10^{-n}) m(U).$$

Let $\{B_j\}$ be disjoint,

$$E \subseteq \cup_{j=1}^\infty \hat{B}_j, \; B_j \subseteq U.$$

Thus

$$m(E) \leq 5^n m(\cup_{j=1}^\infty B_j).$$

Then

$$m(E) > (1 - 10^{-n}) m(U)$$

$$\geq (1 - 10^{-n})[m(E \setminus \cup_{j=1}^\infty B_j) + m(\cup_{j=1}^\infty B_j)]$$

$$\geq (1-10^{-n})[m(E \setminus \cup_{j=1}^{\infty} B_j) + 5^{-n} m(E)].$$

Hence

$$m(E \setminus \cup_{j=1}^{\infty} B_j) \leq (1-10^{-n})^{-1}(1-(1-10^{-n})5^{-n})m(E).$$

Let $(1-10^{-n})^{-1}(1-(1-10^{-n})5^{-n}) < \theta < 1$ and pick N_1 large enough that

$$\theta m(E) \geq m(E \setminus \cup_{j=1}^{N_1} \overline{B_j}).$$

Let $\mathcal{F}_1 = \{B \in \mathcal{F} : B_j \cap B = \emptyset,\ j = 1, \cdots, N_1\}$. If $E \setminus \cup_{j=1}^{N_1} \overline{B_j} \neq \emptyset$, then $\mathcal{F}_1 \neq \emptyset$ and covers $E \setminus \cup_{j=1}^{N_1} \overline{B_j}$ in the sense of Vitali. Repeat the same argument, letting $E \setminus \cup_{j=1}^{N_1} \overline{B_j}$ play the role of E.

9. $\uparrow$Suppose E is a Lebesgue measurable set which has positive measure and let B be an arbitrary open ball and let D be a set dense in $\mathbb{R}^n$. Establish the result of Smítal, [4]which says that under these conditions, $\overline{m_n}\left((E+D) \cap B\right) = m_n(B)$ where here $\overline{m_n}$ denotes the outer measure determined by m_n. Is this also true for X, an arbitrary possibly non measurable set replacing E in which $\overline{m_n}(X) > 0$? **Hint:** Let x be a point of density of E and let D' denote those elements of D, d, such that $d + x \in B$. Thus D' is dense in B. Now use translation invariance of Lebesgue measure to verify that there exists, $R > 0$ such that if $r < R$, we have the following holding for $d \in D'$ and $r_d < R$.

$$\overline{m_n}\left((E+D) \cap B(x+d, r_d)\right) \geq$$

$$m_n\left((E+d) \cap B(x+d, r_d)\right) \geq (1-\varepsilon)\, m_n\left(B(x+d, r_d)\right).$$

Argue the balls, $m_n\left(B(x+d, r_d)\right)$, form a Vitali cover of B.

10. Suppose λ is a Radon measure on $\mathbb{R}^n$, and $\lambda(S) < \infty$ where $m_n(S) = 0$ and $\lambda(E) = \lambda(E \cap S)$. (If $\lambda(E) = \lambda(E \cap S)$ where $m_n(S) = 0$ we say $\lambda \perp m_n$.) Show that for m_n a.e. $\mathbf{x}$ the following holds. If $B_i \downarrow \{\mathbf{x}\}$, then $\lim_{i\to\infty} \frac{\lambda(B_i)}{m_n(B_i)} = 0$. **Hint**: You might try this. Set ε, $r > 0$, and let

$$E_\varepsilon = \{\mathbf{x} \in S^C : \text{ there exists } \{B_i^{\mathbf{x}}\}, B_i^{\mathbf{x}} \downarrow \{\mathbf{x}\} \text{ with } \frac{\lambda(B_i^{\mathbf{x}})}{m_n(B_i^{\mathbf{x}})} \geq \varepsilon\}.$$

Let K be compact, $\lambda(S \setminus K) < r\varepsilon$. Let $\mathcal{F}$ consist of those balls just described that do not intersect K and which have radius < 1. This is a Vitali cover of E_ε. Let $B_1, \cdots, B_k$ be disjoint balls from $\mathcal{F}$ and

$$\overline{m}_n(E_\varepsilon \setminus \cup_{i=1}^k B_i) < r.$$

Then

$$\overline{m}_n(E_\varepsilon) < r + \sum_{i=1}^k m_n(B_i) < r + \varepsilon^{-1} \sum_{i=1}^k \lambda(B_i) =$$

$$r + \varepsilon^{-1} \sum_{i=1}^k \lambda(B_i \cap S) \leq r + \varepsilon^{-1} \lambda(S \setminus K) < 2r.$$

Since r is arbitrary, $m_n(E_\varepsilon) = 0$. Consider $E = \cup_{k=1}^{\infty} E_{k^{-1}}$ and let $\mathbf{x} \notin S \cup E$.

11. $\uparrow$ Is it necessary to assume $\lambda(S) < \infty$ in Problem 10? Explain.

12. $\uparrow$ Let S be an increasing function on $\mathbb{R}$ which is right continuous,

$$\lim_{x\to-\infty} S(x) = 0,$$

and S is bounded. Let λ be the measure representing $\int f dS$. Thus $\lambda((-\infty, x]) = S(x)$. Suppose $\lambda \perp m$. Show $S'(x) = 0$ m a.e. **Hint**:

$$0 \leq h^{-1}(S(x+h) - S(x))$$

$$= \frac{\lambda((x, x+h])}{m((x, x+h])} \leq 3\frac{\lambda((x-h, x+2h))}{m((x-h, x+2h))}.$$

Now apply Problem 10. Similarly $h^{-1}(S(x) - S(x-h)) \to 0$.

13. $\uparrow$ Let f be increasing, bounded above and below, and right continuous. Show $f'(x)$ exists a.e. **Hint**: See Problem 6 of Chapter 18.

14. $\uparrow$ Suppose $|f(x) - f(y)| \leq K|x-y|$. Show there exists $g \in L^\infty(\mathbb{R})$, $||g||_\infty \leq K$, and

$$f(y) - f(x) = \int_x^y g(t)dt.$$

Hint: Let $F(x) = Kx + f(x)$ and let λ be the measure representing $\int f dF$. Show $\lambda << m$. What does this imply about the differentiability of a Lipschitz continuous function?

15. $\uparrow$ Let f be increasing. Show $f'(x)$ exists a.e.

16. Let $f(x) = x^2$. Thus $\int_{-1}^1 f(x)dx = 2/3$. Let's change variables. $u = x^2$, $du = 2xdx = 2u^{1/2}dx$. Thus

$$2/3 = \int_{-1}^1 x^2 dx = \int_1^1 u/2u^{1/2} du = 0.$$

Can this be done correctly using a change of variables theorem?

17. Consider the construction employed to obtain the Cantor set, but instead of removing the middle third interval, remove only enough that the sum of the lengths of all the open intervals which are removed is less than one. That which remains is called a fat Cantor set. Show it is a compact set which has measure greater than zero which contains no interval and has the property that every point is a limit point of the set. Let P be such a fat Cantor set and consider

$$f(x) = \int_0^x \mathcal{X}_{P^C}(t)\, dt.$$

Show that f is a strictly increasing function which has the property that its derivative equals zero on a set of positive measure.

18. Let f be a function defined on an interval, (a, b). The Dini derivates are defined as

$$D_+ f(x) \equiv \lim \inf_{h\to 0+} \frac{f(x+h) - f(x)}{h}, D^+ f(x) \equiv \lim \sup_{h\to 0+} \frac{f(x+h) - f(x)}{h}$$

$$D_- f(x) \equiv \lim \inf_{h\to 0+} \frac{f(x) - f(x-h)}{h}, D^- f(x) \equiv \lim \sup_{h\to 0+} \frac{f(x) - f(x-h)}{h}.$$

Suppose f is continuous on (a,b) and for all $x \in (a,b)$, $D_+ f(x) \geq 0$. Show that then f is increasing on (a,b). **Hint:** Consider the function, $H(x) \equiv f(x)(d-c) - x(f(d) - f(c))$ where $a < c < d < b$. Thus $H(c) = H(d)$. Also it is easy to see that H cannot be constant if $f(d) < f(c)$ due to the assumption that $D_+ f(x) \geq 0$. If there exists $x_1 \in (a,b)$ where $H(x_1) > H(c)$, then let $x_0 \in (c,d)$ be the point where the maximum of f occurs. Consider $D_+ f(x_0)$. If, on the other hand, $H(x) < H(c)$ for all $x \in (c,d)$, then consider $D_+ H(c)$.

19. $\uparrow$ Suppose in the situation of the above problem we only know $D_+ f(x) \geq 0$ a.e. Does the conclusion still follow? What if we only know $D_+ f(x) \geq 0$ for every x outside a countable set? **Hint:** In the case of $D_+ f(x) \geq 0$,consider the bad function in the exercises for the chapter on the construction of measures which was based on the Cantor set. In the case where $D_+ f(x) \geq 0$ for all but countably many x, by replacing $f(x)$ with $\widetilde{f}(x) \equiv f(x) + \varepsilon x$, consider the situation where $D_+ \widetilde{f}(x) > 0$ for all but countably many x. If in this situation, $\widetilde{f}(c) > \widetilde{f}(d)$ for some $c < d$, and $y \in \left(\widetilde{f}(d), \widetilde{f}(c)\right)$,let
$$z \equiv \sup \left\{ x \in [c,d] : \widetilde{f}(x) > y_0 \right\}.$$
Show that $\widetilde{f}(z) = y_0$ and $D_+ \widetilde{f}(z) \leq 0$. Conclude that if $\widetilde{f}$ fails to be increasing, then $D_+ \widetilde{f}(z) \leq 0$ for uncountably many points, z. Now draw a conclusion about f.

20. Consider in the formula for $\Gamma(\alpha+1)$ the following change of variables. $t = \alpha + \alpha^{1/2} s$. Then in terms of the new variable, s, the formula for $\Gamma(\alpha+1)$ is
$$e^{-\alpha} \alpha^{\alpha+\frac{1}{2}} \int_{-\sqrt{\alpha}}^{\infty} e^{-\sqrt{\alpha} s} \left(1 + \frac{s}{\sqrt{\alpha}}\right)^{\alpha} ds = e^{-\alpha} \alpha^{\alpha+\frac{1}{2}} \int_{-\sqrt{\alpha}}^{\infty} e^{\alpha \left[\ln\left(1+\frac{s}{\sqrt{\alpha}}\right) - \frac{s}{\sqrt{\alpha}}\right]} ds$$
Show the integrand converges to $e^{-\frac{s^2}{2}}$. Show that then
$$\lim_{\alpha \to \infty} \frac{\Gamma(\alpha+1)}{e^{-\alpha} \alpha^{\alpha+(1/2)}} = \int_{-\infty}^{\infty} e^{\frac{-s^2}{2}} ds = \sqrt{2\pi}.$$
You will need to obtain a dominating function for the integral so that you can use the dominated convergence theorem. This formula is known as Stirling's formula.

21. Let Ω be an oriented Lipschitz n manifold in $\mathbb{R}^n$ for which
$$\frac{\partial\left(x^1 \cdots x^n\right)}{\partial\left(u^1 \cdots u^n\right)} \geq 0 \text{ a.e.}$$
for all the charts, $\mathbf{x} = \mathbf{R}_r(\mathbf{u})$, and suppose $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ for $m \geq n$ is a Lipschitz continuous function such that there exists a Lipschitz continuous function, $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^n$ such that $\mathbf{G} \circ \mathbf{F}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \Omega$. Show that $\mathbf{F}(\Omega)$ is an oriented Lipschitz n manifold. Now suppose $\omega \equiv \sum_I a_I(\mathbf{y}) d\mathbf{y}^I$ is a differential form on $\mathbf{F}(\Omega)$. Show
$$\int_{\mathbf{F}(\Omega)} \omega = \int_{\Omega} \sum_I a_I(\mathbf{F}(\mathbf{x})) \frac{\partial\left(y^{i_1} \cdots y^{i_n}\right)}{\partial\left(x^1 \cdots x^n\right)} dx.$$
In this case, we say that $\mathbf{F}(\Omega)$ is a parametrically defined manifold. Note this shows how to compute the integral of a differential form on such a manifold without dragging in a partition of unity. Also note that Ω could be a box or some collection of boxes pased togenter along edges. Can you get a similar result in the case where $\mathbf{F}$ satisfies the conditions of Theorem 20.29?

22. Let $\mathbf{h}$:$\mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{h}$ is Lipschitz. Let
$$A = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$$
where $\mathbf{c}$ is a constant vector. Show $J(\mathbf{x}) = 0$ a.e. on A. **Hint:** Use Theorem 20.22.

23. Let U be an open subset of $\mathbb{R}^n$ and let $\mathbf{h} : U \to \mathbb{R}^n$ be differentiable on $A \subseteq U$ for some A a Lebesgue measurable set. Show that if $T \subseteq A$ and $m_n(T) = 0$, then $m_n(\mathbf{h}(T)) = 0$. **Hint:** Let

$$T_k \equiv \{\mathbf{x} \in T : ||D\mathbf{h}(\mathbf{x})|| < k\}$$

and let $\epsilon > 0$ be given. Now let V be an open set containing T_k which is contained in U such that $m_n(V) < \frac{\epsilon}{k^n 5^n}$ and let $\delta > 0$ be given. Using differentiability of $\mathbf{h}$, for each $\mathbf{x} \in T_k$ there exists $r_{\mathbf{x}} < \delta$ such that $B(\mathbf{x}, 5r_{\mathbf{x}}) \subseteq V$ and

$$\mathbf{h}(B(\mathbf{x}, r_{\mathbf{x}})) \subseteq B(\mathbf{h}(\mathbf{x}), 5kr_{\mathbf{x}}).$$

Use the same argument found in Lemma 20.13 to conclude

$$m_n(\mathbf{h}(T_k)) = 0.$$

Now

$$m_n(\mathbf{h}(T)) = \lim_{k\to\infty} m_n(\mathbf{h}(T_k)) = 0.$$

24. $\uparrow$In the context of 23 show that if S is a Lebesgue measurable subset of A, then $\mathbf{h}(S)$ is m_n measurable. **Hint:** Use the same argument found in Lemma 20.14.

25. $\uparrow$ Suppose also that $\mathbf{h}$ is differentiable on U. Show the following holds. Let $\mathbf{x} \in A$ be a point where $D\mathbf{h}(\mathbf{x})^{-1}$ exists. Then if $\epsilon \in (0, 1)$ the following hold for all r small enough.

$$m_n\left(\mathbf{h}\left(\overline{B(\mathbf{x}, r)}\right)\right) = m_n(\mathbf{h}(B(\mathbf{x}, r))) \geq m_n(D\mathbf{h}(\mathbf{x}) B(\mathbf{0}, r(1-\epsilon))), \tag{20.53}$$

$$\mathbf{h}(B(\mathbf{x}, r)) \subseteq \mathbf{h}(\mathbf{x}) + D\mathbf{h}(\mathbf{x}) B(\mathbf{0}, r(1+\epsilon)), \tag{20.54}$$

$$m_n(\mathbf{h}(B(\mathbf{x}, r))) \leq m_n(D\mathbf{h}(\mathbf{x}) B(\mathbf{0}, r(1+\epsilon))). \tag{20.55}$$

If $U \setminus A$ has measure 0, then for $\mathbf{x} \in A$,

$$\lim_{r\to 0} \frac{m_n(\mathbf{h}(B(\mathbf{x}, r) \cap A))}{m_n(\mathbf{h}(B(\mathbf{x}, r)))} = 1. \tag{20.56}$$

Also show that for $\mathbf{x} \in A$,

$$J(\mathbf{x}) = \lim_{r\to 0} \frac{m_n(\mathbf{h}(B(\mathbf{x}, r)))}{m_n(B(\mathbf{x}, r))}, \tag{20.57}$$

where $J(\mathbf{x}) \equiv \det(D\mathbf{h}(\mathbf{x}))$.

26. $\uparrow$ Assuming the context of the above problems, let $\mathbf{h}$ be one to one on A and establish that for F Borel measurable in $\mathbb{R}^n$

$$\int_{\mathbf{h}(A)} \mathcal{X}_F(\mathbf{y})\, dm_n = \int_A \mathcal{X}_F(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm.$$

This is like (20.26). Next show, using the arguments of (20.27) - (20.31), that a change of variables formula of the form

$$\int_{\mathbf{h}(A)} g(\mathbf{y})\, dm_n = \int_A g(\mathbf{h}(\mathbf{x}))\, J(\mathbf{x})\, dm$$

holds whenever $g : \mathbf{h}(A) \to [0, \infty]$ is m_n measurable.

27. Extend the theorem about integration and the Brouwer degree to more general classes of mappings than C^1 mappings.

The complex numbers

In this chapter we consider the complex numbers, $\mathbb{C}$ and a few basic topics such as the roots of a complex number. Just as a real number should be considered as a point on the line, a complex number is considered a point in the plane. We can identify a point in the plane in the usual way using the Cartesian coordinates of the point. Thus (a, b) identifies a point whose x coordinate is a and whose y coordinate is b. In dealing with complex numbers, we write such a point as $a + ib$ and multiplication and addition are defined in the most obvious way subject to the convention that $i^2 = -1$. Thus,

$$(a + ib) + (c + id) = (a + c) + i (b + d)$$

and

$$(a + ib)(c + id) = (ac - bd) + i (bc + ad).$$

We can also verify that every non zero complex number, $a + ib$, with $a^2 + b^2 \neq 0$, has a unique multiplicative inverse.

$$\frac{1}{a + ib} = \frac{a - ib}{a^2 + b^2} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}.$$

Theorem 21.1 *The complex numbers with multiplication and addition defined as above form a field.*

The field of complex numbers is denoted as $\mathbb{C}$. An important construction regarding complex numbers is the complex conjugate denoted by a horizontal line above the number. It is defined as follows.

$$\overline{a + ib} = a - ib.$$

What it does is reflect a given complex number across the x axis. Algebraically, the following formula is easy to obtain.

$$\left(\overline{a + ib}\right)(a + ib) = a^2 + b^2.$$

The length of a complex number, refered to as the modulus of z and denoted by $|z|$ is given by

$$|z| \equiv \left(x^2 + y^2\right)^{1/2} = (z\overline{z})^{1/2},$$

and we make $\mathbb{C}$ into a metric space by defining the distance between two complex numbers, z and w as

$$d(z, w) \equiv |z - w|.$$

We see therefore, that this metric on $\mathbb{C}$ is the same as the usual metric of $\mathbb{R}^2$. A sequence, $z_n \to z$ if and only if $x_n \to x$ in $\mathbb{R}$ and $y_n \to y$ in $\mathbb{R}$ where $z = x + iy$ and $z_n = x_n + iy_n$. For example if $z_n = \frac{n}{n+1} + i\frac{1}{n}$, then $z_n \to 1 + 0i = 1$.

Definition 21.2 *A sequence of complex numbers, $\{z_n\}$ is a Cauchy sequence if for every $\varepsilon > 0$ there exists N such that $n, m > N$ implies $|z_n - z_m| < \varepsilon$.*

This is the usual definition of Cauchy sequence. There are no new ideas here.

Proposition 21.3 *The complex numbers with the norm just mentioned forms a complete normed linear space.*

Proof: Let $\{z_n\}$ be a Cauchy sequence of complex numbers with $z_n = x_n + iy_n$. Then $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences of real numbers and so they converge to real numbers, x and y respectively. Thus $z_n = x_n + iy_n \to x + iy$. By Theorem 21.1 $\mathbb{C}$ is a linear space with the field of scalars equal to $\mathbb{C}$. It only remains to verify that $|\ |$ satisfies the axioms of a norm which are:

$$|z + w| \leq |z| + |w|$$

$$|z| \geq 0 \text{ for all } z$$

$$|z| = 0 \text{ if and only if } z = 0$$

$$|\alpha z| = |\alpha| \, |z| \, .$$

We leave this as an exercise.

Definition 21.4 *An infinite sum of complex numbers is defined as the limit of the sequence of partial sums. Thus,*

$$\sum_{k=1}^{\infty} a_k \equiv \lim_{n \to \infty} \sum_{k=1}^{n} a_k.$$

Just as in the case of sums of real numbers, we see that an infinite sum converges if and only if the sequence of partial sums is a Cauchy sequence.

Definition 21.5 *We say a sequence of functions of a complex variable, $\{f_n\}$ converges uniformly to a function, g for $z \in S$ if for every $\varepsilon > 0$ there exists N_ε such that if $n > N_\varepsilon$, then*

$$|f_n(z) - g(z)| < \varepsilon$$

for all $z \in S$. The infinite sum $\sum_{k=1}^{\infty} f_n$ converges uniformly on S if the partial sums converge uniformly on S.

Proposition 21.6 *A sequence of functions, $\{f_n\}$ defined on a set S, converges uniformly to some function, g if and only if for all $\varepsilon > 0$ there exists N_ε such that whenever $m, n > N_\varepsilon$,*

$$||f_n - f_m||_\infty < \varepsilon.$$

Here $||f||_\infty \equiv \sup \{|f(z)| : z \in S\}$.

Just as in the case of functions of a real variable, we have the Weierstrass M test.

Proposition 21.7 *Let $\{f_n\}$ be a sequence of complex valued functions defined on $S \subseteq \mathbb{C}$. Suppose there exists M_n such that $||f_n||_\infty < M_n$ and $\sum M_n$ converges. Then $\sum f_n$ converges uniformly on S.*

Since every complex number can be considered a point in $\mathbb{R}^2$, we define the polar form of a complex number as follows. If $z = x + iy$ then $\left(\frac{x}{|z|}, \frac{y}{|z|}\right)$ is a point on the unit circle because

$$\left(\frac{x}{|z|}\right)^2 + \left(\frac{y}{|z|}\right)^2 = 1.$$

Therefore, there is an angle θ such that

$$\left(\frac{x}{|z|}, \frac{y}{|z|}\right) = (\cos\theta, \sin\theta).$$

It follows that

$$z = x + iy = |z|(\cos\theta + i\sin\theta).$$

This is the polar form of the complex number, $z = x + iy$.

One of the most important features of the complex numbers is that you can always obtain n *nth* roots of any complex number. To begin with we need a fundamental result known as De Moivre's theorem.

Theorem 21.8 *Let $r > 0$ be given. Then if n is a positive integer,*

$$[r(\cos t + i\sin t)]^n = r^n(\cos nt + i\sin nt).$$

Proof: It is clear the formula holds if $n = 1$. Suppose it is true for n.

$$[r(\cos t + i\sin t)]^{n+1} = [r(\cos t + i\sin t)]^n [r(\cos t + i\sin t)]$$

which by induction equals

$$= r^{n+1}(\cos nt + i\sin nt)(\cos t + i\sin t)$$

$$= r^{n+1}((\cos nt\cos t - \sin nt\sin t) + i(\sin nt\cos t + \cos nt\sin t))$$

$$= r^{n+1}(\cos(n+1)t + i\sin(n+1)t)$$

by standard trig. identities.

Corollary 21.9 *Let z be a non zero complex number. Then there are always exactly k kth roots of z in $\mathbb{C}$.*

Proof: Let $z = x + iy$. Then

$$z = |z|\left(\frac{x}{|z|} + i\frac{y}{|z|}\right)$$

and from the definition of $|z|$,

$$\left(\frac{x}{|z|}\right)^2 + \left(\frac{y}{|z|}\right)^2 = 1.$$

Thus $\left(\frac{x}{|z|}, \frac{y}{|z|}\right)$ is a point on the unit circle and so

$$\frac{y}{|z|} = \sin t,\ \frac{x}{|z|} = \cos t$$

for a unique $t \in [0, 2\pi)$. By De Moivre's theorem, a number is a *kth* root of z if and only if it is of the form

$$|z|^{1/k}\left(\cos\left(\frac{t + 2l\pi}{k}\right) + i\sin\left(\frac{t + 2l\pi}{k}\right)\right)$$

for l an integer. By the fact that the cos and sin are 2π periodic, if $l = k$ in the above formula the same complex number is obtained as if $l = 0$. Thus there are exactly k of these numbers.

If $S \subseteq \mathbb{C}$ and $f : S \to \mathbb{C}$, we say f is continuous if whenever $z_n \to z \in S$, it follows that $f(z_n) \to f(z)$. Thus f is continuous if it takes converging sequences to converging sequences.

21.1 Exercises

1. Let $z = 3 + 4i$. Find the polar form of z and obtain all cube roots of z.

2. Prove Propositions 21.6 and 21.7.

3. Verify the complex numbers form a field.

4. Prove that $\overline{\prod_{k=1}^{n} z_k} = \prod_{k=1}^{n} \overline{z}_k$. In words, show the conjugate of a product is equal to the product of the conjugates.

5. Prove that $\overline{\sum_{k=1}^{n} z_k} = \sum_{k=1}^{n} \overline{z}_k$. In words, show the conjugate of a sum equals the sum of the conjugates.

6. Let $P(z)$ be a polynomial having real coefficients. Show the zeros of $P(z)$ occur in conjugate pairs.

7. If A is a real $n \times n$ matrix and $A\mathbf{x} = \lambda \mathbf{x}$, show that $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$.

8. Tell what is wrong with the following proof that $-1 = 1$.
$$-1 = i^2 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)^2} = \sqrt{1} = 1.$$

9. If $z = |z|(\cos\theta + i\sin\theta)$ and $w = |w|(\cos\alpha + i\sin\alpha)$, show
$$zw = |z|\,|w|\,(\cos(\theta+\alpha) + i\sin(\theta+\alpha)).$$

10. Since each complex number, $z = x + iy$ can be considered a vector in $\mathbb{R}^2$, we can also consider it a vector in $\mathbb{R}^3$ and consider the cross product of two complex numbers. Recall from calculus that for $\mathbf{x} \equiv (a,b,c)$ and $\mathbf{y} \equiv (d,e,f)$, two vectors in $\mathbb{R}^3$,
$$\mathbf{x} \times \mathbf{y} \equiv \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & b & c \\ d & e & f \end{pmatrix}$$
and that geometrically $|\mathbf{x} \times \mathbf{y}| = |\mathbf{x}|\,|\mathbf{y}|\sin\theta$, the area of the parallelogram spanned by the two vectors, $\mathbf{x}, \mathbf{y}$ and the triple, $\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y}$ forms a right handed system. Show
$$z_1 \times z_2 = \operatorname{Im}(\overline{z}_1 z_2)\,\mathbf{k}.$$
Thus the area of the parallelogram spanned by z_1 and z_2 equals $|\operatorname{Im}(\overline{z}_1 z_2)|$.

11. Prove that $f : S \subseteq \mathbb{C} \to \mathbb{C}$ is continuous at $z \in S$ if and only if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $w \in S$ and $|w - z| < \delta$, it follows that $|f(w) - f(z)| < \varepsilon$.

12. Verify that every polynomial $p(z)$ is continuous on $\mathbb{C}$.

13. Show that if $\{f_n\}$ is a sequence of functions converging uniformly to a function, f on $S \subseteq \mathbb{C}$ and if f_n is continuous on S, then so is f.

14. Show that if $|z| < 1$, then $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$.

15. Show that whenever $\sum a_n$ converges it follows that $\lim_{n\to\infty} a_n = 0$. Give an example in which $\lim_{n\to\infty} a_n = 0$, $a_n \geq a_{n+1}$ and yet $\sum a_n$ fails to converge to a number.

16. Prove the root test for series of complex numbers. If $a_k \in \mathbb{C}$ and $r \equiv \limsup_{n\to\infty} |a_n|^{1/n}$ then
$$\sum_{k=0}^{\infty} a_k \begin{cases} \text{converges absolutely if } r < 1 \\ \text{diverges if } r > 1 \\ \text{test fails if } r = 1. \end{cases}$$

17. Does $\lim_{n\to\infty} n\left(\frac{2+i}{3}\right)^n$ exist? Tell why and find the limit if it does exist.

18. Let $A_0 = 0$ and let $A_n \equiv \sum_{k=1}^{n} a_k$ if $n > 0$. Prove the partial summation formula,

$$\sum_{k=p}^{q} a_k b_k = A_q b_q - A_{p-1} b_p + \sum_{k=p}^{q-1} A_k \left(b_k - b_{k+1}\right).$$

Now using this formula, suppose $\{b_n\}$ is a sequence of real numbers which converges to 0 and is decreasing. Determine those values of ω such that $|\omega| = 1$ and $\sum_{k=1}^{\infty} b_k \omega^k$ converges. **Hint:** From Problem 15 you have an example of a sequence $\{b_n\}$ which shows that $\omega = 1$ is not one of those values of ω.

19. Let $f : U \subseteq \mathbb{C} \to \mathbb{C}$ be given by $f(x+iy) = u(x,y) + iv(x,y)$. Show f is continuous on U if and only if $u : U \to \mathbb{R}$ and $v : U \to \mathbb{R}$ are both continuous.

21.2 The extended complex plane

The set of complex numbers has already been considered along with the topology of $\mathbb{C}$ which is nothing but the topology of $\mathbb{R}^2$. Thus, for $z_n = x_n + iy_n$ we say $z_n \to z \equiv x + iy$ if and only if $x_n \to x$ and $y_n \to y$. The norm in $\mathbb{C}$ is given by

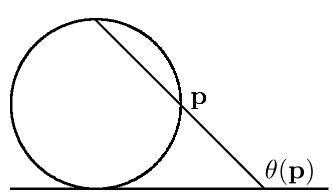
$$|x+iy| \equiv ((x+iy)(x-iy))^{1/2} = \left(x^2+y^2\right)^{1/2}$$

which is just the usual norm in $\mathbb{R}^2$ identifying (x,y) with $x+iy$. Therefore, $\mathbb{C}$ is a complete metric space and we have the Heine Borel theorem that compact sets are those which are closed and bounded. Thus, as far as topology is concerned, there is nothing new about $\mathbb{C}$.

We need to consider another general topological space which is related to $\mathbb{C}$. It is called the extended complex plane, denoted by $\widehat{\mathbb{C}}$ and consisting of the complex plane, $\mathbb{C}$ along with another point not in $\mathbb{C}$ known as ∞. For example, ∞ could be any point in $\mathbb{R}^3$. We say a sequence of complex numbers, z_n, converges to ∞ if, whenever K is a compact set in $\mathbb{C}$, there exists a number, N such that for all $n > N$, $z_n \notin K$. Since compact sets in $\mathbb{C}$ are closed and bounded, this is equivalent to saying that for all $R > 0$, there exists N such that if $n > N$, then $z_n \notin B(0,R)$ which is the same as saying $\lim_{n\to\infty} |z_n| = \infty$ where this last symbol has the same meaning as it does in calculus.

A geometric way of understanding this in terms of more familiar objects involves a concept known as the Riemann sphere.

Consider the unit sphere, S^2 given by $(z-1)^2 + y^2 + x^2 = 1$. We define a map from the unit sphere with the point, $(0,0,2)$ left out which is one to one onto $\mathbb{R}^2$ as follows.



We extend a line from the north pole of the sphere, the point $(0,0,2)$, through the point on the sphere, $\mathbf{p}$, until it intersects a unique point on $\mathbb{R}^2$. This mapping, known as stereographic projection, which we will denote for now by θ, is clearly continuous because it takes converging sequences, to converging sequences. Furthermore, it is clear that θ^{-1} is also continuous. In terms of the extended complex plane, $\widehat{\mathbb{C}}$, we see a sequence, z_n converges to ∞ if and only if $\theta^{-1} z_n$ converges to $(0,0,2)$ and a sequence, z_n converges to $z \in \mathbb{C}$ if and only if $\theta^{-1}(z_n) \to \theta^{-1}(z)$.

21.3 Exercises

1. Try to find an explicit formula for θ and θ^{-1}.

2. What does the mapping θ^{-1} do to lines and circles?

3. Show that S^2 is compact but $\mathbb{C}$ is not. Thus $\mathbb{C} \neq S^2$. Show that a set, K is compact (connected) in $\mathbb{C}$ if and only if $\theta^{-1}(K)$ is compact (connected) in $S^2 \setminus \{(0,0,2)\}$.

4. Let K be a compact set in $\mathbb{C}$. Show that $\mathbb{C} \setminus K$ has exactly one unbounded component and that this component is the one which is a subset of the component of $S^2 \setminus K$ which contains ∞. If you need to rewrite using the mapping, θ to make sense of this, it is fine to do so.

5. Make $\widehat{\mathbb{C}}$ into a topological space as follows. We define a basis for a topology on $\widehat{\mathbb{C}}$ to be all open sets and all complements of compact sets, the latter type being those which are said to contain the point ∞. Show this is a basis for a topology which makes $\widehat{\mathbb{C}}$ into a compact Hausdorff space. Also verify that $\widehat{\mathbb{C}}$ with this topology is homeomorphic to the sphere, S^2.

Riemann Stieltjes integrals

In the theory of functions of a complex variable, the most important results are those involving contour integration. Before we define what we mean by contour integration, it is necessary to define the notion of a Riemann Steiltjes integral, a generalization of the usual Riemann integral and the notion of a function of bounded variation.

Definition 22.1 *Let $\gamma : [a, b] \to \mathbb{C}$ be a function. We say γ is of bounded variation if*

$$\sup \left\{ \sum_{i=1}^{n} |\gamma(t_i) - \gamma(t_{i-1})| : a = t_0 < \cdots < t_n = b \right\} \equiv V(\gamma, [a, b]) < \infty$$

where the sums are taken over all possible lists, $\{a = t_0 < \cdots < t_n = b\}$.

The idea is that it makes sense to talk of the length of the curve $\gamma([a, b])$, defined as $V(\gamma, [a, b])$. For this reason, in the case that γ is continuous, such an image of a bounded variation function is called a rectifiable curve.

Definition 22.2 *Let $\gamma : [a, b] \to \mathbb{C}$ be of bounded variation and let $f : [a, b] \to \mathbb{C}$. Letting $\mathcal{P} \equiv \{t_0, \cdots, t_n\}$ where $a = t_0 < t_1 < \cdots < t_n = b$, we define*

$$||\mathcal{P}|| \equiv \max\{|t_j - t_{j-1}| : j = 1, \cdots, n\}$$

and the Riemann Steiltjes sum by

$$S(\mathcal{P}) \equiv \sum_{j=1}^{n} f(\tau_j)(\gamma(t_j) - \gamma(t_{j-1}))$$

where $\tau_j \in [t_{j-1}, t_j]$. (Note this notation is a little sloppy because it does not identify the specific point, τ_j used. It is understood that this point is arbitrary.) We define $\int_\gamma f(t)\, d\gamma(t)$ as the unique number which satisfies the following condition. For all $\varepsilon > 0$ there exists a $\delta > 0$ such that if $||\mathcal{P}|| \leq \delta$, then

$$\left| \int_\gamma f(t)\, d\gamma(t) - S(\mathcal{P}) \right| < \varepsilon.$$

Sometimes this is written as

$$\int_\gamma f(t)\, d\gamma(t) \equiv \lim_{||\mathcal{P}|| \to 0} S(\mathcal{P}).$$

The function, $\gamma([a, b])$ is a set of points in $\mathbb{C}$ and as t moves from a to b, $\gamma(t)$ moves from $\gamma(a)$ to $\gamma(b)$. Thus $\gamma([a, b])$ has a first point and a last point. If $\phi : [c, d] \to [a, b]$ is a continuous nondecreasing function, then $\gamma \circ \phi : [c, d] \to \mathbb{C}$ is also of bounded variation and yields the same set of points in $\mathbb{C}$ with the same first and last points. In the case where the values of the function, f, which are of interest are those on $\gamma([a, b])$, we have the following important theorem on change of parameters.

Theorem 22.3 *Let ϕ and γ be as just described. Then assuming that*

$$\int_\gamma f(\gamma(t))\, d\gamma(t)$$

exists, so does

$$\int_{\gamma\circ\phi} f(\gamma(\phi(s)))\, d(\gamma\circ\phi)(s)$$

and

$$\int_\gamma f(\gamma(t))\, d\gamma(t) = \int_{\gamma\circ\phi} f(\gamma(\phi(s)))\, d(\gamma\circ\phi)(s). \tag{22.1}$$

Proof: There exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of $[a,b]$ such that $||\mathcal{P}|| < \delta$, then

$$\left|\int_\gamma f(\gamma(t))\, d\gamma(t) - S(\mathcal{P})\right| < \varepsilon.$$

By continuity of ϕ, there exists $\sigma > 0$ such that if $\mathcal{Q}$ is a partition of $[c,d]$ with $||\mathcal{Q}|| < \sigma, \mathcal{Q} = \{s_0, \cdots, s_n\}$, then $|\phi(s_j) - \phi(s_{j-1})| < \delta$. Thus letting $\mathcal{P}$ denote the points in $[a,b]$ given by $\phi(s_j)$ for $s_j \in \mathcal{Q}$, it follows that $||\mathcal{P}|| < \delta$ and so

$$\left|\int_\gamma f(\gamma(t))\, d\gamma(t) - \sum_{j=1}^n f(\gamma(\phi(\tau_j)))(\gamma(\phi(s_j)) - \gamma(\phi(s_{j-1})))\right| < \varepsilon$$

where $\tau_j \in [s_{j-1}, s_j]$. Therefore, from the definition we see that (22.1) holds and that

$$\int_{\gamma\circ\phi} f(\gamma(\phi(s)))\, d(\gamma\circ\phi)(s)$$

exists.

This theorem shows that $\int_\gamma f(\gamma(t))\, d\gamma(t)$ is independent of the particular γ used in its computation to the extent that if ϕ is any nondecreasing function from another interval, $[c,d]$, mapping to $[a,b]$, then the same value is obtained by replacing γ with $\gamma\circ\phi$.

The fundamental result in this subject is the following theorem.

Theorem 22.4 *Let $f : [a,b] \to \mathbb{C}$ be continuous and let $\gamma : [a,b] \to \mathbb{C}$ be of bounded variation. Then $\int_\gamma f(t)\, d\gamma(t)$ exists. Also if $\delta_m > 0$ is such that $|t-s| < \delta_m$ implies $|f(t) - f(s)| < \frac{1}{m}$, then*

$$\left|\int_\gamma f(t)\, d\gamma(t) - S(\mathcal{P})\right| \leq \frac{2V(\gamma,[a,b])}{m}$$

whenever $||\mathcal{P}|| < \delta_m$.

Proof: The function, f, is uniformly continuous because it is defined on a compact set. Therefore, there exists a decreasing sequence of positive numbers, $\{\delta_m\}$ such that if $|s-t| < \delta_m$, then

$$|f(t) - f(s)| < \frac{1}{m}.$$

Let

$$F_m \equiv \overline{\{S(\mathcal{P}) : ||\mathcal{P}|| < \delta_m\}}.$$

Thus F_m is a closed set. (When we write $S(\mathcal{P})$ in the above definition, we mean to include all sums corresponding to $\mathcal{P}$ for any choice of τ_j.) We wish to show that

$$diam\,(F_m) \leq \frac{2V(\gamma,[a,b])}{m} \tag{22.2}$$

because then there will exist a unique point, $I \in \cap_{m=1}^{\infty} F_m$. It will then follow that $I = \int_\gamma f(t)\,d\gamma(t)$. To verify (22.2), it suffices to verify that whenever $\mathcal{P}$ and $\mathcal{Q}$ are partitions satisfying $||\mathcal{P}|| < \delta_m$ and $||\mathcal{Q}|| < \delta_m$,

$$|S(\mathcal{P}) - S(\mathcal{Q})| \leq \frac{2}{m} V(\gamma,[a,b]). \tag{22.3}$$

Suppose $||\mathcal{P}|| < \delta_m$ and $\mathcal{Q} \supseteq \mathcal{P}$. Then also $||\mathcal{Q}|| < \delta_m$. To begin with, suppose that $\mathcal{P} \equiv \{t_0,\cdots,t_p,\cdots,t_n\}$ and $\mathcal{Q} \equiv \{t_0,\cdots,t_{p-1},t^*,t_p,\cdots,t_n\}$. Thus $\mathcal{Q}$ contains only one more point than $\mathcal{P}$. Letting $S(\mathcal{Q})$ and $S(\mathcal{P})$ be Riemann Steiltjes sums,

$$S(\mathcal{Q}) \equiv \sum_{j=1}^{p-1} f(\sigma_j)(\gamma(t_j) - \gamma(t_{j-1})) + f(\sigma_*)(\gamma(t^*) - \gamma(t_{p-1}))$$

$$+ f(\sigma^*)(\gamma(t_p) - \gamma(t^*)) + \sum_{j=p+1}^{n} f(\sigma_j)(\gamma(t_j) - \gamma(t_{j-1})),$$

$$S(\mathcal{P}) \equiv \sum_{j=1}^{p-1} f(\tau_j)(\gamma(t_j) - \gamma(t_{j-1})) + f(\tau_p)(\gamma(t^*) - \gamma(t_{p-1}))$$

$$+ f(\tau_p)(\gamma(t_p) - \gamma(t^*)) + \sum_{j=p+1}^{n} f(\tau_j)(\gamma(t_j) - \gamma(t_{j-1})).$$

Therefore,

$$|S(\mathcal{P}) - S(\mathcal{Q})| \leq \sum_{j=1}^{p-1} \frac{1}{m} |\gamma(t_j) - \gamma(t_{j-1})| + \frac{1}{m} |\gamma(t^*) - \gamma(t_{p-1})| +$$

$$\frac{1}{m} |\gamma(t_p) - \gamma(t^*)| + \sum_{j=p+1}^{n} \frac{1}{m} |\gamma(t_j) - \gamma(t_{j-1})| \leq \frac{1}{m} V(\gamma,[a,b]). \tag{22.4}$$

Clearly the extreme inequalities would be valid in (22.4) if $\mathcal{Q}$ had more than one extra point. Let $S(\mathcal{P})$ and $S(\mathcal{Q})$ be Riemann Steiltjes sums for which $||\mathcal{P}||$ and $||\mathcal{Q}||$ are less than δ_m and let $\mathcal{R} \equiv \mathcal{P} \cup \mathcal{Q}$. Then

$$|S(\mathcal{P}) - S(\mathcal{Q})| \leq |S(\mathcal{P}) - S(\mathcal{R})| + |S(\mathcal{R}) - S(\mathcal{Q})| \leq \frac{2}{m} V(\gamma,[a,b]).$$

and this shows (22.3) which proves (22.2). Therefore, there exists a unique complex number, $I \in \cap_{m=1}^{\infty} F_m$ which satisfies the definition of $\int_\gamma f(t)\,d\gamma(t)$. This proves the theorem.

The following theorem follows easily from the above definitions and theorem.

Theorem 22.5 *Let $f \in C([a,b])$ and let $\gamma : [a,b] \to \mathbb{C}$ be of bounded variation. Let*

$$M \geq \max\{|f(t)| : t \in [a,b]\}. \tag{22.5}$$

Then

$$\left|\int_\gamma f(t)\, d\gamma(t)\right| \leq MV(\gamma,[a,b]). \tag{22.6}$$

Also if $\{f_n\}$ is a sequence of functions of $C([a,b])$ which is converging uniformly to the function, f, then

$$\lim_{n\to\infty}\int_\gamma f_n(t)\, d\gamma(t) = \int_\gamma f(t)\, d\gamma(t). \tag{22.7}$$

Proof: Let (22.5) hold. From the proof of the above theorem we know that when $||\mathcal{P}|| < \delta_m$,

$$\left|\int_\gamma f(t)\, d\gamma(t) - S(\mathcal{P})\right| \leq \frac{2}{m}V(\gamma,[a,b])$$

and so

$$\left|\int_\gamma f(t)\, d\gamma(t)\right| \leq |S(\mathcal{P})| + \frac{2}{m}V(\gamma,[a,b])$$

$$\begin{aligned} &\leq \sum_{j=1}^n M|\gamma(t_j) - \gamma(t_{j-1})| + \frac{2}{m}V(\gamma,[a,b]) \\ &\leq MV(\gamma,[a,b]) + \frac{2}{m}V(\gamma,[a,b]). \end{aligned}$$

This proves (22.6) since m is arbitrary. To verify (22.7) we use the above inequality to write

$$\left|\int_\gamma f(t)\, d\gamma(t) - \int_\gamma f_n(t)\, d\gamma(t)\right| = \left|\int_\gamma (f(t) - f_n(t))\, d\gamma(t)\right|$$

$$\leq \max\{|f(t) - f_n(t)| : t \in [a,b]\} V(\gamma,[a,b]).$$

Since the convergence is assumed to be uniform, this proves (22.7).

It turns out that we will be mainly interested in the case where γ is also continuous in addition to being of bounded variation. Also, it turns out to be much easier to evaluate such integrals in the case where γ is also $C^1([a,b])$. The following theorem about approximation will be very useful.

Theorem 22.6 *Let $\gamma : [a,b] \to \mathbb{C}$ be continuous and of bounded variation, let $f : [a,b] \times K \to \mathbb{C}$ be continuous for K a compact set in $\mathbb{C}$, and let $\varepsilon > 0$ be given. Then there exists $\eta : [a,b] \to \mathbb{C}$ such that $\eta(a) = \gamma(a)$, $\gamma(b) = \eta(b)$, $\eta \in C^1([a,b])$, and*

$$||\gamma - \eta|| < \varepsilon, \tag{22.8}$$

$$\left|\int_\gamma f(t,z)\, d\gamma(t) - \int_\eta f(t,z)\, d\eta(t)\right| < \varepsilon, \tag{22.9}$$

$$V(\eta,[a,b]) \leq V(\gamma,[a,b]), \tag{22.10}$$

where $||\gamma - \eta|| \equiv \max\{|\gamma(t) - \eta(t)| : t \in [a,b]\}$.

Proof: We extend γ to be defined on all $\mathbb{R}$ according to $\gamma(t) = \gamma(a)$ if $t < a$ and $\gamma(t) = \gamma(b)$ if $t > b$. Now we define

$$\gamma_h(t) \equiv \frac{1}{2h} \int_{-2h+t+\frac{2h}{(b-a)}(t-a)}^{t+\frac{2h}{(b-a)}(t-a)} \gamma(s)\, ds.$$

where the integral is defined in the obvious way. That is,

$$\int_a^b \alpha(t) + i\beta(t)\, dt \equiv \int_a^b \alpha(t)\, dt + i \int_a^b \beta(t)\, dt.$$

Therefore,

$$\gamma_h(b) = \frac{1}{2h} \int_b^{b+2h} \gamma(s)\, ds = \gamma(b),$$

$$\gamma_h(a) = \frac{1}{2h} \int_{a-2h}^{a} \gamma(s)\, ds = \gamma(a).$$

Also, because of continuity of γ and the fundamental theorem of calculus,

$$\gamma_h'(t) = \frac{1}{2h} \left\{ \gamma\left(t + \frac{2h}{b-a}(t-a)\right)\left(1 + \frac{2h}{b-a}\right) - \right.$$

$$\left. \gamma\left(-2h + t + \frac{2h}{b-a}(t-a)\right)\left(1 + \frac{2h}{b-a}\right) \right\}$$

and so $\gamma_h \in C^1([a,b])$. The following lemma is significant.

Lemma 22.7 $V(\gamma_h, [a,b]) \leq V(\gamma, [a,b])$.

Proof: Let $a = t_0 < t_1 < \cdots < t_n = b$. Then using the definition of γ_h and changing the variables to make all integrals over $[0, 2h]$,

$$\sum_{j=1}^{n} |\gamma_h(t_j) - \gamma_h(t_{j-1})| =$$

$$\sum_{j=1}^{n} \left| \frac{1}{2h} \int_0^{2h} \left[\gamma\left(s - 2h + t_j + \frac{2h}{b-a}(t_j - a)\right) - \right.\right.$$

$$\left.\left. \gamma\left(s - 2h + t_{j-1} + \frac{2h}{b-a}(t_{j-1} - a)\right) \right] \right|$$

$$\leq \frac{1}{2h} \int_0^{2h} \sum_{j=1}^{n} \left| \gamma\left(s - 2h + t_j + \frac{2h}{b-a}(t_j - a)\right) - \right.$$

$$\left. \gamma\left(s - 2h + t_{j-1} + \frac{2h}{b-a}(t_{j-1} - a)\right) \right| ds.$$

For a given $s \in [0, 2h]$, the points, $s - 2h + t_j + \frac{2h}{b-a}(t_j - a)$ for $j = 1, \cdots, n$ form an increasing list of points in the interval $[a - 2h, b + 2h]$ and so the integrand is bounded above by $V(\gamma, [a - 2h, b + 2h]) = V(\gamma, [a, b])$. It follows

$$\sum_{j=1}^{n} |\gamma_h(t_j) - \gamma_h(t_{j-1})| \leq V(\gamma, [a, b])$$

which proves the lemma.

With this lemma the proof of the theorem can be completed without too much trouble. First of all, if $\varepsilon > 0$ is given, there exists δ_1 such that if $h < \delta_1$, then for all t,

$$\begin{aligned} |\gamma(t) - \gamma_h(t)| &\leq \frac{1}{2h} \int_{-2h+t+\frac{2h}{(b-a)}(t-a)}^{t+\frac{2h}{(b-a)}(t-a)} |\gamma(s) - \gamma(t)| \, ds \\ &< \frac{1}{2h} \int_{-2h+t+\frac{2h}{(b-a)}(t-a)}^{t+\frac{2h}{(b-a)}(t-a)} \varepsilon ds = \varepsilon \end{aligned} \tag{22.11}$$

due to the uniform continuity of γ. This proves (22.8). From (22.2) there exists δ_2 such that if $||\mathcal{P}|| < \delta_2$, then for all $z \in K$,

$$\left| \int_{\gamma} f(t, z) \, d\gamma(t) - S(\mathcal{P}) \right| < \frac{\varepsilon}{3}, \left| \int_{\gamma_h} f(t, z) \, d\gamma_h(t) - S_h(\mathcal{P}) \right| < \frac{\varepsilon}{3}$$

for all h. Here $S(\mathcal{P})$ is a Riemann Steiltjes sum of the form

$$\sum_{i=1}^{n} f(\tau_i, z) (\gamma(t_i) - \gamma(t_{i-1}))$$

and $S_h(\mathcal{P})$ is a similar Riemann Steiltjes sum taken with respect to γ_h instead of γ. Therefore, fix the partition, $\mathcal{P}$, and choose h small enough that in addition to this, we have the following inequality valid for all $z \in K$.

$$|S(\mathcal{P}) - S_h(\mathcal{P})| < \frac{\varepsilon}{3}$$

We can do this thanks to (22.11) and the uniform continuity of f on $[a, b] \times K$. It follows

$$\left| \int_{\gamma} f(t, z) \, d\gamma(t) - \int_{\gamma_h} f(t, z) \, d\gamma_h(t) \right| \leq$$

$$\left| \int_{\gamma} f(t, z) \, d\gamma(t) - S(\mathcal{P}) \right| + |S(\mathcal{P}) - S_h(\mathcal{P})|$$

$$+ \left| S_h(\mathcal{P}) - \int_{\gamma_h} f(t, z) \, d\gamma_h(t) \right| < \varepsilon.$$

Formula (22.10) follows from the lemma. This proves the theorem.

Of course the same result is obtained without the explicit dependence of f on z.

This is a very useful theorem because if γ is $C^1([a, b])$, it is easy to calculate $\int_{\gamma} f(t) \, d\gamma(t)$. We will typically reduce to the case where γ is C^1 by using the above theorem. The next theorem shows how easy it is to compute these integrals in the case where γ is C^1.

Theorem 22.8 *If $f : [a,b] \to \mathbb{C}$ and $\gamma : [a,b] \to \mathbb{C}$ is in $C^1([a,b])$, then*

$$\int_\gamma f(t)\, d\gamma(t) = \int_a^b f(t)\, \gamma'(t)\, dt. \tag{22.12}$$

Proof: Let $\mathcal{P}$ be a partition of $[a,b]$, $\mathcal{P} = \{t_0, \cdots, t_n\}$ and $||\mathcal{P}||$ is small enough that whenever $|t-s| < ||\mathcal{P}||$,

$$|f(t) - f(s)| < \varepsilon \tag{22.13}$$

and

$$\left| \int_\gamma f(t)\, d\gamma(t) - \sum_{j=1}^n f(\tau_j)\left(\gamma(t_j) - \gamma(t_{j-1})\right) \right| < \varepsilon.$$

Now

$$\sum_{j=1}^n f(\tau_j)\left(\gamma(t_j) - \gamma(t_{j-1})\right) = \int_a^b \sum_{j=1}^n f(\tau_j)\, \mathcal{X}_{(t_{j-1},t_j]}(s)\, \gamma'(s)\, ds$$

and thanks to (22.13),

$$\left| \int_a^b \sum_{j=1}^n f(\tau_j)\, \mathcal{X}_{(t_{j-1},t_j]}(s)\, \gamma'(s)\, ds - \int_a^b f(s)\, \gamma'(s)\, ds \right|$$

$$< \int_a^b \varepsilon\, |\gamma'(s)|\, ds.$$

It follows that

$$\left| \int_\gamma f(t)\, d\gamma(t) - \int_a^b f(s)\, \gamma'(s)\, ds \right| < \varepsilon \int_a^b |\gamma'(s)|\, ds + \varepsilon.$$

Since ε is arbitrary, this verifies (22.12).

Definition 22.9 *Let $\gamma : [a,b] \to U$ be a continuous function with bounded variation and let $f : U \to \mathbb{C}$ be a continuous function. Then we define,*

$$\int_\gamma f(z)\, dz \equiv \int_\gamma f(\gamma(t))\, d\gamma(t).$$

The expression, $\int_\gamma f(z)\, dz$, is called a contour integral and γ is referred to as the contour. We also say that a function $f : U \to \mathbb{C}$ for U an open set in $\mathbb{C}$ has a primitive if there exists a function, F, the primitive, such that $F'(z) = f(z)$. Thus F is just an antiderivative. Also if $\gamma_k : [a_k, b_k] \to \mathbb{C}$ is continuous and of bounded variation, for $k = 1, \cdots, m$ and $\gamma_k(b_k) = \gamma_{k+1}(a_k)$, we define

$$\int_{\sum_{k=1}^m \gamma_k} f(z)\, dz \equiv \sum_{k=1}^m \int_{\gamma_k} f(z)\, dz. \tag{22.14}$$

In addition to this, for $\gamma : [a,b] \to \mathbb{C}$, we define $-\gamma : [a,b] \to \mathbb{C}$ by $-\gamma(t) \equiv \gamma(b+a-t)$. Thus γ simply traces out the points of $\gamma([a,b])$ in the opposite order.

The following lemma is useful and follows quickly from Theorem 22.3.

Lemma 22.10 *In the above definition, there exists a continuous bounded variation function, γ defined on some closed interval, $[c,d]$, such that $\gamma([c,d]) = \cup_{k=1}^{m} \gamma_k([a_k,b_k])$ and $\gamma(c) = \gamma_1(a_1)$ while $\gamma(d) = \gamma_m(b_m)$. Furthermore,*

$$\int_\gamma f(z)\,dz = \sum_{k=1}^{m} \int_{\gamma_k} f(z)\,dz.$$

If $\gamma : [a,b] \to \mathbb{C}$ is of bounded variation and continuous, then

$$\int_\gamma f(z)\,dz = -\int_{-\gamma} f(z)\,dz.$$

Theorem 22.11 *Let K be a compact set in $\mathbb{C}$ and let $f : U \times K \to \mathbb{C}$ be continuous for U an open set in $\mathbb{C}$. Also let $\gamma : [a,b] \to U$ be continuous with bounded variation. Then if $r > 0$ is given, there exists $\eta : [a,b] \to U$ such that $\eta(a) = \gamma(a)$, $\eta(b) = \gamma(b)$, η is $C^1([a,b])$, and*

$$\left| \int_\gamma f(z,w)\,dz - \int_\eta f(z,w)\,dz \right| < r, \ ||\eta - \gamma|| < r.$$

Proof: Let $\varepsilon > 0$ be given and let H be an open set containing $\gamma([a,b])$ such that $\overline{H}$ is compact. Then f is uniformly continuous on $\overline{H} \times K$ and so there exists a $\delta > 0$ such that if $z_j \in H, j = 1,2$ and $w_j \in K$ for $j = 1,2$ such that if

$$|z_1 - z_2| + |w_1 - w_2| < \delta,$$

then

$$|f(z_1,w_1) - f(z_2,w_2)| < \varepsilon.$$

By Theorem 22.6, let $\eta : [a,b] \to \mathbb{C}$ be such that $\eta([a,b]) \subseteq H$, $\eta(x) = \gamma(x)$ for $x = a,b$, $\eta \in C^1([a,b])$, $||\eta - \gamma|| < \min(\delta, r)$, $V(\eta,[a,b]) < V(\gamma,[a,b])$, and

$$\left| \int_\eta f(\gamma(t),w)\,d\eta(t) - \int_\gamma f(\gamma(t),w)\,d\gamma(t) \right| < \varepsilon$$

for all $w \in K$. Then, since $|f(\gamma(t),w) - f(\eta(t),w)| < \varepsilon$ for all $t \in [a,b]$,

$$\left| \int_\eta f(\gamma(t),w)\,d\eta(t) - \int_\eta f(\eta(t),w)\,d\eta(t) \right| < \varepsilon V(\eta,[a,b]) \le \varepsilon V(\gamma,[a,b]).$$

Therefore,

$$\left| \int_\eta f(z,w)\,dz - \int_\gamma f(z,w)\,dz \right| =$$

$$\left| \int_\eta f(\eta(t),w)\,d\eta(t) - \int_\gamma f(\gamma(t),w)\,d\gamma(t) \right| < \varepsilon + \varepsilon V(\gamma,[a,b]).$$

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

We will be very interested in the functions which have primitives. It turns out, it is not enough for f to be continuous in order to possess a primitive. This is in stark contrast to the situation for functions of a real variable in which the fundamental theorem of calculus will deliver a primitive for any continuous function. The reason for our interest in such functions is the following theorem and its corollary.

Theorem 22.12 *Let $\gamma : [a,b] \to \mathbb{C}$ be continuous and of bounded variation. Also suppose $F'(z) = f(z)$ for all $z \in U$, an open set containing $\gamma([a,b])$ and f is continuous on U. Then*

$$\int_\gamma f(z)\,dz = F(\gamma(b)) - F(\gamma(a)).$$

Proof: By Theorem 22.11 there exists $\eta \in C^1([a,b])$ such that $\gamma(a) = \eta(a)$, and $\gamma(b) = \eta(b)$ such that

$$\left| \int_\gamma f(z)\,dz - \int_\eta f(z)\,dz \right| < \varepsilon.$$

Then since η is in $C^1([a,b])$, we may write

$$\begin{aligned} \int_\eta f(z)\,dz &= \int_a^b f(\eta(t))\,\eta'(t)\,dt = \int_a^b \frac{dF(\eta(t))}{dt}dt \\ &= F(\eta(b)) - F(\eta(a)) = F(\gamma(b)) - F(\gamma(a)). \end{aligned}$$

Therefore,

$$\left| (F(\gamma(b)) - F(\gamma(a))) - \int_\gamma f(z)\,dz \right| < \varepsilon$$

and since $\varepsilon > 0$ is arbitrary, this proves the theorem.

Corollary 22.13 *If $\gamma : [a,b] \to \mathbb{C}$ is continuous, has bounded variation, is a closed curve, $\gamma(a) = \gamma(b)$, and $\gamma([a,b]) \subseteq U$ where U is an open set on which $F'(z) = f(z)$, then*

$$\int_\gamma f(z)\,dz = 0.$$

22.1 Exercises

1. Let $\gamma : [a,b] \to \mathbb{R}$ be increasing. Show $V(\gamma, [a,b]) = \gamma(b) - \gamma(a)$.
2. Suppose $\gamma : [a,b] \to \mathbb{C}$ satisfies a Lipschitz condition, $|\gamma(t) - \gamma(s)| \le K|s-t|$. Show γ is of bounded variation and that $V(\gamma, [a,b]) \le K|b-a|$.
3. We say $\gamma : [c_0, c_m] \to \mathbb{C}$ is piecewise smooth if there exist numbers, $c_k, k = 1, \cdots, m$ such that $c_0 < c_1 < \cdots < c_{m-1} < c_m$ such that γ is continuous and $\gamma : [c_k, c_{k+1}] \to \mathbb{C}$ is C^1. Show that such piecewise smooth functions are of bounded variation and give an estimate for $V(\gamma, [c_0, c_m])$.
4. Let $\gamma : [0, 2\pi] \to \mathbb{C}$ be given by $\gamma(t) = r(\cos mt + i \sin mt)$ for m an integer. Find $\int_\gamma \frac{dz}{z}$.
5. Show that if $\gamma : [a,b] \to \mathbb{C}$ then there exists an increasing function $h : [0,1] \to [a,b]$ such that $\gamma \circ h([0,1]) = \gamma([a,b])$.
6. Let $\gamma : [a,b] \to \mathbb{C}$ be an arbitrary continuous curve having bounded variation and let f, g have continuous derivatives on some open set containing $\gamma([a,b])$. Prove the usual integration by parts formula.
$$\int_\gamma f g' dz = f(\gamma(b))\,g(\gamma(b)) - f(\gamma(a))\,g(\gamma(a)) - \int_\gamma f' g dz.$$
7. Let $f(z) \equiv |z|^{-(1/2)} e^{-i\frac{\theta}{2}}$ where $z = |z|e^{i\theta}$. This function is called the principle branch of $z^{-(1/2)}$. Find $\int_\gamma f(z)\,dz$ where γ is the semicircle in the upper half plane which goes from $(1,0)$ to $(-1,0)$ in the counter clockwise direction. Next do the integral in which γ goes in the clockwise direction along the semicircle in the lower half plane.

8. Prove an open set, U is connected if and only if for every two points in U, there exists a C^1 curve having values in U which joins them.

9. Let $\mathcal{P}, \mathcal{Q}$ be two partitions of $[a,b]$ with $\mathcal{P} \subseteq \mathcal{Q}$. Each of these partitions can be used to form an approximation to $V(\gamma, [a,b])$ as described above. Recall the total variation was the supremum of sums of a certain form determined by a partition. How is the sum associated with $\mathcal{P}$ related to the sum associated with $\mathcal{Q}$? Explain.

10. Consider the curve,

$$\gamma(t) = \begin{cases} t + it^2 \sin\left(\frac{1}{t}\right) \text{ if } t \in (0,1] \\ 0 \text{ if } t = 0 \end{cases}.$$

Is γ a continuous curve having bounded variation? What if the t^2 is replaced with t? Is the resulting curve continuous? Is it a bounded variation curve?

11. Suppose $\gamma : [a,b] \to \mathbb{R}$ is given by $\gamma(t) = t$. What is $\int_\gamma f(t)\, d\gamma$? Explain.

Analytic functions

In this chapter we define what we mean by an analytic function and give a few important examples of functions which are analytic.

Definition 23.1 *Let U be an open set in $\mathbb{C}$ and let $f : U \to \mathbb{C}$. We say f is analytic on U if for every $z \in U$,*

$$\lim_{h \to 0} \frac{f(z+h) - f(z)}{h} \equiv f'(z)$$

exists and is a continuous function of $z \in U$. Here $h \in \mathbb{C}$.

Note that if f is analytic, it must be the case that f is continuous. It is more common to not include the requirement that f' is continuous but we will show later that the continuity of f' follows.

What are some examples of analytic functions? The simplest example is any polynomial. Thus

$$p(z) \equiv \sum_{k=0}^{n} a_k z^k$$

is an analytic function and

$$p'(z) = \sum_{k=1}^{n} a_k k z^{k-1}.$$

We leave the verification of this as an exercise. More generally, power series are analytic. We will show this later. For now, we consider the very important Cauchy Riemann equations which give conditions under which complex valued functions of a complex variable are analytic.

Theorem 23.2 *Let U be an open subset of $\mathbb{C}$ and let $f : U \to \mathbb{C}$ be a function, such that for $z = x + iy \in U$,*

$$f(z) = u(x,y) + iv(x,y).$$

Then f is analytic if and only if u, v are $C^1(U)$ and

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

Furthermore, we have the formula,

$$f'(z) = \frac{\partial u}{\partial x}(x,y) + i\frac{\partial v}{\partial x}(x,y).$$

Proof: Suppose f is analytic first. Then letting $t \in \mathbb{R}$,

$$f'(z) = \lim_{t \to 0} \frac{f(z+t) - f(z)}{t} =$$

$$\lim_{t \to 0} \left(\frac{u(x+t,y) + iv(x+t,y)}{t} - \frac{u(x,y) + iv(x,y)}{t} \right)$$

$$= \frac{\partial u(x,y)}{\partial x} + i \frac{\partial v(x,y)}{\partial x}.$$

But also

$$f'(z) = \lim_{t \to 0} \frac{f(z+it) - f(z)}{it} =$$

$$\lim_{t \to 0} \left(\frac{u(x,y+t) + iv(x,y+t)}{it} - \frac{u(x,y) + iv(x,y)}{it} \right)$$

$$\frac{1}{i} \left(\frac{\partial u(x,y)}{\partial y} + i \frac{\partial v(x,y)}{\partial y} \right)$$

$$= \frac{\partial v(x,y)}{\partial y} - i \frac{\partial u(x,y)}{\partial y}.$$

This verifies the Cauchy Riemann equations. We are assuming that $z \to f'(z)$ is continuous. Therefore, the partial derivatives of u and v are also continuous. To see this, note that from the formulas for $f'(z)$ given above, and letting $z_1 = x_1 + iy_1$

$$\left| \frac{\partial v(x,y)}{\partial y} - \frac{\partial v(x_1,y_1)}{\partial y} \right| \leq |f'(z) - f'(z_1)|,$$

showing that $(x,y) \to \frac{\partial v(x,y)}{\partial y}$ is continuous since $(x_1,y_1) \to (x,y)$ if and only if $z_1 \to z$. The other cases are similar.

Now suppose the Cauchy Riemann equations hold and the functions, u and v are $C^1(U)$. Then letting $h = h_1 + ih_2$,

$$f(z+h) - f(z) = u(x+h_1, y+h_2)$$

$$+iv(x+h_1, y+h_2) - (u(x,y) + iv(x,y))$$

We know u and v are both differentiable and so

$$f(z+h) - f(z) = \frac{\partial u}{\partial x}(x,y) h_1 + \frac{\partial u}{\partial y}(x,y) h_2 +$$

$$i \left(\frac{\partial v}{\partial x}(x,y) h_1 + \frac{\partial v}{\partial y}(x,y) h_2 \right) + o(h).$$

Dividing by h and using the Cauchy Riemann equations,

$$\frac{f(z+h)-f(z)}{h}=\frac{\frac{\partial u}{\partial x}(x,y)\,h_1+i\frac{\partial v}{\partial y}(x,y)\,h_2}{h}+$$

$$\frac{i\frac{\partial v}{\partial x}(x,y)\,h_1+\frac{\partial u}{\partial y}(x,y)\,h_2}{h}+\frac{o(h)}{h}$$

$$=\frac{\partial u}{\partial x}(x,y)\,\frac{h_1+ih_2}{h}+i\frac{\partial v}{\partial x}(x,y)\,\frac{h_1+ih_2}{h}+\frac{o(h)}{h}$$

Taking the limit as $h\to 0$, we obtain

$$f'(z)=\frac{\partial u}{\partial x}(x,y)+i\frac{\partial v}{\partial x}(x,y).$$

It follows from this formula and the assumption that u,v are $C^1(U)$ that f' is continuous.

It is routine to verify that all the usual rules of derivatives hold for analytic functions. In particular, we have the product rule, the chain rule, and quotient rule.

23.1 Exercises

1. Verify all the usual rules of differentiation including the product and chain rules.

2. Suppose f and $f':U\to\mathbb{C}$ are analytic and $f(z)=u(x,y)+iv(x,y)$. Verify $u_{xx}+u_{yy}=0$ and $v_{xx}+v_{yy}=0$. This partial differential equation satisfied by the real and imaginary parts of an analytic function is called Laplace's equation. We say these functions satisfying Laplace's equation are harmonic functions. If u is a harmonic function defined on $B(0,r)$ show that $v(x,y)\equiv\int_0^y u_x(x,t)\,dt-\int_0^x u_y(t,0)\,dt$ is such that $u+iv$ is analytic.

3. Define a function $f(z)\equiv\overline{z}\equiv x-iy$ where $z=x+iy$. Is f analytic?

4. If $f(z)=u(x,y)+iv(x,y)$ and f is analytic, verify that

$$\det\begin{pmatrix}u_x & u_y\\ v_x & v_y\end{pmatrix}=|f'(z)|^2.$$

5. Show that if $u(x,y)+iv(x,y)=f(z)$ is analytic, then $\nabla u\cdot\nabla v=0$. Recall

$$\nabla u(x,y)=\langle u_x(x,y),u_y(x,y)\rangle.$$

6. Show that every polynomial is analytic.

7. If $\gamma(t)=x(t)+iy(t)$ is a C^1 curve having values in U, an open set of $\mathbb{C}$, and if $f:U\to\mathbb{C}$ is analytic, we can consider $f\circ\gamma$, another C^1 curve having values in $\mathbb{C}$. Also, $\gamma'(t)$ and $(f\circ\gamma)'(t)$ are complex numbers so these can be considered as vectors in $\mathbb{R}^2$ as follows. The complex number, $x+iy$ corresponds to the vector, $\langle x,y\rangle$. Suppose that γ and η are two such C^1 curves having values in U and that $\gamma(t_0)=\eta(s_0)=z$ and suppose that $f:U\to\mathbb{C}$ is analytic. Show that the angle between $(f\circ\gamma)'(t_0)$ and $(f\circ\eta)'(s_0)$ is the same as the angle between $\gamma'(t_0)$ and $\eta'(s_0)$ assuming that $f'(z)\neq 0$. Thus analytic mappings preserve angles at points where the derivative is nonzero. Such mappings are called isogonal. . **Hint:** To make this easy to show, first observe that $\langle x,y\rangle\cdot\langle a,b\rangle=\frac{1}{2}(z\overline{w}+\overline{z}w)$ where $z=x+iy$ and $w=a+ib$.

8. Analytic functions are even better than what is described in Problem 7. In addition to preserving angles, they also preserve orientation. To verify this show that if $z = x + iy$ and $w = a + ib$ are two complex numbers, then $\langle x, y, 0\rangle$ and $\langle a, b, 0\rangle$ are two vectors in $\mathbb{R}^3$. Recall that the cross product, $\langle x, y, 0\rangle \times \langle a, b, 0\rangle$, yields a vector normal to the two given vectors such that the triple, $\langle x, y, 0\rangle, \langle a, b, 0\rangle$, and $\langle x, y, 0\rangle \times \langle a, b, 0\rangle$ satisfies the right hand rule and has magnitude equal to the product of the sine of the included angle times the product of the two norms of the vectors. In this case, the cross product either points in the direction of the positive z axis or in the direction of the negative z axis. Thus, either the vectors $\langle x, y, 0\rangle, \langle a, b, 0\rangle, \mathbf{k}$ form a right handed system or the vectors $\langle a, b, 0\rangle, \langle x, y, 0\rangle, \mathbf{k}$ form a right handed system. These are the two possible orientations. Show that in the situation of Problem 7 the orientation of $\gamma'(t_0), \eta'(s_0), \mathbf{k}$ is the same as the orientation of the vectors $(f \circ \gamma)'(t_0), (f \circ \eta)'(s_0), \mathbf{k}$. Such mappings are called conformal. **Hint:** You can do this by verifying that $(f \circ \gamma)'(t_0) \times (f \circ \eta)'(s_0) = \gamma'(t_0) \times \eta'(s_0)$. To make the verification easier, you might first establish the following simple formula for the cross product where here $x + iy = z$ and $a + ib = w$.
$$\langle x, y, 0\rangle \times \langle a, b, 0\rangle = \operatorname{Re}(zi\overline{w})\,\mathbf{k}.$$

9. Write the Cauchy Riemann equations in terms of polar coordinates. Recall the polar coordinates are given by
$$x = r\cos\theta,\ y = r\sin\theta.$$

23.2 Examples of analytic functions

A very important example of an analytic function is $e^z \equiv e^x(\cos y + i\sin y) \equiv \exp(z)$. We can verify this is an analytic function by considering the Cauchy Riemann equations. Here $u(x, y) = e^x\cos y$ and $v(x, y) = e^x\sin y$. The Cauchy Riemann equations hold and the two functions u and v are $C^1(\mathbb{C})$. Therefore, $z \to e^z$ is an analytic function on all of $\mathbb{C}$. Also from the formula for $f'(z)$ given above for an analytic function,
$$\frac{d}{dz}e^z = e^x(\cos y + i\sin y) = e^z.$$
We also see that $e^z = 1$ if and only if $z = 2\pi k$ for k an integer. Other properties of e^z follow from the formula for it. For example, let $z_j = x_j + iy_j$ where $j = 1, 2$.
$$\begin{aligned} e^{z_1}e^{z_2} &\equiv e^{x_1}(\cos y_1 + i\sin y_1)e^{x_2}(\cos y_2 + i\sin y_2) \\ &= e^{x_1+x_2}(\cos y_1\cos y_2 - \sin y_1\sin y_2) + \\ &\quad ie^{x_1+x_2}(\sin y_1\cos y_2 + \sin y_2\cos y_1) \end{aligned}$$
$$= e^{x_1+x_2}(\cos(y_1 + y_2) + i\sin(y_1 + y_2)) = e^{z_1+z_2}.$$

Another example of an analytic function is any polynomial. We can also define the functions $\cos z$ and $\sin z$ by the usual formulas.
$$\sin z \equiv \frac{e^{iz} - e^{-iz}}{2i}, \cos z \equiv \frac{e^{iz} + e^{-iz}}{2}.$$
By the rules of differentiation, it is clear these are analytic functions which agree with the usual functions in the case where z is real. Also the usual differentiation formulas hold. However,
$$\cos ix = \frac{e^{-x} + e^x}{2} = \cosh x$$
and so $\cos z$ is not bounded. Similarly $\sin z$ is not bounded.

A more interesting example is the log function. We cannot define the log for all values of z but if we leave out the ray, $(-\infty, 0]$, then it turns out we can do so. On $\mathbb{R} + i(-\pi, \pi)$ it is easy to see that e^z is one to one, mapping onto $\mathbb{C} \setminus (-\infty, 0]$. Therefore, we can define the log on $\mathbb{C} \setminus (-\infty, 0]$ in the usual way,

$$e^{\log z} \equiv z = e^{\ln |z|} e^{i \arg(z)},$$

where $\arg(z)$ is the unique angle in $(-\pi, \pi)$ for which the equal sign in the above holds. Thus we need

$$\log z = \ln |z| + i \arg(z). \tag{23.1}$$

There are many other ways to define a logarithm. In fact, we could take any ray from 0 and define a logarithm on what is left. It turns out that all these logarithm functions are analytic. This will be clear from the open mapping theorem presented later but for now you may verify by brute force that the usual definition of the logarithm, given in (23.1) and referred to as the principle branch of the logarithm is analytic. This can be done by verifying the Cauchy Riemann equations in the following.

$$\log z = \ln \left(x^2 + y^2\right)^{1/2} + i \left(-\arccos\left(\frac{x}{\sqrt{x^2 + y^2}}\right)\right) \text{ if } y < 0,$$

$$\log z = \ln \left(x^2 + y^2\right)^{1/2} + i \left(\arccos\left(\frac{x}{\sqrt{x^2 + y^2}}\right)\right) \text{ if } y > 0,$$

$$\log z = \ln \left(x^2 + y^2\right)^{1/2} + i \left(\arctan\left(\frac{y}{x}\right)\right) \text{ if } x > 0.$$

With the principle branch of the logarithm defined, we may define the principle branch of z^α for any $\alpha \in \mathbb{C}$. We define

$$z^\alpha \equiv e^{\alpha \log(z)}.$$

23.3 Exercises

1. Verify the principle branch of the logarithm is an analytic function.
2. Find i^i corresponding to the principle branch of the logarithm.
3. Show that $\sin(z + w) = \sin z \cos w + \cos z \sin w$.
4. If f is analytic on U, an open set in $\mathbb{C}$, when can it be concluded that $|f|$ is analytic? When can it be concluded that $|f|$ is continuous? Prove your assertions.
5. Let $f(z) = \overline{z}$ where $\overline{z} \equiv x - iy$ for $z = x + iy$. Describe geometrically what f does and discuss whether f is analytic.
6. A fractional linear transformation is a function of the form

$$f(z) = \frac{az + b}{cz + d}$$

where $ad - bc \neq 0$. Note that if $c = 0$, this reduces to a linear transformation $(a/d) z + (b/d)$. Special cases of these are given defined as follows.

$$\text{dilations: } z \to \delta z,\ \delta \neq 0,\ \text{inversions: } z \to \frac{1}{z},$$

$$\text{translations: } z \to z + \rho.$$

In the case where $c \neq 0$, let $S_1(z) = z + \frac{d}{c}$, $S_2(z) = \frac{1}{z}$, $S_3(z) = \frac{(bc-ad)}{c^2}z$ and $S_4(z) = z + \frac{a}{c}$. Verify that $f(z) = S_4 \circ S_3 \circ S_2 \circ S_1$. Now show that in the case where $c = 0$, f is still a finite composition of dilations, inversions, and translations.

7. Show that for a fractional linear transformation described in Problem 6 circles and lines are mapped to circles or lines. **Hint:** This is obvious for dilations, and translations. It only remains to verify this for inversions. Note that all circles and lines may be put in the form

$$\alpha\left(x^2 + y^2\right) - 2ax - 2by = r^2 - \left(a^2 + b^2\right)$$

where $\alpha = 1$ gives a circle centered at (a, b) with radius r and $\alpha = 0$ gives a line. In terms of complex variables we may consider all possible circles and lines in the form

$$\alpha z\overline{z} + \beta z + \overline{\beta}\overline{z} + \gamma = 0,$$

Verify every circle or line is of this form and that conversely, every expression of this form yields either a circle or a line. Then verify that inversions do what is claimed.

8. It is desired to find an analytic function, $L(z)$ defined for all $z \in \mathbb{C} \setminus \{0\}$ such that $e^{L(z)} = z$. Is this possible? Explain why or why not.

9. If f is analytic, show that $z \to \overline{f(\overline{z})}$ is also analytic.

10. Find the real and imaginary parts of the principle branch of $z^{1/2}$.

Cauchy's formula for a disk

In this chapter we prove the Cauchy formula for a disk. Later we will generalize this formula to much more general situations but the version given here will suffice to prove many interesting theorems needed in the later development of the theory. First we give a few preliminary results from advanced calculus.

Lemma 24.1 *Let $f : [a,b] \to \mathbb{C}$. Then $f'(t)$ exists if and only if $\operatorname{Re} f'(t)$ and $\operatorname{Im} f'(t)$ exist. Furthermore,*

$$f'(t) = \operatorname{Re} f'(t) + i \operatorname{Im} f'(t).$$

Proof: The if part of the equivalence is obvious.

Now suppose $f'(t)$ exists. Let both t and $t+h$ be contained in $[a,b]$

$$\left| \frac{\operatorname{Re} f(t+h) - \operatorname{Re} f(t)}{h} - \operatorname{Re}(f'(t)) \right| \leq \left| \frac{f(t+h) - f(t)}{h} - f'(t) \right|$$

and this converges to zero as $h \to 0$. Therefore, $\operatorname{Re} f'(t) = \operatorname{Re}(f'(t))$. Similarly, $\operatorname{Im} f'(t) = \operatorname{Im}(f'(t))$.

Lemma 24.2 *If $g : [a,b] \to \mathbb{C}$ and g is continuous on $[a,b]$ and differentiable on (a,b) with $g'(t) = 0$, then $g(t)$ is a constant.*

Proof: From the above lemma, we can apply the mean value theorem to the real and imaginary parts of g.

Lemma 24.3 *Let $\phi : [a,b] \times [c,d] \to \mathbb{R}$ be continuous and let*

$$g(t) \equiv \int_a^b \phi(s,t)\, ds. \tag{24.1}$$

Then g is continuous. If $\frac{\partial \phi}{\partial t}$ exists and is continuous on $[a,b] \times [c,d]$, then

$$g'(t) = \int_a^b \frac{\partial \phi(s,t)}{\partial t} ds. \tag{24.2}$$

Proof: The first claim follows from the uniform continuity of ϕ on $[a,b] \times [c,d]$, which uniform continuity results from the set being compact. To establish (24.2), let t and $t+h$ be contained in $[c,d]$ and form, using the mean value theorem,

$$\begin{aligned} \frac{g(t+h) - g(t)}{h} &= \frac{1}{h} \int_a^b [\phi(s,t+h) - \phi(s,t)]\, ds \\ &= \frac{1}{h} \int_a^b \frac{\partial \phi(s,t+\theta h)}{\partial t} h ds \\ &= \int_a^b \frac{\partial \phi(s,t+\theta h)}{\partial t} ds, \end{aligned}$$

where θ may depend on s but is some number between 0 and 1. Then by the uniform continuity of $\frac{\partial \phi}{\partial t}$, it follows that (24.2) holds.

Corollary 24.4 *Let $\phi : [a,b] \times [c,d] \to \mathbb{C}$ be continuous and let*

$$g(t) \equiv \int_a^b \phi(s,t)\, ds. \tag{24.3}$$

Then g is continuous. If $\frac{\partial \phi}{\partial t}$ exists and is continuous on $[a,b] \times [c,d]$, then

$$g'(t) = \int_a^b \frac{\partial \phi(s,t)}{\partial t} ds. \tag{24.4}$$

Proof: Apply Lemma 24.3 to the real and imaginary parts of ϕ.

With this preparation we are ready to prove Cauchy's formula for a disk.

Theorem 24.5 *Let $f : U \to \mathbb{C}$ be analytic on the open set, U and let*

$$\overline{B(z_0, r)} \subseteq U.$$

Let $\gamma(t) \equiv z_0 + re^{it}$ for $t \in [0, 2\pi]$. Then if $z \in B(z_0, r)$,

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w - z} dw. \tag{24.5}$$

Proof: Consider for $\alpha \in [0,1]$,

$$g(\alpha) \equiv \int_0^{2\pi} \frac{f\left(z + \alpha\left(z_0 + re^{it} - z\right)\right)}{re^{it} + z_0 - z} rie^{it} dt.$$

If α equals one, this reduces to the integral in (24.5). We will show g is a constant and that $g(0) = f(z)\, 2\pi i$. First we consider the claim about $g(0)$.

$$\begin{aligned} g(0) &= \left(\int_0^{2\pi} \frac{re^{it}}{re^{it} + z_0 - z} dt\right) if(z) \\ &= if(z) \left(\int_0^{2\pi} \frac{1}{1 - \frac{z - z_0}{re^{it}}} dt\right) \\ &= if(z) \int_0^{2\pi} \sum_{n=0}^{\infty} r^{-n} e^{-int} (z - z_0)^n dt \end{aligned}$$

because $\left|\frac{z - z_0}{re^{it}}\right| < 1$. Since this sum converges uniformly we may interchange the sum and the integral to obtain

$$\begin{aligned} g(0) &= if(z) \sum_{n=0}^{\infty} r^{-n} (z - z_0)^n \int_0^{2\pi} e^{-int} dt \\ &= 2\pi i f(z) \end{aligned}$$

because $\int_0^{2\pi} e^{-int} dt = 0$ if $n > 0$.

Next we show that g is constant. By Corollary 24.4, for $\alpha \in (0,1)$,

$$\begin{aligned} g'(\alpha) &= \int_0^{2\pi} \frac{f'\left(z+\alpha\left(z_0+re^{it}-z\right)\right)\left(re^{it}+z_0-z\right)}{re^{it}+z_0-z} rie^{it}dt \\ &= \int_0^{2\pi} f'\left(z+\alpha\left(z_0+re^{it}-z\right)\right) rie^{it}dt \\ &= \int_0^{2\pi} \frac{d}{dt}\left(f\left(z+\alpha\left(z_0+re^{it}-z\right)\right)\frac{1}{\alpha}\right)dt \\ &= f\left(z+\alpha\left(z_0+re^{i2\pi}-z\right)\right)\frac{1}{\alpha} - f\left(z+\alpha\left(z_0+re^{0}-z\right)\right)\frac{1}{\alpha} = 0. \end{aligned}$$

Now g is continuous on $[0,1]$ and $g'(t)=0$ on $(0,1)$ so by Lemma 24.2, g equals a constant. This constant can only be $g(0)=2\pi i f(z)$. Thus,

$$g(1) = \int_\gamma \frac{f(w)}{w-z}dw = g(0) = 2\pi i f(z).$$

This proves the theorem.

This is a very significant theorem. We give a few applications next.

Theorem 24.6 *Let $f : U \to \mathbb{C}$ be analytic where U is an open set in $\mathbb{C}$. Then f has infinitely many derivatives on U. Furthermore, for all $z \in B(z_0, r)$,*

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_\gamma \frac{f(w)}{(w-z)^{n+1}}dw \tag{24.6}$$

where $\gamma(t) \equiv z_0 + re^{it}, t \in [0, 2\pi]$ for r small enough that $B(z_0,r) \subseteq U$.

Proof: Let $z \in B(z_0,r) \subseteq U$ and let $\overline{B(z_0,r)} \subseteq U$. Then, letting $\gamma(t) \equiv z_0 + re^{it}, t \in [0,2\pi]$, and h small enough,

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-z}dw,\ f(z+h) = \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-z-h}dw$$

Now

$$\frac{1}{w-z-h} - \frac{1}{w-z} = \frac{h}{(-w+z+h)(-w+z)}$$

and so

$$\begin{aligned} \frac{f(z+h)-f(z)}{h} &= \frac{1}{2\pi hi}\int_\gamma \frac{hf(w)}{(-w+z+h)(-w+z)}dw \\ &= \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{(-w+z+h)(-w+z)}dw. \end{aligned}$$

Now for all h sufficiently small, there exists a constant C independent of such h such that

$$\begin{aligned} &\left|\frac{1}{(-w+z+h)(-w+z)} - \frac{1}{(-w+z)(-w+z)}\right| \\ = &\left|\frac{h}{(w-z-h)(w-z)^2}\right| \leq C|h| \end{aligned}$$

and so, the integrand converges uniformly as $h \to 0$ to

$$= \frac{f(w)}{(w-z)^2}$$

Therefore, we may take the limit as $h \to 0$ inside the integral to obtain

$$f'(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{(w-z)^2} dw.$$

Continuing in this way, we obtain (24.6).

This is a very remarkable result. We just showed that the existence of one continuous derivative implies the existence of all derivatives, in contrast to the theory of functions of a real variable. Actually, we just showed a little more than what the theorem states. The above proof establishes the following corollary.

Corollary 24.7 *Suppose f is continuous on $\partial B(z_0, r)$ and suppose that for all $z \in B(z_0, r)$,*

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w-z} dw,$$

where $\gamma(t) \equiv z + re^{it}, t \in [0, 2\pi]$. Then f is analytic on $B(z_0, r)$ and in fact has infinitely many derivatives on $B(z_0, r)$.

We also have the following simple lemma as an application of the above.

Lemma 24.8 *Let $\gamma(t) = z_0 + re^{it}$, for $t \in [0, 2\pi]$, suppose $f_n \to f$ uniformly on $\overline{B(z_0, r)}$, and suppose*

$$f_n(z) = \frac{1}{2\pi i} \int_\gamma \frac{f_n(w)}{w-z} dw \tag{24.7}$$

for $z \in B(z_0, r)$. Then

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w-z} dw, \tag{24.8}$$

implying that f is analytic on $B(z_0, r)$.

Proof: From (24.7) and the uniform convergence of f_n to f on $\gamma([0, 2\pi])$, we have that the integrals in (24.7) converge to

$$\frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w-z} dw.$$

Therefore, the formula (24.8) follows.

Proposition 24.9 *Let $\{a_n\}$ denote a sequence of complex numbers. Then there exists $R \in [0, \infty]$ such that*

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

converges absolutely if $|z - z_0| < R$, diverges if $|z - z_0| > R$ and converges uniformly on $B(z_0, r)$ for all $r < R$. Furthermore, if $R > 0$, the function,

$$f(z) \equiv \sum_{k=0}^{\infty} a_k (z - z_0)^k$$

is analytic on $B(z_0, R)$.

Proof: The assertions about absolute convergence are routine from the root test if we define

$$R \equiv \left(\limsup_{n\to\infty} |a_n|^{1/n}\right)^{-1}$$

with $R = \infty$ if the quantity in parenthesis equals zero. The assertion about uniform convergence follows from the Weierstrass M test if we use $M_n \equiv |a_n| r^n$. ($\sum_{n=0}^{\infty} |a_n| r^n < \infty$ by the root test). It only remains to verify the assertion about $f(z)$ being analytic in the case where $R > 0$. Let $0 < r < R$ and define $f_n(z) \equiv \sum_{k=0}^{n} a_k (z - z_0)^k$. Then f_n is a polynomial and so it is analytic. Thus, by the Cauchy integral formula above,

$$f_n(z) = \frac{1}{2\pi i} \int_\gamma \frac{f_n(w)}{w - z} dw$$

where $\gamma(t) = z_0 + re^{it}$, for $t \in [0, 2\pi]$. By Lemma 24.8 and the first part of this proposition involving uniform convergence, we obtain

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w - z} dw.$$

Therefore, f is analytic on $B(z_0, r)$ by Corollary 24.7. Since $r < R$ is arbitrary, this shows f is analytic on $B(z_0, R)$.

This proposition shows that all functions which are given as power series are analytic on their circle of convergence, the set of complex numbers, z, such that $|z - z_0| < R$. Next we show that every analytic function can be realized as a power series.

Theorem 24.10 *If $f : U \to \mathbb{C}$ is analytic and if $B(z_0, r) \subseteq U$, then*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{24.9}$$

for all $|z - z_0| < r$. Furthermore,

$$a_n = \frac{f^{(n)}(z_0)}{n!}. \tag{24.10}$$

Proof: Consider $|z - z_0| < r$ and let $\gamma(t) = z_0 + re^{it}$, $t \in [0, 2\pi]$. Then for $w \in \gamma([0, 2\pi])$,

$$\left|\frac{z - z_0}{w - z_0}\right| < 1$$

and so, by the Cauchy integral formula, we may write

$$\begin{aligned}
f(z) &= \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w - z} dw \\
&= \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{(w - z_0)\left(1 - \frac{z - z_0}{w - z_0}\right)} dw \\
&= \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{(w - z_0)} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{w - z_0}\right)^n dw.
\end{aligned}$$

Since the series converges uniformly, we may interchange the integral and the sum to obtain

$$\begin{aligned}
f(z) &= \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_\gamma \frac{f(w)}{(w - z_0)^{n+1}}\right) (z - z_0)^n \\
&\equiv \sum_{n=0}^{\infty} a_n (z - z_0)^n
\end{aligned}$$

By Theorem 24.6 we see that (24.10) holds.

The following theorem pertains to functions which are analytic on all of $\mathbb{C}$, "entire" functions.

Theorem 24.11 *(Liouville's theorem) If f is a bounded entire function then f is a constant.*

Proof: Since f is entire, we can pick any $z \in \mathbb{C}$ and write

$$f'(z) = \frac{1}{2\pi i} \int_{\gamma_R} \frac{f(w)}{(w-z)^2} dw$$

where $\gamma_R(t) = z + Re^{it}$ for $t \in [0, 2\pi]$. Therefore,

$$|f'(z)| \leq C \frac{1}{R}$$

where C is some constant depending on the assumed bound on f. Since R is arbitrary, we can take $R \to \infty$ to obtain $f'(z) = 0$ for any $z \in \mathbb{C}$. It follows from this that f is constant for if z_j $j = 1, 2$ are two complex numbers, we can consider $h(t) = f(z_1 + t(z_2 - z_1))$ for $t \in [0, 1]$. Then $h'(t) = f'(z_1 + t(z_2 - z_1))(z_2 - z_1) = 0$. By Lemma 24.2 h is a constant on $[0, 1]$ which implies $f(z_1) = f(z_2)$.

With Liouville's theorem it becomes possible to give an easy proof of the fundamental theorem of algebra. It is ironic that all the best proofs of this theorem in algebra come from the subjects of analysis or topology. Out of all the proofs that have been given of this very important theorem, the following one based on Liouville's theorem is the easiest.

Theorem 24.12 *(Fundamental theorem of Algebra) Let*

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

be a polynomial where $n \geq 1$ and each coefficient is a complex number. Then there exists $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$.

Proof: Suppose not. Then $p(z)^{-1}$ is an entire function. Also

$$|p(z)| \geq |z|^n - \left(|a_{n-1}| |z|^{n-1} + \cdots + |a_1| |z| + |a_0|\right)$$

and so $\lim_{|z| \to \infty} |p(z)| = \infty$ which implies $\lim_{|z| \to \infty} \left|p(z)^{-1}\right| = 0$. It follows that, since $p(z)^{-1}$ is bounded for z in any bounded set, we must have that $p(z)^{-1}$ is a bounded entire function. But then it must be constant. However since $p(z)^{-1} \to 0$ as $|z| \to \infty$, this constant can only be 0. However, $\frac{1}{p(z)}$ is never equal to zero. This proves the theorem.

24.1 Exercises

1. Show that if $|e_k| \leq \varepsilon$, then $\left|\sum_{k=m}^{\infty} e_k \left(r^k - r^{k+1}\right)\right| < \varepsilon$ if $0 \leq r < 1$. **Hint:** Let $|\theta| = 1$ and verify that

$$\theta \sum_{k=m}^{\infty} e_k \left(r^k - r^{k+1}\right) = \left|\sum_{k=m}^{\infty} e_k \left(r^k - r^{k+1}\right)\right| = \sum_{k=m}^{\infty} \operatorname{Re}(\theta e_k) \left(r^k - r^{k+1}\right)$$

where $-\varepsilon < \operatorname{Re}(\theta e_k) < \varepsilon$.

2. Abel's theorem says that if $\sum_{n=0}^{\infty} a_n (z-a)^n$ has radius of convergence equal to 1 and if $A = \sum_{n=0}^{\infty} a_n$, then $\lim_{r\to 1-} \sum_{n=0}^{\infty} a_n r^n = A$. **Hint:** Show $\sum_{k=0}^{\infty} a_k r^k = \sum_{k=0}^{\infty} A_k \left(r^k - r^{k+1}\right)$ where A_k denotes the kth partial sum of $\sum a_j$. Thus

$$\sum_{k=0}^{\infty} a_k r^k = \sum_{k=m+1}^{\infty} A_k \left(r^k - r^{k+1}\right) + \sum_{k=0}^{m} A_k \left(r^k - r^{k+1}\right),$$

where $|A_k - A| < \varepsilon$ for all $k \le m$. In the first sum, write $A_k = A + e_k$ and use Problem 1. Use this theorem to verify that $\arctan(1) = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}$.

3. Find the integrals using the Cauchy integral formula.

   (a) $\int_\gamma \frac{\sin z}{z-i} dz$ where $\gamma(t) = 2e^{it} : t \in [0, 2\pi]$.

   (b) $\int_\gamma \frac{1}{z-a} dz$ where $\gamma(t) = a + re^{it} : t \in [0, 2\pi]$

   (c) $\int_\gamma \frac{\cos z}{z^2} dz$ where $\gamma(t) = e^{it} : t \in [0, 2\pi]$

   (d) $\int_\gamma \frac{\log(z)}{z^n} dz$ where $\gamma(t) = 1 + \frac{1}{2} e^{it} : t \in [0, 2\pi]$ and $n = 0, 1, 2$.

4. Let $\gamma(t) = 4e^{it} : t \in [0, 2\pi]$ and find $\int_\gamma \frac{z^2+4}{z(z^2+1)} dz$.

5. Suppose $f(z) = \sum_{n=0}^{\infty} a_n z^n$ for all $|z| < R$. Show that then

$$\frac{1}{2\pi} \int_0^{2\pi} \left| f\left(re^{i\theta}\right)\right|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2 r^{2n}$$

for all $r \in [0, R)$. **Hint:** Let $f_n(z) \equiv \sum_{k=0}^{n} a_k z^k$, show $\frac{1}{2\pi} \int_0^{2\pi} \left| f_n\left(re^{i\theta}\right)\right|^2 d\theta = \sum_{k=0}^{n} |a_k|^2 r^{2k}$ and then take limits as $n \to \infty$ using uniform convergence.

6. The Cauchy integral formula, marvelous as it is, can actually be improved upon. The Cauchy integral formula involves representing f by the values of f on the boundary of the disk, $B(a, r)$. It is possible to represent f by using only the values of $\operatorname{Re} f$ on the boundary. This leads to the Schwarz formula . Supply the details in the following outline.

   Suppose f is analytic on $|z| < R$ and

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{24.11}$$

   with the series converging uniformly on $|z| = R$. Then letting $|w| = R$,

$$2u(w) = f(w) + \overline{f(w)}$$

   and so

$$2u(w) = \sum_{k=0}^{\infty} a_k w^k + \sum_{k=0}^{\infty} \overline{a_k} \left(\overline{w}\right)^k. \tag{24.12}$$

   Now letting $\gamma(t) = Re^{it},\ t \in [0, 2\pi]$

$$\begin{aligned} \int_\gamma \frac{2u(w)}{w} dw &= \left(a_0 + \overline{a_0}\right) \int_\gamma \frac{1}{w} dw \\ &= 2\pi i \left(a_0 + \overline{a_0}\right). \end{aligned}$$

Thus, multiplying (24.12) by w^{-1},

$$\frac{1}{\pi i}\int_\gamma \frac{u(w)}{w}dw = a_0 + \overline{a_0}.$$

Now multiply (24.12) by $w^{-(n+1)}$ and integrate again to obtain

$$a_n = \frac{1}{\pi i}\int_\gamma \frac{u(w)}{w^{n+1}}dw.$$

Using these formulas for a_n in (24.11), we can interchange the sum and the integral (Why can we do this?) to write the following for $|z| < R$.

$$\begin{aligned} f(z) &= \frac{1}{\pi i}\int_\gamma \frac{1}{z}\sum_{k=0}^{\infty}\left(\frac{z}{w}\right)^{k+1} u(w)\,dw - \overline{a_0} \\ &= \frac{1}{\pi i}\int_\gamma \frac{u(w)}{w-z}dw - \overline{a_0}, \end{aligned}$$

which is the Schwarz formula. Now $\operatorname{Re} a_0 = \frac{1}{2\pi i}\int_\gamma \frac{u(w)}{w}dw$ and $\overline{a_0} = \operatorname{Re} a_0 - i \operatorname{Im} a_0$. Therefore, we can also write the Schwarz formula as

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{u(w)(w+z)}{(w-z)w}dw + i \operatorname{Im} a_0. \tag{24.13}$$

7. Take the real parts of the second form of the Schwarz formula to derive the Poisson formula for a disk,

$$u\left(re^{i\alpha}\right) = \frac{1}{2\pi}\int_0^{2\pi} \frac{u\left(Re^{i\theta}\right)\left(R^2 - r^2\right)}{R^2 + r^2 - 2Rr\cos(\theta - \alpha)}d\theta. \tag{24.14}$$

8. Suppose that $u(w)$ is a given real continuous function defined on $\partial B(0,R)$ and define $f(z)$ for $|z| < R$ by (24.13). Show that f, so defined is analytic. Explain why u given in (24.14) is harmonic. Show that

$$\lim_{r \to R-} u\left(re^{i\alpha}\right) = u\left(Re^{i\alpha}\right).$$

Thus u is a harmonic function which approaches a given function on the boundary and is therefore, a solution to the Dirichlet problem.

9. Suppose $f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$ for all $|z - z_0| < R$. Show that $f'(z) = \sum_{k=0}^{\infty} a_k k (z - z_0)^{k-1}$ for all $|z - z_0| < R$. **Hint:** Let $f_n(z)$ be a partial sum of f. Show that f_n' converges uniformly to some function, g on $|z - z_0| \le r$ for any $r < R$. Now use the Cauchy integral formula for a function and its derivative to identify g with f'.

10. Use Problem 9 to find the exact value of $\sum_{k=0}^{\infty} k^2 \left(\frac{1}{3}\right)^k$.

11. Prove the binomial formula,

$$(1+z)^\alpha = \sum_{n=0}^{\infty}\binom{\alpha}{n} z^n$$

where

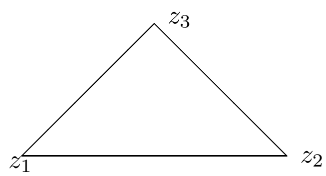
$$\binom{\alpha}{n} \equiv \frac{\alpha \cdots (\alpha - n + 1)}{n!}.$$

Can this be used to give a proof of the binomial formula, $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k} a^{n-k} b^k$? Explain.

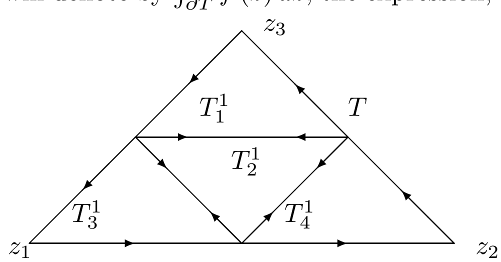
The general Cauchy integral formula

25.1 The Cauchy Goursat theorem

In this section we prove a fundamental theorem which is essential to the development which follows and is closely related to the question of when a function has a primitive. First of all, if we are given two points in $\mathbb{C}$, z_1 and z_2, we may consider $\gamma(t) \equiv z_1 + t(z_2 - z_1)$ for $t \in [0,1]$ to obtain a continuous bounded variation curve from z_1 to z_2. More generally, if $z_1, \cdots, z_m$ are points in $\mathbb{C}$ we can obtain a continuous bounded variation curve from z_1 to z_m which consists of first going from z_1 to z_2 and then from z_2 to z_3 and so on, till in the end one goes from z_{m-1} to z_m. We denote this piecewise linear curve as $\gamma(z_1, \cdots, z_m)$. Now let T be a triangle with vertices z_1, z_2 and z_3 encountered in the counter clockwise direction as shown.



Then we will denote by $\int_{\partial T} f(z)\,dz$, the expression, $\int_{\gamma(z_1,z_2,z_3,z_1)} f(z)\,dz$. Consider the following picture.



By Lemma 22.10 we may conclude that

$$\int_{\partial T} f(z)\,dz = \sum_{k=1}^{4} \int_{\partial T_k^1} f(z)\,dz. \tag{25.1}$$

On the "inside lines" the integrals cancel as claimed in Lemma 22.10 because there are two integrals going in opposite directions for each of these inside lines. Now we are ready to prove the Cauchy Goursat theorem.

Theorem 25.1 *(Cauchy Goursat) Let $f : U \to \mathbb{C}$ have the property that $f'(z)$ exists for all $z \in U$ and let T be a triangle contained in U. Then*

$$\int_{\partial T} f(w)\,dw = 0.$$

Proof: Suppose not. Then

$$\left|\int_{\partial T} f(w)\,dw\right| = \alpha \neq 0.$$

From (25.1) it follows

$$\alpha \leq \sum_{k=1}^{4} \left|\int_{\partial T_k^1} f(w)\,dw\right|$$

and so for at least one of these T_k^1, denoted from now on as T_1, we must have

$$\left|\int_{\partial T_1} f(w)\,dw\right| \geq \frac{\alpha}{4}.$$

Now let T_1 play the same role as T, subdivide as in the above picture, and obtain T_2 such that

$$\left|\int_{\partial T_2} f(w)\,dw\right| \geq \frac{\alpha}{4^2}.$$

Continue in this way, obtaining a sequence of triangles,

$$T_k \supseteq T_{k+1}, diam\,(T_k) \leq diam\,(T)\,2^{-k},$$

and

$$\left|\int_{\partial T_k} f(w)\,dw\right| \geq \frac{\alpha}{4^k}.$$

Then let $z \in \cap_{k=1}^{\infty} T_k$ and note that by assumption, $f'(z)$ exists. Therefore, for all k large enough,

$$\int_{\partial T_k} f(w)\,dw = \int_{\partial T_k} f(z) + f'(z)(w-z) + g(w)\,dw$$

where $|g(w)| < \varepsilon |w-z|$. Now observe that $w \to f(z) + f'(z)(w-z)$ has a primitive, namely,

$$F(w) = f(z)\,w + f'(z)(w-z)^2/2.$$

Therefore, by Corollary 22.13.

$$\int_{\partial T_k} f(w)\,dw = \int_{\partial T_k} g(w)\,dw.$$

From the definition, of the integral, we see

$$\begin{aligned}\frac{\alpha}{4^k} &\leq \left|\int_{\partial T_k} g(w)\,dw\right| \leq \varepsilon diam\,(T_k)\,(\text{length of }\partial T_k) \\ &\leq \varepsilon 2^{-k}\,(\text{length of }T)\,diam\,(T)\,2^{-k},\end{aligned}$$

and so

$$\alpha \leq \varepsilon\,(\text{length of }T)\,diam\,(T).$$

Since ε is arbitrary, this shows $\alpha = 0$, a contradiction. Thus $\int_{\partial T} f(w)\,dw = 0$ as claimed.

This fundamental result yields the following important theorem.

Theorem 25.2 *(Morera) Let U be an open set and let $f'(z)$ exist for all $z \in U$. Let $D \equiv \overline{B(z_0, r)} \subseteq U$. Then there exists $\varepsilon > 0$ such that f has a primitive on $B(z_0, r+\varepsilon)$.*

Proof: Choose $\varepsilon > 0$ small enough that $B(z_0, r+\varepsilon) \subseteq U$. Then for $w \in B(z_0, r+\varepsilon)$, define

$$F(w) \equiv \int_{\gamma(z_0,w)} f(u)\, du.$$

Then by the Cauchy Goursat theorem, and $w \in B(z_0, r+\varepsilon)$, it follows that for $|h|$ small enough,

$$\frac{F(w+h) - F(w)}{h} = \frac{1}{h} \int_{\gamma(w,w+h)} f(u)\, du$$

$$= \frac{1}{h} \int_0^1 f(w+th)\, h dt = \int_0^1 f(w+th)\, dt$$

which converges to $f(w)$ due to the continuity of f at w. This proves the theorem.

We can also give the following corollary whose proof is similar to the proof of the above theorem.

Corollary 25.3 *Let U be an open set and suppose that whenever*

$$\gamma(z_1, z_2, z_3, z_1)$$

is a closed curve bounding a triangle T, which is contained in U, and f is a continuous function defined on U, it follows that

$$\int_{\gamma(z_1,z_2,z_3,z_1)} f(z)\, dz = 0,$$

then f is analytic on U.

Proof: As in the proof of Morera's theorem, let $\overline{B(z_0, r)} \subseteq U$ and use the given condition to construct a primitive, F for f on $B(z_0, r)$. Then F is analytic and so by Theorem 24.6, it follows that F and hence f have infinitely many derivatives, implying that f is analytic on $B(z_0, r)$. Since z_0 is arbitrary, this shows f is analytic on U.

Theorem 25.4 *Let U be an open set in $\mathbb{C}$ and suppose $f : U \to \mathbb{C}$ has the property that $f'(z)$ exists for each $z \in U$. Then f is analytic on U.*

Proof: Let $z_0 \in U$ and let $B(z_0, r) \subseteq U$. By Morera's theorem f has a primitive, F on $B(z_0, r)$. It follows that F is analytic because it has a derivative, f, and this derivative is continuous. Therefore, by Theorem 24.6 F has infinitely many derivatives on $B(z_0, r)$ implying that f also has infinitely many derivatives on $B(z_0, r)$. Thus f is analytic as claimed.

It follows that we can say a function is analytic on an open set, U if and only if $f'(z)$ exists for $z \in U$. We just proved the derivative, if it exists, is automatically continuous.

The same proof used to prove Theorem 25.2 implies the following corollary.

Corollary 25.5 *Let U be a convex open set and suppose that $f'(z)$ exists for all $z \in U$. Then f has a primitive on U.*

Note that this implies that if U is a convex open set on which $f'(z)$ exists and if $\gamma : [a, b] \to U$ is a closed, continuous curve having bounded variation, then letting F be a primitive of f Theorem 22.12 implies

$$\int_\gamma f(z)\, dz = F(\gamma(b)) - F(\gamma(a)) = 0.$$

Notice how different this is from the situation of a function of a real variable. It is possible for a function of a real variable to have a derivative everywhere and yet the derivative can be discontinuous. A simple example is the following.

$$f(x) \equiv \begin{cases} x^2 \sin\left(\frac{1}{x}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}.$$

Then $f'(x)$ exists for all $x \in \mathbb{R}$. Indeed, if $x \neq 0$, the derivative equals $2x \sin \frac{1}{x} - \cos \frac{1}{x}$ which has no limit as $x \to 0$. However, from the definition of the derivative of a function of one variable, we see easily that $f'(0) = 0$.

25.2 The Cauchy integral formula

Here we develop the general version of the Cauchy integral formula valid for arbitrary closed rectifiable curves. The key idea in this development is the notion of the winding number. This is the number defined in the following theorem, also called the index. We make use of this winding number along with the earlier results, especially Liouville's theorem, to give an extremely general Cauchy integral formula.

Theorem 25.6 *Let $\gamma : [a,b] \to \mathbb{C}$ be continuous and have bounded variation with $\gamma(a) = \gamma(b)$. Also suppose that $z \notin \gamma([a,b])$. We define*

$$n(\gamma, z) \equiv \frac{1}{2\pi i} \int_\gamma \frac{dw}{w - z}. \tag{25.2}$$

Then $n(\gamma, \cdot)$ is continuous and integer valued. Furthermore, there exists a sequence, $\eta_k : [a,b] \to \mathbb{C}$ such that η_k is $C^1([a,b])$,

$$||\eta_k - \gamma|| < \frac{1}{k}, \eta_k(a) = \eta_k(b) = \gamma(a) = \gamma(b),$$

and $n(\eta_k, z) = n(\gamma, z)$ for all k large enough. Also $n(\gamma, \cdot)$ is constant on every component of $\mathbb{C} \setminus \gamma([a,b])$ and equals zero on the unbounded component of $\mathbb{C} \setminus \gamma([a,b])$.

Proof: First we verify the assertion about continuity.

$$\begin{aligned} |n(\gamma, z) - n(\gamma, z_1)| &\leq C \left| \int_\gamma \left(\frac{1}{w - z} - \frac{1}{w - z_1} \right) dw \right| \\ &\leq \widetilde{C} (\text{Length of } \gamma) |z_1 - z| \end{aligned}$$

whenever z_1 is close enough to z. This proves the continuity assertion.

Next we need to show the winding number equals an integer. To do so, use Theorem 22.11 to obtain η_k, a function in $C^1([a,b])$ such that $z \notin \eta_k([a,b])$ for all k large enough, $\eta_k(x) = \gamma(x)$ for $x = a, b$, and

$$\left| \frac{1}{2\pi i} \int_\gamma \frac{dw}{w - z} - \frac{1}{2\pi i} \int_{\eta_k} \frac{dw}{w - z} \right| < \frac{1}{k}, \ ||\eta_k - \gamma|| < \frac{1}{k}.$$

We will show each of $\frac{1}{2\pi i} \int_{\eta_k} \frac{dw}{w-z}$ is an integer. To simplify the notation, we write η instead of η_k.

$$\int_\eta \frac{dw}{w - z} = \int_a^b \frac{\eta'(s)\, ds}{\eta(s) - z}.$$

We define

$$g(t) \equiv \int_a^t \frac{\eta'(s)\,ds}{\eta(s)-z}. \tag{25.3}$$

Then

$$\begin{aligned}\left(e^{-g(t)}(\eta(t)-z)\right)' &= e^{-g(t)}\eta'(t) - e^{-g(t)}g'(t)(\eta(t)-z) \\ &= e^{-g(t)}\eta'(t) - e^{-g(t)}\eta'(t) = 0.\end{aligned}$$

It follows that $e^{-g(t)}(\eta(t)-z)$ equals a constant. In particular, using the fact that $\eta(a) = \eta(b)$,

$$e^{-g(b)}(\eta(b)-z) = e^{-g(a)}(\eta(a)-z) = (\eta(a)-z) = (\eta(b)-z)$$

and so $e^{-g(b)} = 1$. This happens if and only if $-g(b) = 2m\pi i$ for some integer m. Therefore, (25.3) implies

$$2m\pi i = \int_a^b \frac{\eta'(s)\,ds}{\eta(s)-z} = \int_\eta \frac{dw}{w-z}.$$

Therefore, $\frac{1}{2\pi i}\int_{\eta_k}\frac{dw}{w-z}$ is a sequence of integers converging to $\frac{1}{2\pi i}\int_\gamma \frac{dw}{w-z} \equiv n(\gamma, z)$ and so $n(\gamma, z)$ must also be an integer and $n(\eta_k, z) = n(\gamma, z)$ for all k large enough.

Since $n(\gamma, \cdot)$ is continuous and integer valued, it follows that it must be constant on every connected component of $\mathbb{C} \setminus \gamma([a,b])$. It is clear that $n(\gamma, z)$ equals zero on the unbounded component because from the formula,

$$\lim_{z\to\infty} |n(\gamma, z)| \leq \lim_{z\to\infty} V(\gamma, [a,b])\left(\frac{1}{|z|-c}\right)$$

where $c \geq \max\{|w| : w \in \gamma([a,b])\}$ .This proves the theorem.

It is a good idea to consider a simple case to get an idea of what the winding number is measuring. To do so, consider $\gamma : [a,b] \to \mathbb{C}$ such that γ is continuous, closed and bounded variation. Suppose also that γ is one to one on (a,b). Such a curve is called a simple closed curve. It can be shown that such a simple closed curve divides the plane into exactly two components, an "inside" bounded component and an "outside" unbounded component. This is called the Jordan Curve theorem or the Jordan separation theorem. For a proof of this difficult result, see the chapter on degree theory. For now, it suffices to simply assume that γ is such that this result holds. This will usually be obvious anyway. We also suppose that it is possible to change the parameter to be in $[0, 2\pi]$, in such a way that $\gamma(t) + \lambda\left(z + re^{it} - \gamma(t)\right) - z \neq 0$ for all $t \in [0, 2\pi]$ and $\lambda \in [0,1]$. (As t goes from 0 to 2π the point $\gamma(t)$ traces the curve $\gamma([0, 2\pi])$ in the counter clockwise direction.) Suppose $z \in D$, the inside of the simple closed curve and consider the curve $\delta(t) = z + re^{it}$ for $t \in [0, 2\pi]$ where r is chosen small enough that $\overline{B(z,r)} \subseteq D$. Then we claim that $n(\delta, z) = n(\gamma, z)$.

Proposition 25.7 *Under the above conditions, $n(\delta, z) = n(\gamma, z)$ and $n(\delta, z) = 1$.*

Proof: By changing the parameter, we may assume that $[a,b] = [0, 2\pi]$. From Theorem 25.6 it suffices to assume also that γ is C^1. Define $h_\lambda(t) \equiv \gamma(t) + \lambda\left(z + re^{it} - \gamma(t)\right)$ for $\lambda \in [0,1]$. (This function is called a homotopy of the curves γ and δ.) Note that for each $\lambda \in [0,1]$, $t \to h_\lambda(t)$ is a closed C^1 curve. Also,

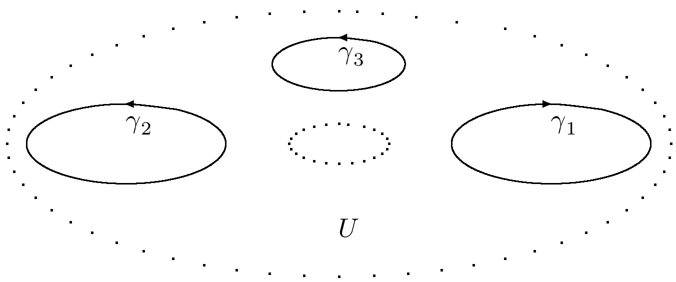
$$\frac{1}{2\pi i}\int_{h_\lambda}\frac{1}{w-z}dw = \frac{1}{2\pi i}\int_0^{2\pi}\frac{\gamma'(t) + \lambda\left(rie^{it} - \gamma'(t)\right)}{\gamma(t) + \lambda(z + re^{it} - \gamma(t)) - z}dt.$$

We know this number is an integer and it is routine to verify that it is a continuous function of λ. When $\lambda = 0$ it equals $n(\gamma, z)$ and when $\lambda = 1$ it equals $n(\delta, z)$. Therefore, $n(\delta, z) = n(\gamma, z)$. It only remains to compute $n(\delta, z)$.

$$n(\delta, z) = \frac{1}{2\pi i}\int_0^{2\pi}\frac{rie^{it}}{re^{it}}dt = 1.$$

This proves the proposition.

Now if γ was not one to one but caused the point, $\gamma(t)$ to travel around $\gamma([a,b])$ twice, we could modify the above argument to have the parameter interval, $[0,4\pi]$ and still find $n(\delta,z) = n(\gamma,z)$ only this time, $n(\delta,z) = 2$. Thus the winding number is just what its name suggests. It measures the number of times the curve winds around the point. One might ask why bother with the winding number if this is all it does. The reason is that the notion of counting the number of times a curve winds around a point is rather vague. The winding number is precise. It is also the natural thing to consider in the general Cauchy integral formula presented below. We have in mind a situation typified by the following picture in which U is the open set between the dotted curves and γ_j are closed rectifiable curves in U.



The following theorem is the general Cauchy integral formula.

Theorem 25.8 *Let U be an open subset of the plane and let $f : U \to \mathbb{C}$ be analytic. If $\gamma_k : [a_k, b_k] \to U$, $k = 1, \cdots, m$ are continuous closed curves having bounded variation such that for all $z \notin U$,*

$$\sum_{k=1}^{m} n(\gamma_k, z) = 0,$$

then for all $z \in U \setminus \cup_{k=1}^{m} \gamma_k([a_k, b_k])$,

$$f(z) \sum_{k=1}^{m} n(\gamma_k, z) = \sum_{k=1}^{m} \frac{1}{2\pi i} \int_{\gamma_k} \frac{f(w)}{w-z} dw.$$

Proof: Let ϕ be defined on $U \times U$ by

$$\phi(z,w) \equiv \begin{cases} \frac{f(w)-f(z)}{w-z} \text{ if } w \neq z \\ f'(z) \text{ if } w = z \end{cases}.$$

Then ϕ is analytic as a function of both z and w and is continuous in $U \times U$. The claim that this function is analytic as a function of both z and w is obvious at points where $z \neq w$, and is most easily seen using Theorem 24.10 at points, where $z = w$. Indeed, if (z,z) is such a point, we need to verify that $w \to \phi(z,w)$ is analytic even at $w = z$. But by Theorem 24.10, for all h small enough,

$$\frac{\phi(z, z+h) - \phi(z,z)}{h} = \frac{1}{h}\left[\frac{f(z+h) - f(z)}{h} - f'(z)\right]$$

$$\begin{aligned} &= \frac{1}{h}\left[\frac{1}{h}\sum_{k=1}^{\infty} \frac{f^{(k)}(z)}{k!} h^k - f'(z)\right] \\ &= \left[\sum_{k=2}^{\infty} \frac{f^{(k)}(z)}{k!} h^{k-2}\right] \to \frac{f''(z)}{2!}. \end{aligned}$$

Similarly, $z \to \phi(z, w)$ is analytic even if $z = w$.

We define

$$h(z) \equiv \frac{1}{2\pi i} \sum_{k=1}^{m} \int_{\gamma_k} \phi(z, w)\, dw.$$

We wish to show that h is analytic on U. To do so, we verify

$$\int_{\partial T} h(z)\, dz = 0$$

for every triangle, T, contained in U and apply Corollary 25.3. To do this we use Theorem 22.11 to obtain for each k, a sequence of functions, $\eta_{kn} \in C^1([a_k, b_k])$ such that

$$\eta_{kn}(x) = \gamma_k(x) \text{ for } x \in \{a_k, b_k\}$$

and

$$\eta_{kn}([a_k, b_k]) \subseteq U, \;\; ||\eta_{kn} - \gamma_k|| < \frac{1}{n},$$

$$\left| \int_{\eta_{kn}} \phi(z, w)\, dw - \int_{\gamma_k} \phi(z, w)\, dw \right| < \frac{1}{n}, \tag{25.4}$$

for all $z \in T$. Then applying Fubini's theorem, we can write

$$\int_{\partial T} \int_{\eta_{kn}} \phi(z, w)\, dwdz = \int_{\eta_{kn}} \int_{\partial T} \phi(z, w)\, dzdw = 0$$

because ϕ is given to be analytic. By (25.4),

$$\int_{\partial T} \int_{\gamma_k} \phi(z, w)\, dwdz = \lim_{n\to\infty} \int_{\partial T} \int_{\eta_{kn}} \phi(z, w)\, dwdz = 0$$

and so h is analytic on U as claimed.

Now let H denote the set,

$$H \equiv \left\{ z \in \mathbb{C} \setminus \cup_{k=1}^{m} \gamma_k([a_k, b_k]) : \sum_{k=1}^{m} n(\gamma_k, z) = 0 \right\}.$$

We know that H is an open set because $z \to \sum_{k=1}^{m} n(\gamma_k, z)$ is integer valued and continuous. Define

$$g(z) \equiv \left\{ \begin{array}{l} h(z) \text{ if } z \in U \\ \frac{1}{2\pi i} \sum_{k=1}^{m} \int_{\gamma_k} \frac{f(w)}{w-z} dw \text{ if } z \in H \end{array} \right. . \tag{25.5}$$

We need to verify that $g(z)$ is well defined. For $z \in U \cap H$, we know $z \notin \cup_{k=1}^{m} \gamma_k([a_k, b_k])$ and so

$$\begin{aligned} g(z) &= \frac{1}{2\pi i} \sum_{k=1}^{m} \int_{\gamma_k} \frac{f(w) - f(z)}{w - z} dw \\ &= \frac{1}{2\pi i} \sum_{k=1}^{m} \int_{\gamma_k} \frac{f(w)}{w - z} dw - \frac{1}{2\pi i} \sum_{k=1}^{m} \int_{\gamma_k} \frac{f(z)}{w - z} dw \\ &= \frac{1}{2\pi i} \sum_{k=1}^{m} \int_{\gamma_k} \frac{f(w)}{w - z} dw \end{aligned}$$

because $z \in H$. This shows $g(z)$ is well defined. Also, g is analytic on U because it equals h there. It is routine to verify that g is analytic on H also. By assumption, $U^C \subseteq H$ and so $U \cup H = \mathbb{C}$ showing that g is an entire function.

Now note that $\sum_{k=1}^m n(\gamma_k, z) = 0$ for all z contained in the unbounded component of $\mathbb{C} \setminus \cup_{k=1}^m \gamma_k([a_k, b_k])$ which component contains $B(0,r)^C$ for r large enough. It follows that for $|z| > r$, it must be the case that $z \in H$ and so for such z, the bottom description of $g(z)$ found in (25.5) is valid. Therefore, it follows

$$\lim_{|z| \to \infty} |g(z)| = 0$$

and so g is bounded and entire. By Liouville's theorem, g is a constant. Hence, from the above equation, the constant can only equal zero.

For $z \in U \setminus \cup_{k=1}^m \gamma_k([a_k, b_k])$,

$$0 = \frac{1}{2\pi i} \sum_{k=1}^m \int_{\gamma_k} \frac{f(w) - f(z)}{w - z} dw =$$

$$\frac{1}{2\pi i} \sum_{k=1}^m \int_{\gamma_k} \frac{f(w)}{w - z} dw - f(z) \sum_{k=1}^m n(\gamma_k, z).$$

This proves the theorem.

Corollary 25.9 *Let U be an open set and let $\gamma_k : [a_k, b_k] \to U$, $k = 1, \cdots, m$, be closed, continuous and of bounded variation. Suppose also that*

$$\sum_{k=1}^m n(\gamma_k, z) = 0$$

for all $z \notin U$. Then if $f : U \to \mathbb{C}$ is analytic, we have

$$\sum_{k=1}^m \int_{\gamma_k} f(w) \, dw = 0.$$

Proof: This follows from Theorem 25.8 as follows. Let

$$g(w) = f(w)(w - z)$$

where $z \in U \setminus \cup_{k=1}^m \gamma_k([a_k, b_k])$. Then by this theorem,

$$0 = 0 \sum_{k=1}^m n(\gamma_k, z) = g(z) \sum_{k=1}^m n(\gamma_k, z) =$$

$$\sum_{k=1}^m \frac{1}{2\pi i} \int_{\gamma_k} \frac{g(w)}{w - z} dw = \frac{1}{2\pi i} \sum_{k=1}^m \int_{\gamma_k} f(w) \, dw.$$

Another simple corollary to the above theorem is Cauchy's theorem for a simply connected region.

Definition 25.10 *We say an open set, $U \subseteq \mathbb{C}$ is a region if it is open and connected. We say U is simply connected if $\widehat{\mathbb{C}} \setminus U$ is connected.*

Corollary 25.11 *Let $\gamma : [a, b] \to U$ be a continuous closed curve of bounded variation where U is a simply connected region in $\mathbb{C}$ and let $f : U \to \mathbb{C}$ be analytic. Then*

$$\int_\gamma f(w)\,dw = 0.$$

Proof: Let D denote the unbounded component of $\widehat{\mathbb{C}} \setminus \gamma([a, b])$. Thus $\infty \in \widehat{\mathbb{C}} \setminus \gamma([a, b])$.Then the connected set, $\widehat{\mathbb{C}} \setminus U$ is contained in D since every point of $\widehat{\mathbb{C}} \setminus U$ must be in some component of $\widehat{\mathbb{C}} \setminus \gamma([a, b])$ and ∞ is contained in both $\widehat{\mathbb{C}} \setminus U$ and D. Thus D must be the component that contains $\widehat{\mathbb{C}} \setminus U$. It follows that $n(\gamma, \cdot)$ must be constant on $\widehat{\mathbb{C}} \setminus U$, its value being its value on D. However, for $z \in D$,

$$n(\gamma, z) = \frac{1}{2\pi i} \int_\gamma \frac{1}{w - z}\,dw$$

and so $\lim_{|z| \to \infty} n(\gamma, z) = 0$ showing $n(\gamma, z) = 0$ on D. Therefore we have verified the hypothesis of Theorem 25.8. Let $z \in U \cap D$ and define

$$g(w) \equiv f(w)(w - z).$$

Thus g is analytic on U and by Theorem 25.8,

$$0 = n(z, \gamma)\,g(z) = \frac{1}{2\pi i} \int_\gamma \frac{g(w)}{w - z}\,dw = \frac{1}{2\pi i} \int_\gamma f(w)\,dw.$$

This proves the corollary.

The following is a very significant result which will be used later.

Corollary 25.12 *Suppose U is a simply connected open set and $f : U \to \mathbb{C}$ is analytic. Then f has a primitive, F, on U. Recall this means there exists F such that $F'(z) = f(z)$ for all $z \in U$.*

Proof: Pick a point, $z_0 \in U$ and let V denote those points, z of U for which there exists a curve, $\gamma : [a, b] \to U$ such that γ is continuous, of bounded variation, $\gamma(a) = z_0$, and $\gamma(b) = z$. Then it is easy to verify that V is both open and closed in U and therefore, $V = U$ because U is connected. Denote by $\gamma_{z_0,z}$ such a curve from z_0 to z and define

$$F(z) \equiv \int_{\gamma_{z_0,z}} f(w)\,dw.$$

Then F is well defined because if $\gamma_j, j = 1, 2$ are two such curves, it follows from Corollary 25.11 that

$$\int_{\gamma_1} f(w)\,dw + \int_{-\gamma_2} f(w)\,dw = 0,$$

implying that

$$\int_{\gamma_1} f(w)\,dw = \int_{\gamma_2} f(w)\,dw.$$

Now this function, F is a primitive because, thanks to Corollary 25.11

$$\begin{aligned}(F(z+h) - F(z))\,h^{-1} &= \frac{1}{h} \int_{\gamma_{z,z+h}} f(w)\,dw \\ &= \frac{1}{h} \int_0^1 f(z + th)\,h\,dt\end{aligned}$$

and so, taking the limit as $h \to 0$, we see $F'(z) = f(z)$.

25.3 Exercises

1. If U is simply connected, f is analytic on U and f has no zeros in U, show there exists an analytic function, F, defined on U such that $e^F = f$.

2. Let U be an open set and let f be analytic on U. Show that if $a \in U$, then $f(z) = \sum_{k=0}^{\infty} b_k (z-a)^k$ whenever $|z-a| < R$ where R is the distance between a and the nearest point where f fails to have a derivative. The number R, is called the radius of convergence and the power series is said to be expanded about a.

3. Find the radius of convergence of the function $\frac{1}{1+z^2}$ expanded about $a = 2$. Note there is nothing wrong with the function, $\frac{1}{1+x^2}$ when considered as a function of a real variable, x for any value of x. However, if we insist on using power series, we find that there is a limitation on the values of x for which the power series converges due to the presence in the complex plane of a point, i, where the function fails to have a derivative.

4. What if we defined an open set, U to be simply connected if $\mathbb{C} \setminus U$ is connected. Would it amount to the same thing? **Hint:** Consider the outside of $B(0,1)$.

5. Let $\gamma(t) = e^{it} : t \in [0, 2\pi]$. Find $\int_\gamma \frac{1}{z^n} dz$ for $n = 1, 2, \cdots$.

6. Show $i \int_0^{2\pi} (2\cos\theta)^{2n} d\theta = \int_\gamma \left(z + \frac{1}{z}\right)^{2n} \left(\frac{1}{z}\right) dz$ where $\gamma(t) = e^{it} : t \in [0, 2\pi]$. Then evaluate this integral using the binomial theorem and the previous problem.

7. Let $f : U \to \mathbb{C}$ be analytic and $f(z) = u(x,y) + iv(x,y)$. Show u, v and uv are all harmonic although it can happen that u^2 is not. Recall that a function, w is harmonic if $w_{xx} + w_{yy} = 0$.

8. Suppose that for some constants $a, b \neq 0$, $a, b \in \mathbb{R}$, $f(z + ib) = f(z)$ for all $z \in \mathbb{C}$ and $f(z + a) = f(z)$ for all $z \in \mathbb{C}$. If f is analytic, show that f must be constant. Can you generalize this? **Hint:** This uses Liouville's theorem.

The open mapping theorem

In this chapter we present the open mapping theorem for analytic functions. This important result states that analytic functions map connected open sets to connected open sets or else to single points. It is very different than the situation for a function of a real variable.

26.1 Zeros of an analytic function

In this section we give a very surprising property of analytic functions which is in stark contrast to what takes place for functions of a real variable. It turns out the zeros of an analytic function which is not constant on some region cannot have a limit point.

Theorem 26.1 *Let U be a connected open set (region) and let $f : U \to \mathbb{C}$ be analytic. Then the following are equivalent.*

1. *$f(z) = 0$ for all $z \in U$*
2. *There exists $z_0 \in U$ such that $f^{(n)}(z_0) = 0$ for all n.*
3. *There exists $z_0 \in U$ which is a limit point of the set,*

$$Z \equiv \{z \in U : f(z) = 0\}.$$

Proof: It is clear the first condition implies the second two. Suppose the third holds. Then for z near z_0 we have

$$f(z) = \sum_{n=k}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$$

where $k \geq 1$ since z_0 is a zero of f. Suppose $k < \infty$. Then,

$$f(z) = (z - z_0)^k g(z)$$

where $g(z_0) \neq 0$. Letting $z_n \to z_0$ where $z_n \in Z, z_n \neq z_0$, it follows

$$0 = (z_n - z_0)^k g(z_n)$$

which implies $g(z_n) = 0$. Then by continuity of g, we see that $g(z_0) = 0$ also, contrary to the choice of k. Therefore, k cannot be less than ∞ and so z_0 is a point satisfying the second condition.

Now suppose the second condition and let

$$S \equiv \left\{z \in U : f^{(n)}(z) = 0 \text{ for all } n\right\}.$$

It is clear that S is a closed set which by assumption is nonempty. However, this set is also open. To see this, let $z \in S$. Then for all w close enough to z,

$$f(w) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z)}{k!} (w-z)^k = 0.$$

Thus f is identically equal to zero near $z \in S$. Therefore, all points near z are contained in S also, showing that S is an open set. Now $U = S \cup (U \setminus S)$, the union of two disjoint open sets, S being nonempty. It follows the other open set, $U \setminus S$, must be empty because U is connected. Therefore, the first condition is verified. This proves the theorem. (See the following diagram.)

$$\begin{array}{ccccc} & & 1.) & & \\ & \swarrow\nearrow & & \searrow & \\ 2.) & & \longleftarrow & & 3.) \end{array}$$

Note how radically different this from the theory of functions of a real variable. Consider, for example the function

$$f(x) \equiv \begin{cases} x^2 \sin\left(\frac{1}{x}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$$

which has a derivative for all $x \in \mathbb{R}$ and for which 0 is a limit point of the set, Z, even though f is not identically equal to zero.

26.2 The open mapping theorem

With this preparation we are ready to prove the open mapping theorem, an even more surprising result than the theorem about the zeros of an analytic function.

Theorem 26.2 *(Open mapping theorem) Let U be a region in $\mathbb{C}$ and suppose $f : U \to \mathbb{C}$ is analytic. Then $f(U)$ is either a point or a region. In the case where $f(U)$ is a region, it follows that for each $z_0 \in U$, there exists an open set, V containing z_0 such that for all $z \in V$,*

$$f(z) = f(z_0) + \phi(z)^m \tag{26.1}$$

where $\phi : V \to B(0,\delta)$ is one to one, analytic and onto, $\phi(z_0) = 0$, $\phi'(z) \neq 0$ on V and ϕ^{-1} analytic on $B(0,\delta)$. If f is one to one, then $m = 1$ for each z_0 and $f^{-1} : f(U) \to U$ is analytic.

Proof: Suppose $f(U)$ is not a point. Then if $z_0 \in U$ it follows there exists $r > 0$ such that $f(z) \neq f(z_0)$ for all $z \in B(z_0, r) \setminus \{z_0\}$. Otherwise, z_0 would be a limit point of the set,

$$\{z \in U : f(z) - f(z_0) = 0\}$$

which would imply from Theorem 26.1 that $f(z) = f(z_0)$ for all $z \in U$. Therefore, making r smaller if necessary, we may write, using the power series of f,

$$f(z) = f(z_0) + (z - z_0)^m g(z)$$

for all $z \in B(z_0, r)$, where $g(z) \neq 0$ on $B(z_0, r)$. Then $\frac{g'}{g}$ is an analytic function on $B(z_0, r)$ and so by Corollary 25.5 it has a primitive on $B(z_0, r)$, h. Therefore, using the product rule and the chain rule, $\left(ge^{-h}\right)' = 0$ and so there exists a constant, $C = e^{a+ib}$ such that on $B(z_0, r)$,

$$ge^{-h} = e^{a+ib}.$$

Therefore,

$$g(z) = e^{h(z)+a+ib}$$

and so, modifying h by adding in the constant, $a+ib$, we see $g(z) = e^{h(z)}$ where $h'(z) = \frac{g'(z)}{g(z)}$ on $B(z_0, r)$. Letting

$$\phi(z) = (z - z_0) e^{\frac{h(z)}{m}}$$

we obtain the formula (26.1) valid on $B(z_0, r)$. Now

$$\phi'(z_0) = e^{\frac{h(z_0)}{m}} \neq 0$$

and so, restricting r we may assume that $\phi'(z) \neq 0$ for all $z \in B(z_0, r)$. We need to verify that there is an open set, V contained in $B(z_0, r)$ such that ϕ maps V onto $B(0, \delta)$ for some $\delta > 0$.

Let $\phi(z) = u(x, y) + iv(x, y)$ where $z = x + iy$. Then

$$\begin{pmatrix} u(x_0, y_0) \\ v(x_0, y_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

because for $z_0 = x_0 + iy_0$, $\phi(z_0) = 0$. In addition to this, the functions u and v are in $C^1(B(0, r))$ because ϕ is analytic. By the Cauchy Riemann equations,

$$\begin{vmatrix} u_x(x_0, y_0) & u_y(x_0, y_0) \\ v_x(x_0, y_0) & v_y(x_0, y_0) \end{vmatrix} = \begin{vmatrix} u_x(x_0, y_0) & -v_x(x_0, y_0) \\ v_x(x_0, y_0) & u_x(x_0, y_0) \end{vmatrix}$$

$$= u_x^2(x_0, y_0) + v_x^2(x_0, y_0) = \left|\phi'(z_0)\right|^2 \neq 0.$$

Therefore, by the inverse function theorem there exists an open set, V, containing z_0 and $\delta > 0$ such that $(u, v)^T$ maps V one to one onto $B(0, \delta)$. Thus ϕ is one to one onto $B(0, \delta)$ as claimed. It follows that ϕ^m maps V onto $B(0, \delta^m)$. Therefore, the formula (26.1) implies that f maps the open set, V, containing z_0 to an open set. This shows $f(U)$ is an open set. It is connected because f is continuous and U is connected. Thus $f(U)$ is a region. It only remains to verify that ϕ^{-1} is analytic on $B(0, \delta)$. We show this by verifying the Cauchy Riemann equations.

Let

$$\begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} \tag{26.2}$$

for $(u, v)^T \in B(0, \delta)$. Then, letting $w = u + iv$, it follows that $\phi^{-1}(w) = x(u, v) + iy(u, v)$. We need to verify that

$$x_u = y_v,\ x_v = -y_u. \tag{26.3}$$

The inverse function theorem has already given us the continuity of these partial derivatives. From the equations (26.2), we have the following systems of equations.

$$\begin{array}{c} u_x x_u + u_y y_u = 1 \\ v_x x_u + v_y y_u = 0 \end{array},\quad \begin{array}{c} u_x x_v + u_y y_v = 0 \\ v_x x_v + v_y y_v = 1 \end{array}.$$

Solving these for $x_u, y_v,\ x_v$, and y_u, and using the Cauchy Riemann equations for u and v, yields (26.3).

It only remains to verify the assertion about the case where f is one to one. If $m > 1$, then $e^{\frac{2\pi i}{m}} \neq 1$ and so for $z_1 \in V$,

$$e^{\frac{2\pi i}{m}} \phi(z_1) \neq \phi(z_1).$$

But $e^{\frac{2\pi i}{m}}\phi(z_1) \in B(0,\delta)$ and so there exists $z_2 \neq z_1$(since ϕ is one to one) such that $\phi(z_2) = e^{\frac{2\pi i}{m}}\phi(z_1)$. But then

$$\phi(z_2)^m = \left(e^{\frac{2\pi i}{m}}\phi(z_1)\right)^m = \phi(z_1)^m$$

implying $f(z_2) = f(z_1)$ contradicting the assumption that f is one to one. Thus $m = 1$ and $f'(z) = \phi'(z) \neq 0$ on V. Since f maps open sets to open sets, it follows that f^{-1} is continuous and so we may write

$$\begin{aligned}\left(f^{-1}\right)'(f(z)) &= \lim_{f(z_1)\to f(z)} \frac{f^{-1}(f(z_1)) - f^{-1}(f(z))}{f(z_1) - f(z)} \\ &= \lim_{z_1\to z} \frac{z_1 - z}{f(z_1) - f(z)} = \frac{1}{f'(z)}.\end{aligned}$$

This proves the theorem.

One does not have to look very far to find that this sort of thing does not hold for functions mapping $\mathbb{R}$ to $\mathbb{R}$. Take for example, the function $f(x) = x^2$. Then $f(\mathbb{R})$ is neither a point nor a region. In fact $f(\mathbb{R})$ fails to be open.

26.3 Applications of the open mapping theorem

Definition 26.3 *We will denote by ρ a ray starting at 0. Thus ρ is a straight line of infinite length extending in one direction with its initial point at 0.*

As a simple application of the open mapping theorem, we give the following theorem about branches of the logarithm.

Theorem 26.4 *Let ρ be a ray starting at 0. Then there exists an analytic function, $L(z)$ defined on $\mathbb{C}\setminus\rho$ such that*

$$e^{L(z)} = z.$$

We call L a branch of the logarithm.

Proof: Let θ be an angle of the ray, ρ. The function, e^z is a one to one and onto mapping from $\mathbb{R} + i(\theta, \theta + 2\pi)$ to $\mathbb{C}\setminus\rho$ and so we may define $L(z)$ for $z \in \mathbb{C}\setminus\rho$ such that $e^{L(z)} = z$ and we see that L defined in this way is analytic on $\mathbb{C}\setminus\rho$ because of the open mapping theorem. Note we could just as well have considered $\mathbb{R} + i(\theta - 2\pi, \theta)$. This would have given another branch of the logarithm valid on $\mathbb{C}\setminus\rho$. Also, there are infinitely many choices for θ, each of which leads to a branch of the logarithm by the process just described.

Here is another very significant theorem known as the maximum modulus theorem which follows immediately from the open mapping theorem.

Theorem 26.5 *(maximum modulus theorem) Let U be a bounded region and let $f : U \to \mathbb{C}$ be analytic and $f : \overline{U} \to \mathbb{C}$ continuous. Then if $z \in U$,*

$$|f(z)| \leq \max\{|f(w)| : w \in \partial U\}. \tag{26.4}$$

If equality is achieved for any $z \in U$, then f is a constant.

Proof: Suppose f is not a constant. Then $f(U)$ is a region and so if $z \in U$, there exists $r > 0$ such that $B(f(z), r) \subseteq f(U)$. It follows there exists $z_1 \in U$ with $|f(z_1)| > |f(z)|$. Hence $\max\left\{|f(w)| : w \in \overline{U}\right\}$ is not achieved at any interior point of U. Therefore, the point at which the maximum is achieved must lie on the boundary of U and so

$$\max\{|f(w)| : w \in \partial U\} = \max\left\{|f(w)| : w \in \overline{U}\right\} > |f(z)|$$

for all $z \in U$ or else f is a constant. This proves the theorem.

26.4 Counting zeros

The above proof of the open mapping theorem relies on the very important inverse function theorem from real analysis. The proof features this and the Cauchy Riemann equations to indicate how the assumption f is analytic is used. There are other approaches to this important theorem which do not rely on the big theorems from real analysis and are more oriented toward the use of the Cauchy integral formula and specialized techniques from complex analysis. We give one of these approaches next which involves the notion of "counting zeros". The next theorem is the one about counting zeros. We will use the theorem later in the proof of the Riemann mapping theorem.

Theorem 26.6 *Let U be a region and let $\gamma : [a,b] \to U$ be closed, continuous, bounded variation, and $n(\gamma, z) = 0$ for all $z \notin U$. Suppose also that f is analytic on U having zeros $a_1, \cdots, a_m$ where the zeros are repeated according to multiplicity, and suppose that none of these zeros are on $\gamma([a,b])$. Then*

$$\frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} dz = \sum_{k=1}^m n(\gamma, a_k).$$

Proof: We are given $f(z) = \prod_{j=1}^m (z - a_j) g(z)$ where $g(z) \neq 0$ on U. Hence

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^m \frac{1}{z - a_j} + \frac{g'(z)}{g(z)}$$

and so

$$\frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} dz = \sum_{j=1}^m n(\gamma, a_j) + \frac{1}{2\pi i} \int_\gamma \frac{g'(z)}{g(z)} dz.$$

But the function, $z \to \frac{g'(z)}{g(z)}$ is analytic and so by Corollary 25.9, the last integral in the above expression equals 0. Therefore, this proves the theorem.

Theorem 26.7 *Let U be a region, let $\gamma : [a,b] \to U$ be continuous, closed and bounded variation such that $n(\gamma, z) = 0$ for all $z \notin U$. Also suppose $f : U \to \mathbb{C}$ be analytic and that $\alpha \notin f(\gamma([a,b]))$. Then $f \circ \gamma : [a,b] \to \mathbb{C}$ is continuous, closed, and bounded variation. Also suppose $\{a_1, \cdots, a_m\} = f^{-1}(\alpha)$ where these points are counted according to their multiplicities as zeros of the function $f - \alpha$ Then*

$$n(f \circ \gamma, \alpha) = \sum_{k=1}^m n(\gamma, a_k).$$

Proof: It is clear that $f \circ \gamma$ is closed and continuous. It only remains to verify that it is of bounded variation. Suppose first that $\gamma([a,b]) \subseteq B \subseteq \overline{B} \subseteq U$ where B is a ball. Then

$$\begin{aligned} &|f(\gamma(t)) - f(\gamma(s))| = \\ &\left| \int_0^1 f'(\gamma(s) + \lambda(\gamma(t) - \gamma(s)))(\gamma(t) - \gamma(s)) d\lambda \right| \\ \leq \quad & C|\gamma(t) - \gamma(s)| \end{aligned}$$

where $C \geq \max\left\{|f'(z)| : z \in \overline{B}\right\}$. Hence, in this case,

$$V(f \circ \gamma, [a,b]) \leq CV(\gamma, [a,b]).$$

Now let ε denote the distance between $\gamma([a,b])$ and $\mathbb{C}\setminus U$. Since $\gamma([a,b])$ is compact, $\varepsilon > 0$. By uniform continuity there exists $\delta = \frac{b-a}{p}$ for p a positive integer such that if $|s-t| < \delta$, then $|\gamma(s) - \gamma(t)| < \frac{\varepsilon}{2}$. Then

$$\gamma([t, t+\delta]) \subseteq \overline{B\left(\gamma(t), \frac{\varepsilon}{2}\right)} \subseteq U.$$

Let $C \geq \max\left\{|f'(z)| : z \in \cup_{j=1}^{p} \overline{B\left(\gamma(t_j), \frac{\varepsilon}{2}\right)}\right\}$ where $t_j \equiv \frac{j}{p}(b-a) + a$. Then from what was just shown,

$$\begin{aligned} V(f\circ\gamma, [a,b]) &\leq \sum_{j=0}^{p-1} V(f\circ\gamma, [t_j, t_{j+1}]) \\ &\leq C\sum_{j=0}^{p-1} V(\gamma, [t_j, t_{j+1}]) < \infty \end{aligned}$$

showing that $f\circ\gamma$ is bounded variation as claimed. Now from Theorem 25.6 there exists $\eta \in C^1([a,b])$ such that

$$\eta(a) = \gamma(a) = \gamma(b) = \eta(b), \ \eta([a,b]) \subseteq U,$$

and

$$n(\eta, a_k) = n(\gamma, a_k), \ n(f\circ\gamma, \alpha) = n(f\circ\eta, \alpha) \tag{26.5}$$

for $k = 1, \cdots, m$. Then

$$n(f\circ\gamma, \alpha) = n(f\circ\eta, \alpha)$$

$$\begin{aligned} &= \frac{1}{2\pi i}\int_{f\circ\eta} \frac{dw}{w-\alpha} \\ &= \frac{1}{2\pi i}\int_a^b \frac{f'(\eta(t))}{f(\eta(t)) - \alpha}\eta'(t)\,dt \\ &= \frac{1}{2\pi i}\int_\eta \frac{f'(z)}{f(z)-\alpha}dz \\ &= \sum_{k=1}^{m} n(\eta, a_k) \end{aligned}$$

By Theorem 26.6. By (26.5), this equals $\sum_{k=1}^{m} n(\gamma, a_k)$ which proves the theorem.

The next theorem is very interesting for its own sake.

Theorem 26.8 *Let $f : B(a,R) \to \mathbb{C}$ be analytic and let*

$$f(z) - \alpha = (z-a)^m g(z), \ \infty > m \geq 1$$

where $g(z) \neq 0$ in $B(a,R)$. ($f(z) - \alpha$ has a zero of order m at $z = a$.) Then there exist $\varepsilon, \delta > 0$ with the property that for each z satisfying $0 < |z-\alpha| < \delta$, there exist points,

$$\{a_1, \cdots, a_m\} \subseteq B(a, \varepsilon),$$

such that

$$f^{-1}(z) \cap B(a,\varepsilon) = \{a_1, \cdots, a_m\}$$

and each a_k is a zero of order 1 for the function $f(\cdot) - z$.

Proof: By Theorem 26.1 f is not constant on $B(a,R)$ because it has a zero of order m. Therefore, using this theorem again, there exists $\varepsilon > 0$ such that $\overline{B(a,2\varepsilon)} \subseteq B(a,R)$ and there are no solutions to the equation $f(z) - \alpha = 0$ for $z \in \overline{B(a,2\varepsilon)}$ except a. Also we may assume ε is small enough that for $0 < |z-a| \leq 2\varepsilon$, $f'(z) \neq 0$. Otherwise, a would be a limit point of a sequence of points, z_n, having $f'(z_n) = 0$ which would imply, by Theorem 26.1 that $f' = 0$ on $B(0,R)$, contradicting the assumption that f has a zero of order m and is therefore not constant.

Now pick $\gamma(t) = a + \varepsilon e^{it}, t \in [0,2\pi]$. Then $\alpha \notin f(\gamma([0,2\pi]))$ so there exists $\delta > 0$ with

$$B(\alpha,\delta) \cap f(\gamma([0,2\pi])) = \emptyset. \tag{26.6}$$

Therefore, $B(\alpha,\delta)$ is contained on one component of $\mathbb{C} \setminus f(\gamma([0,2\pi]))$. Therefore, $n(f\circ\gamma,\alpha) = n(f\circ\gamma,z)$ for all $z \in B(\alpha,\delta)$. Now consider f restricted to $\overline{B(a,2\varepsilon)}$. For $z \in B(\alpha,\delta)$, $f^{-1}(z)$ must consist of a finite set of points because $f'(w) \neq 0$ for all w in $\overline{B(a,2\varepsilon)} \setminus \{a\}$ implying that the zeros of $f(\cdot) - z$ in $\overline{B(a,2\varepsilon)}$ are isolated. Since $\overline{B(a,2\varepsilon)}$ is compact, this means there are only finitely many. By Theorem 26.7,

$$n(f\circ\gamma,z) = \sum_{k=1}^{p} n(\gamma,a_k) \tag{26.7}$$

where $\{a_1,\cdots,a_p\} = f^{-1}(z)$. Each point, a_k of $f^{-1}(z)$ is either inside the circle traced out by γ, yielding $n(\gamma,a_k) = 1$, or it is outside this circle yielding $n(\gamma,a_k) = 0$ because of (26.6). It follows the sum in (26.7) reduces to the number of points of $f^{-1}(z)$ which are contained in $B(a,\varepsilon)$. Thus, letting those points in $f^{-1}(z)$ which are contained in $B(a,\varepsilon)$ be denoted by $\{a_1,\cdots,a_r\}$

$$n(f\circ\gamma,\alpha) = n(f\circ\gamma,z) = r.$$

We need to verify that $r = m$. We do this by computing $n(f\circ\gamma,\alpha)$. However, this is easy to compute by Theorem 26.6 which states

$$n(f\circ\gamma,\alpha) = \sum_{k=1}^{m} n(\gamma,a) = m.$$

Therefore, $r = m$. Each of these a_k is a zero of order 1 of the function $f(\cdot) - z$ because $f'(a_k) \neq 0$. This proves the theorem.

This is a very fascinating result partly because it implies that for values of f near a value, α, at which $f(\cdot) - \alpha$ has a root of order m for $m > 1$, the inverse image of these values includes at least m points, not just one. Thus the topological properties of the inverse image changes radically. This theorem also shows that $f(B(a,\varepsilon)) \supseteq B(\alpha,\delta)$.

Theorem 26.9 *(open mapping theorem) Let U be a region and $f : U \to \mathbb{C}$ be analytic. Then $f(U)$ is either a point of a region. If f is one to one, then $f^{-1} : f(U) \to U$ is analytic.*

Proof: If f is not constant, then for every $\alpha \in f(U)$, it follows from Theorem 26.1 that $f(\cdot) - \alpha$ has a zero of order $m < \infty$ and so from Theorem 26.8 for each $a \in U$ there exist $\varepsilon, \delta > 0$ such that $f(B(a,\varepsilon)) \supseteq B(\alpha,\delta)$ which clearly implies that f maps open sets to open sets. Therefore, $f(U)$ is open, connected because f is continuous. If f is one to one, Theorem 26.8 implies that for every $\alpha \in f(U)$ the zero of $f(\cdot) - \alpha$ is of order 1. Otherwise, that theorem implies that for z near α, there are m points which f maps to z contradicting the assumption that f is one to one. Therefore, $f'(z) \neq 0$ and since f^{-1} is continuous, due to f being an open map, it follows we may write

$$\begin{aligned}
\left(f^{-1}\right)'(f(z)) &= \lim_{f(z_1)\to f(z)} \frac{f^{-1}(f(z_1)) - f^{-1}(f(z))}{f(z_1) - f(z)} \\
&= \lim_{z_1\to z} \frac{z_1 - z}{f(z_1) - f(z)} = \frac{1}{f'(z)}.
\end{aligned}$$

This proves the theorem.

26.5 Exercises

1. Use Theorem 26.6 to give an alternate proof of the fundamental theorem of algebra. **Hint:** Take a contour of the form $\gamma_r = re^{it}$ where $t \in [0, 2\pi]$. Consider $\int_{\gamma_r} \frac{p'(z)}{p(z)} dz$ and consider the limit as $r \to \infty$.

2. Prove the following version of the maximum modulus theorem. Let $f : U \to \mathbb{C}$ be analytic where U is a region. Suppose there exists $a \in U$ such that $|f(a)| \geq |f(z)|$ for all $z \in U$. Then f is a constant.

3. Let M be an $n \times n$ matrix. Recall that the eigenvalues of M are given by the zeros of the polynomial, $p_M(z) = \det(M - zI)$ where I is the $n \times n$ identity. Formulate a theorem which describes how the eigenvalues depend on small changes in M. **Hint:** You could define a norm on the space of $n \times n$ matrices as $||M|| \equiv tr\ (MM^*)^{1/2}$ where M^* is the conjugate transpose of M. Thus
$$||M|| = \left(\sum_{j,k} |M_{jk}|^2 \right)^{1/2}.$$
Argue that small changes will produce small changes in $p_M(z)$. Then apply Theorem 26.6 using γ_k a very small circle surrounding z_k, the kth eigenvalue.

4. Suppose that two analytic functions defined on a region are equal on some set, S which contains a limit point. (Recall p is a limit point of S if every open set which contains p, also contains infinitely many points of S.) Show the two functions coincide. We defined $e^z \equiv e^x(\cos y + i \sin y)$ earlier and we showed that e^z, defined this way was analytic on $\mathbb{C}$. Is there any other way to define e^z on all of $\mathbb{C}$ such that the function coincides with e^x on the real axis?

5. We know various identities for real valued functions. For example $\cosh^2 x - \sinh^2 x = 1$. If we define $\cosh z \equiv \frac{e^z + e^{-z}}{2}$ and $\sinh z \equiv \frac{e^z - e^{-z}}{2}$, does it follow that
$$\cosh^2 z - \sinh^2 z = 1$$
for all $z \in \mathbb{C}$? What about
$$\sin(z + w) = \sin z \cos w + \cos z \sin w?$$
Can you verify these sorts of identities just from your knowledge about what happens for real arguments?

6. Was it necessary that U be a region in Theorem 26.1? Would the same conclusion hold if U were only assumed to be an open set? Why? What about the open mapping theorem? Would it hold if U were not a region?

7. Let $f : U \to \mathbb{C}$ be analytic and one to one. Show that $f'(z) \neq 0$ for all $z \in U$. Does this hold for a function of a real variable?

8. We say a real valued function, u is subharmonic if $u_{xx} + u_{yy} \geq 0$. Show that if u is subharmonic on a bounded region, (open connected set) U, and continuous on $\overline{U}$ and $u \leq m$ on ∂U, then $u \leq m$ on U. **Hint:** If not, u achieves its maximum at $(x_0, y_0) \in U$. Let $u(x_0, y_0) > m + \delta$ where $\delta > 0$. Now consider $u_\varepsilon(x, y) = \varepsilon x^2 + u(x, y)$ where ε is small enough that $0 < \varepsilon x^2 < \delta$ for all $(x, y) \in U$. Show that u_ε also achieves its maximum at some point of U and that therefore, $u_{\varepsilon xx} + u_{\varepsilon yy} \leq 0$ at that point implying that $u_{xx} + u_{yy} \leq -\varepsilon$, a contradiction.

9. If u is harmonic on some region, U, show that u coincides locally with the real part of an analytic function and that therefore, u has infinitely many derivatives on U. **Hint:** Consider the case where $0 \in U$. You can always reduce to this case by a suitable translation. Now let $B(0, r) \subseteq U$ and use the Schwarz formula to obtain an analytic function whose real part coincides with u on $\partial B(0, r)$. Then use Problem 8.

10. Show the solution to the Dirichlet problem of Problem 8 in the section on the Cauchy integral formula for a disk is unique. You need to formulate this precisely and then prove uniqueness.

Singularities

27.1 The Laurent series

In this chapter we consider the functions which are analytic in some open set except at isolated points. The fundamental formula in this subject which is used to classify isolated singularities is the Laurent series.

Definition 27.1 *We define* $\operatorname{ann}(a, R_1, R_2) \equiv \{z : R_1 < |z-a| < R_2\}$.

Thus $\operatorname{ann}(a, 0, R)$ would denote the punctured ball, $B(a, R) \setminus \{0\}$. We now consider an important lemma which will be used in what follows.

Lemma 27.2 *Let* g *be analytic on* $\operatorname{ann}(a, R_1, R_2)$. *Then if* $\gamma_r(t) \equiv a + re^{it}$ *for* $t \in [0, 2\pi]$ *and* $r \in (R_1, R_2)$, *then* $\int_{\gamma_r} g(z)\, dz$ *is independent of* r.

Proof: Let $R_1 < r_1 < r_2 < R_2$ and denote by $-\gamma_r(t)$ the curve, $-\gamma_r(t) \equiv a + re^{i(2\pi - t)}$ for $t \in [0, 2\pi]$. Then if $z \in \overline{B(a, R_1)}$, we can apply Proposition 25.7 to conclude $n(-\gamma_{r_1}, z) + n(\gamma_{r_2}, z) = 0$. Also if $z \notin B(a, R_2)$, then by Corollary 25.11 we have $n\left(\gamma_{r_j}, z\right) = 0$ for $j = 1, 2$. Therefore, we can apply Theorem 25.8 and conclude that for all $z \in \operatorname{ann}(a, R_1, R_2) \setminus \cup_{j=1}^{2} \gamma_{r_j}([0, 2\pi])$,

$$0\left(n\left(\gamma_{r_2}, z\right) + n\left(-\gamma_{r_1}, z\right)\right) =$$

$$\frac{1}{2\pi i} \int_{\gamma_{r_2}} \frac{g(w)(w-z)}{w-z} dw - \frac{1}{2\pi i} \int_{\gamma_{r_1}} \frac{g(w)(w-z)}{w-z} dw$$

which proves the desired result.

With this preparation we are ready to discuss the Laurent series.

Theorem 27.3 *Let* f *be analytic on* $\operatorname{ann}(a, R_1, R_2)$. *Then there exist numbers,* $a_n \in \mathbb{C}$ *such that for all* $z \in \operatorname{ann}(a, R_1, R_2)$,

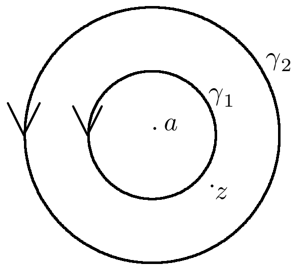
$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z-a)^n, \tag{27.1}$$

where the series converges absolutely and uniformly on $\overline{\operatorname{ann}(a, r_1, r_2)}$ *whenever* $R_1 < r_1 < r_2 < R_2$. *Also*

$$a_n = \frac{1}{2\pi i} \int_{\gamma} \frac{f(w)}{(w-a)^{n+1}} dw \tag{27.2}$$

where $\gamma(t) = a + re^{it}, t \in [0, 2\pi]$ *for any* $r \in (R_1, R_2)$. *Furthermore the series is unique in the sense that if (27.1) holds for* $z \in \operatorname{ann}(a, R_1, R_2)$, *then we obtain (27.2).*

Proof: Let $R_1 < r_1 < r_2 < R_2$ and define $\gamma_1(t) \equiv a + (r_1 - \varepsilon) e^{it}$ and $\gamma_2(t) \equiv a + (r_2 + \varepsilon) e^{it}$ for $t \in [0, 2\pi]$ and ε chosen small enough that $R_1 < r_1 - \varepsilon < r_2 + \varepsilon < R_2$.



Then by Proposition 25.7 and Corollary 25.11, we see that

$$n(-\gamma_1, z) + n(\gamma_2, z) = 0$$

off $\operatorname{ann}(a, R_1, R_2)$ and that on $\operatorname{ann}(a, r_1, r_2)$,

$$n(-\gamma_1, z) + n(\gamma_2, z) = 1.$$

Therefore, by Theorem 25.8,

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\left[\int_{-\gamma_1} \frac{f(w)}{w-z} dw + \int_{\gamma_2} \frac{f(w)}{w-z} dw\right] \\ &= \frac{1}{2\pi i}\left[\int_{\gamma_1} \frac{f(w)}{(z-a)\left[1-\frac{w-a}{z-a}\right]} dw + \int_{\gamma_2} \frac{f(w)}{(w-a)\left[1-\frac{z-a}{w-a}\right]} dw\right] \end{aligned}$$

$$= \frac{1}{2\pi i}\int_{\gamma_2} \frac{f(w)}{w-a} \sum_{n=0}^{\infty}\left(\frac{z-a}{w-a}\right)^n dw +$$

$$\frac{1}{2\pi i}\int_{\gamma_1} \frac{f(w)}{(z-a)} \sum_{n=0}^{\infty}\left(\frac{w-a}{z-a}\right)^n dw. \tag{27.3}$$

From the formula (27.3), it follows that for $z \in \overline{\operatorname{ann}(a, r_1, r_2)}$, the terms in the first sum are bounded by an expression of the form $C\left(\frac{r_2}{r_2+\varepsilon}\right)^n$ while those in the second are bounded by one of the form $C\left(\frac{r_1-\varepsilon}{r_1}\right)^n$ and so by the Weierstrass M test, the convergence is uniform and so we may interchange the integrals and the sums in the above formula and rename the variable of summation to obtain

$$f(z) = \sum_{n=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma_2} \frac{f(w)}{(w-a)^{n+1}} dw\right)(z-a)^n +$$

$$\sum_{n=-\infty}^{-1}\left(\frac{1}{2\pi i}\int_{\gamma_1} \frac{f(w)}{(w-a)^{n+1}}\right)(z-a)^n. \tag{27.4}$$

By Lemma 27.2, we may write this as

$$f(z) = \sum_{n=-\infty}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma_r} \frac{f(w)}{(w-a)^{n+1}} dw\right)(z-a)^n.$$

where $r \in (R_1, R_2)$ is arbitrary.

If $f(z) = \sum_{n=-\infty}^{\infty} a_n (z-a)^n$ on $\operatorname{ann}(a, R_1, R_2)$ let

$$f_n(z) \equiv \sum_{k=-n}^{n} a_k (z-a)^k \tag{27.5}$$

and verify from a repeat of the above argument that

$$f_n(z) = \sum_{k=-\infty}^{\infty} \left(\frac{1}{2\pi i} \int_{\gamma_r} \frac{f_n(w)}{(w-a)^{k+1}} dw \right) (z-a)^k . \tag{27.6}$$

Therefore, using (27.5) directly, we see

$$\frac{1}{2\pi i} \int_{\gamma_r} \frac{f_n(w)}{(w-a)^{k+1}} dw = a_k$$

for each $k \in [-n, n]$. However,

$$\frac{1}{2\pi i} \int_{\gamma_r} \frac{f_n(w)}{(w-a)^{k+1}} dw = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{(w-a)^{k+1}} dw$$

because if $l > n$ or $l < -n$, then it is easy to verify that

$$\int_{\gamma_r} \frac{a_l (w-a)^l}{(w-a)^{k+1}} dw = 0$$

for all $k \in [-n, n]$. Therefore,

$$a_k = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(w)}{(w-a)^{k+1}} dw$$

and so this establishes uniqueness. This proves the theorem.

Definition 27.4 *We say f has an isolated singularity at $a \in \mathbb{C}$ if there exists $R > 0$ such that f is analytic on* $\operatorname{ann}(a, 0, R)$. *Such an isolated singularity is said to be a pole of order m if $a_{-m} \neq 0$ but $a_k = 0$ for all $k < m$. The singularity is said to be removable if $a_n = 0$ for all $n < 0$, and it is said to be essential if $a_m \neq 0$ for infinitely many $m < 0$.*

Note that thanks to the Laurent series, the possibilities enumerated in the above definition are the only ones possible. Also observe that a is removable if and only if $f(z) = g(z)$ for some g analytic near a. How can we recognize a removable singularity or a pole without computing the Laurent series? This is the content of the next theorem.

Theorem 27.5 *Let a be an isolated singularity of f. Then a is removable if and only if*

$$\lim_{z \to a} (z-a) f(z) = 0 \tag{27.7}$$

and a is a pole if and only if

$$\lim_{z \to a} |f(z)| = \infty. \tag{27.8}$$

The pole is of order m if

$$\lim_{z \to a} (z-a)^{m+1} f(z) = 0$$

but

$$\lim_{z \to a} (z-a)^m f(z) \neq 0.$$

Proof: First suppose a is a removable singularity. Then it is clear that (27.7) holds since $a_m = 0$ for all $m < 0$. Now suppose that (27.7) holds and f is analytic on $\operatorname{ann}(a, 0, R)$. Then define

$$h(z) \equiv \begin{cases} (z-a) f(z) & \text{if } z \neq a \\ 0 \text{ if } z = a \end{cases}$$

We verify that h is analytic near a by using Morera's theorem. Let T be a triangle in $B(a, R)$. If T does not contain the point, a, then Corollary 25.11 implies $\int_{\partial T} h(z)\, dz = 0$. Therefore, we may assume $a \in T$. If a is a vertex, then, denoting by b and c the other two vertices, we pick p and q, points on the sides, ab and ac respectively which are close to a. Then by Corollary 25.11,

$$\int_{\gamma(q,c,b,p,q)} h(z)\, dz = 0.$$

But by continuity of h, it follows that as p and q are moved closer to a the above integral converges to $\int_{\partial T} h(z)\, dz$, showing that in this case, $\int_{\partial T} h(z)\, dz = 0$ also. It only remains to consider the case where a is not a vertex but is in T. In this case we subdivide the triangle T into either 3 or 2 subtriangles having a as one vertex, depending on whether a is in the interior or on an edge. Then, applying the above result to these triangles and noting that the integrals over the interior edges cancel out due to the integration being taken in opposite directions, we see that $\int_{\partial T} h(z)\, dz = 0$ in this case also.

Now we know h is analytic. Since h equals zero at a, we can conclude that

$$h(z) = (z-a) g(z)$$

where $g(z)$ is analytic in $B(a, R)$. Therefore, for all $z \neq a$,

$$(z-a) g(z) = (z-a) f(z)$$

showing that $f(z) = g(z)$ for all $z \neq a$ and g is analytic on $B(0, R)$. This proves the converse.

It is clear that if f has a pole at a, then (27.8) holds. Suppose conversely that (27.8) holds. Then we know from the first part of this theorem that $1/f(z)$ has a removable singularity at a. Also, if $g(z) = 1/f(z)$ for z near a, then $g(a) = 0$. Therefore, for $z \neq a$,

$$1/f(z) = (z-a)^m h(z)$$

for some analytic function, $h(z)$ for which $h(a) \neq 0$. It follows that $1/h \equiv r$ is analytic near a with $r(a) \neq 0$. Therefore, for z near a,

$$f(z) = (z-a)^{-m} \sum_{k=0}^{\infty} a_k (z-a)^k,\ a_0 \neq 0,$$

showing that f has a pole of order m. This proves the theorem.

Note that this is very different than what occurs for functions of a real variable. Consider for example, the function, $f(x) = x^{-1/2}$. We see $x\left(|x|^{-1/2}\right) \to 0$ but clearly $|x|^{-1/2}$ cannot equal a differentiable function near 0.

What about rational functions, those which are a quotient of two polynomials? It seems reasonable to suppose, since every finite partial sum of the Laurent series is a rational function just as every finite sum of a power series is a polynomial, it might be the case that something interesting can be said about rational functions in the context of Laurent series. In fact we will show the existence of the partial fraction expansion for rational functions. First we need the following simple lemma.

Lemma 27.6 *If f is a rational function which has no poles in $\mathbb{C}$ then f is a polynomial.*

Proof: We can write

$$f(z) = \frac{p_0 (z-b_1)^{l_1} \cdots (z-b_n)^{l_n}}{(z-a_1)^{r_1} \cdots (z-a_m)^{r_m}},$$

where we can assume the fraction has been reduced to lowest terms. Thus none of the b_j equal any of the a_k. But then, by Theorem 27.5 we would have poles at each a_k. Therefore, the denominator must reduce to 1 and so f is a polynomial.

Theorem 27.7 *Let $f(z)$ be a rational function,*

$$f(z) = \frac{p_0 (z-b_1)^{l_1} \cdots (z-b_n)^{l_n}}{(z-a_1)^{r_1} \cdots (z-a_m)^{r_m}}, \tag{27.9}$$

where the expression is in lowest terms. Then there exist numbers, b_j^k and a polynomial, $p(z)$, such that

$$f(z) = \sum_{l=1}^{m} \sum_{j=1}^{r_l} \frac{b_j^l}{(z-a_l)^j} + p(z). \tag{27.10}$$

Proof: We see that f has a pole at a_1 and it is clear this pole must be of order r_1 since otherwise we could not achieve equality between (27.9) and the Laurent series for f near a_1 due to different rates of growth. Therefore, for $z \in \operatorname{ann}(a_1, 0, R_1)$

$$f(z) = \sum_{j=1}^{r_1} \frac{b_j^1}{(z-a_1)^j} + p_1(z)$$

where p_1 is analytic in $B(a_1, R_1)$. Then define

$$f_1(z) \equiv f(z) - \sum_{j=1}^{r_1} \frac{b_j^1}{(z-a_1)^j}$$

so that f_1 is a rational function coinciding with p_1 near a_1 which has no pole at a_1. We see that f_1 has a pole at a_2 or order r_2 by the same reasoning. Therefore, we may subtract off the principle part of the Laurent series for f_1 near a_2 like we just did for f. This yields

$$f(z) = \sum_{j=1}^{r_1} \frac{b_j^1}{(z-a_1)^j} + \sum_{j=1}^{r_2} \frac{b_j^2}{(z-a_2)^j} + p_2(z).$$

Letting

$$f(z) - \left(\sum_{j=1}^{r_1} \frac{b_j^1}{(z-a_1)^j} + \sum_{j=1}^{r_2} \frac{b_j^2}{(z-a_2)^j} \right) = f_2(z),$$

and continuing in this way we finally obtain

$$f(z) - \sum_{l=1}^{m} \sum_{j=1}^{r_l} \frac{b_j^l}{(z-a_l)^j} = f_m(z)$$

where f_m is a rational function which has no poles. Therefore, it must be a polynomial. This proves the theorem.

How does this relate to the usual partial fractions routine of calculus? Recall in that case we had to consider irreducible quadratics and all the constants were real. In the case from calculus, since the coefficients of the polynomials were real, the roots of the denominator occurred in conjugate pairs. Thus we would have paired terms like

$$\frac{b}{(z-\overline{a})^j}+\frac{c}{(z-a)^j}$$

occurring in the sum. We leave it to the reader to verify this version of partial fractions does reduce to the version from calculus.

We have considered the case of a removable singularity or a pole and proved theorems about this case. What about the case where the singularity is essential? We give an interesting theorem about this case next.

Theorem 27.8 *(Casorati Weierstrass) If f has an essential singularity at a then for all $r>0$,*

$$\overline{f(\operatorname{ann}(a,0,r))}=\mathbb{C}$$

Proof: If not there exists $c\in\mathbb{C}$ and $r>0$ such that $c\notin\overline{f(\operatorname{ann}(a,0,r))}$. Therefore,there exists $\varepsilon>0$ such that $B(c,\varepsilon)\cap f(\operatorname{ann}(a,0,r))=\emptyset$. It follows that

$$\lim_{z\to a}|z-a|^{-1}|f(z)-c|=\infty$$

and so by Theorem 27.5 $z\to(z-a)^{-1}(f(z)-c)$ has a pole at a. It follows that for m the order of the pole,

$$(z-a)^{-1}(f(z)-c)=\sum_{k=1}^{m}\frac{a_k}{(z-a)^k}+g(z)$$

where g is analytic near a. Therefore,

$$f(z)-c=\sum_{k=1}^{m}\frac{a_k}{(z-a)^{k-1}}+g(z)(z-a),$$

showing that f has a pole at a rather than an essential singularity. This proves the theorem.

This theorem is much weaker than the best result known, the Picard theorem which we state next. A proof of this famous theorem may be found in Conway [6].

Theorem 27.9 *If f is an analytic function having an essential singularity at z, then in every open set containing z the function f, assumes each complex number, with one possible exception, an infinite number of times.*

27.2 Exercises

1. Classify the singular points of the following functions according to whether they are poles or essential singularities. If poles, determine the order of the pole.

   (a) $\frac{\cos z}{z^2}$

   (b) $\frac{z^3+1}{z(z-1)}$

   (c) $\cos\left(\frac{1}{z}\right)$

2. Suppose f is defined on an open set, U, and it is known that f is analytic on $U\setminus\{z_0\}$ but continuous at z_0. Show that f is actually analytic on U.

3. A function defined on $\mathbb{C}$ has finitely many poles and $\lim_{|z|\to\infty} f(z)$ exists. Show f is a rational function. **Hint:** First show that if h has only one pole at 0 and if $\lim_{|z|\to\infty} h(z)$ exists, then h is a rational function. Now consider

$$h(z) \equiv \frac{\prod_{k=1}^{m} (z - z_k)^{r_k}}{\prod_{k=1}^{m} z^{r_k}} f(z)$$

where z_k is a pole of order r_k.

Residues and evaluation of integrals

It turns out that the theory presented above about singularities and the Laurent series is very useful in computing the exact value of many hard integrals. First we define what we mean by a residue.

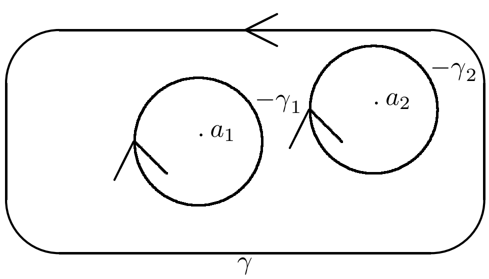
Definition 28.1 *Let a be an isolated singularity of f. Thus*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z-a)^n$$

for all z near a. Then we define the residue of f at a by

$$\text{Res}(f,a) = a_{-1}.$$

Now suppose that U is an open set and $f : U \setminus \{a_1, \cdots, a_m\} \to \mathbb{C}$ is analytic where the a_k are isolated singularities of f.



Let γ be a simple closed continuous, and bounded variation curve enclosing these isolated singularities such that $\gamma([a,b]) \subseteq U$ and $\{a_1, \cdots, a_m\} \subseteq D \subseteq U$, where D is the bounded component (inside) of $\mathbb{C} \setminus \gamma([a,b])$. Also assume $n(\gamma, z) = 1$ for all $z \in D$. As explained earlier, this would occur if $\gamma(t)$ traces out the curve in the counter clockwise direction. Choose r small enough that $B(a_j, r) \cap B(a_k, r) = \emptyset$ whenever $j \neq k$, $B(a_k, r) \subseteq U$ for all k, and define

$$-\gamma_k(t) \equiv a_k + re^{(2\pi - t)i},\ t \in [0, 2\pi].$$

Thus $n(-\gamma_k, a_i) = -1$ and if z is in the unbounded component of $\mathbb{C} \setminus \gamma([a,b])$, $n(\gamma, z) = 0$ and $n(-\gamma_k, z) = 0$. If $z \notin U \setminus \{a_1, \cdots, a_m\}$, then z either equals one of the a_k or else z is in the unbounded component just described. Either way, $\sum_{k=1}^{m} n(\gamma_k, z) + n(\gamma, z) = 0$. Therefore, by Theorem 25.8, if $z \notin D$,

$$\begin{aligned}
\sum_{j=1}^{m} \frac{1}{2\pi i} \int_{-\gamma_j} f(w) \frac{(w-z)}{(w-z)} dw + \frac{1}{2\pi i} \int_{\gamma} f(w) \frac{(w-z)}{(w-z)} dw &= \\
\sum_{j=1}^{m} \frac{1}{2\pi i} \int_{-\gamma_j} f(w)\, dw + \frac{1}{2\pi i} \int_{\gamma} f(w)\, dw &= \\
\left(\sum_{k=1}^{m} n(-\gamma_k, z) + n(\gamma, z) \right) f(z)(z-z) &= 0.
\end{aligned}$$

and so, taking r small enough,

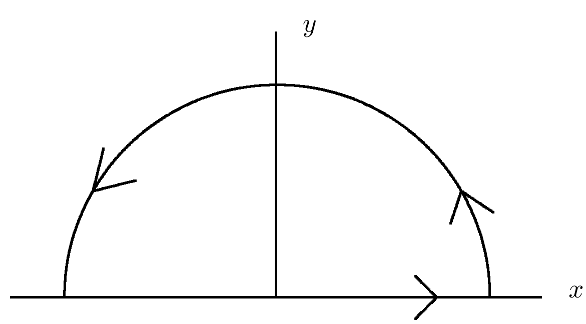
$$\begin{aligned}
\frac{1}{2\pi i} \int_{\gamma} f(w)\, dw &= \sum_{j=1}^{m} \frac{1}{2\pi i} \int_{\gamma_j} f(w)\, dw \\
&= \frac{1}{2\pi i} \sum_{k=1}^{m} \sum_{l=-\infty}^{\infty} a_l^k \int_{\gamma_k} (w-a_k)^l\, dw \\
&= \frac{1}{2\pi i} \sum_{k=1}^{m} a_{-1}^k \int_{\gamma_k} (w-a_k)^{-1}\, dw \\
&= \sum_{k=1}^{m} a_{-1}^k = \sum_{k=1}^{m} \operatorname{Res}(f, a_k).
\end{aligned}$$

Now we give some examples of hard integrals which can be evaluated by using this idea. This will be done by integrating over various closed curves having bounded variation.

Example 28.2 *The first example we consider is the following integral.*

$$\int_{-\infty}^{\infty} \frac{1}{1+x^4} dx$$

One could imagine evaluating this integral by the method of partial fractions and it should work out by that method. However, we will consider the evaluation of this integral by the method of residues instead. To do so, consider the following picture.



Let $\gamma_r(t) = re^{it}, t \in [0, \pi]$ and let $\sigma_r(t) = t : t \in [-r, r]$. Thus γ_r parameterizes the top curve and σ_r parameterizes the straight line from $-r$ to r along the x axis. Denoting by Γ_r the closed curve traced out by these two, we see from simple estimates that

$$\lim_{r\to\infty}\int_{\gamma_r}\frac{1}{1+z^4}dz=0.$$

This follows from the following estimate.

$$\left|\int_{\gamma_r}\frac{1}{1+z^4}dz\right|\leq\frac{1}{r^4-1}\pi r.$$

Therefore,

$$\int_{-\infty}^{\infty}\frac{1}{1+x^4}dx=\lim_{r\to\infty}\int_{\Gamma_r}\frac{1}{1+z^4}dz.$$

We compute $\int_{\Gamma_r}\frac{1}{1+z^4}dz$ using the method of residues. The only residues of the integrand are located at points, z where $1+z^4=0$. These points are

$$\begin{aligned} z &= -\frac{1}{2}\sqrt{2}-\frac{1}{2}i\sqrt{2}, z=\frac{1}{2}\sqrt{2}-\frac{1}{2}i\sqrt{2}, \\ z &= \frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}, z=-\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2} \end{aligned}$$

and it is only the last two which are found in the inside of Γ_r. Therefore, we need to calculate the residues at these points. Clearly this function has a pole of order one at each of these points and so we may calculate the residue at α in this list by evaluating

$$\lim_{z\to\alpha}(z-\alpha)\frac{1}{1+z^4}$$

Thus

$$\begin{aligned} \operatorname{Res}\left(f,\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}\right) &= \\ \lim_{z\to\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}}\left(z-\left(\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}\right)\right)\frac{1}{1+z^4} &= -\frac{1}{8}\sqrt{2}-\frac{1}{8}i\sqrt{2} \end{aligned}$$

Similarly we may find the other residue in the same way

$$\begin{aligned} \operatorname{Res}\left(f,-\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}\right) &= \\ \lim_{z\to-\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}}\left(z-\left(-\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}\right)\right)\frac{1}{1+z^4} &= -\frac{1}{8}i\sqrt{2}+\frac{1}{8}\sqrt{2}. \end{aligned}$$

Therefore,

$$\int_{\Gamma_r}\frac{1}{1+z^4}dz=2\pi i\left(-\frac{1}{8}i\sqrt{2}+\frac{1}{8}\sqrt{2}+\left(-\frac{1}{8}\sqrt{2}-\frac{1}{8}i\sqrt{2}\right)\right)=\frac{1}{2}\pi\sqrt{2}.$$

Thus, taking the limit we obtain $\frac{1}{2}\pi\sqrt{2}=\int_{-\infty}^{\infty}\frac{1}{1+x^4}dx$.

Obviously many different variations of this are possible. The main idea being that the integral over the semicircle converges to zero as $r\to\infty$. Sometimes one must be fairly creative to determine the sort of curve to integrate over as well as the sort of function in the integrand and even the interpretation of the integral which results.

Example 28.3 *This example illustrates the comment about the integral.*

$$\int_0^\infty \frac{\sin x}{x} dx$$

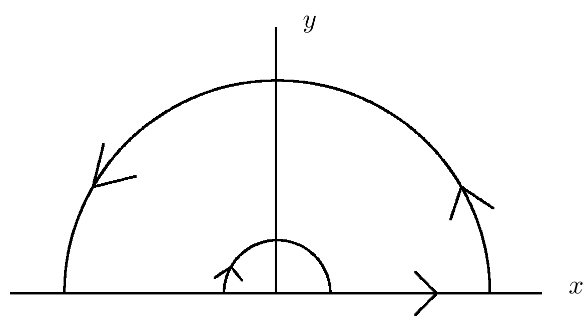
By this integral we mean $\lim_{r\to\infty} \int_0^r \frac{\sin x}{x} dx$. The function is not absolutely integrable so the meaning of the integral is in terms of the limit just described. To do this integral, we note the integrand is even and so it suffices to find

$$\lim_{R\to\infty} \int_{-R}^R \frac{e^{ix}}{x} dx$$

called the Cauchy principle value, take the imaginary part to get

$$\lim_{R\to\infty} \int_{-R}^R \frac{\sin x}{x} dx$$

and then divide by two. In order to do so, we let $R > r$ and consider the curve which goes along the x axis from $(-R,0)$ to $(-r,0)$, from $(-r,0)$ to $(r,0)$ along the semicircle in the upper half plane, from $(r,0)$ to $(R,0)$ along the x axis, and finally from $(R,0)$ to $(-R,0)$ along the semicircle in the upper half plane as shown in the following picture.



On the inside of this curve, the function, $\frac{e^{iz}}{z}$ has no singularities and so it has no residues. Pick R large and let $r \to 0+$. The integral along the small semicircle is

$$\int_\pi^0 \frac{e^{re^{it}} rie^{it}}{re^{it}} dt = \int_\pi^0 e^{\left(re^{it}\right)} dt.$$

and this clearly converges to $-\pi$ as $r \to 0$. Now we consider the top integral. For $z = Re^{it}$,

$$e^{iRe^{it}} = e^{-R\sin t} \cos(R\cos t) + ie^{-R\sin t} \sin(R\cos t)$$

and so

$$\left| e^{iRe^{it}} \right| \le e^{-R\sin t}.$$

Therefore, along the top semicircle we get the absolute value of the integral along the top is,

$$\left| \int_0^\pi e^{iRe^{it}} dt \right| \le \int_0^\pi e^{-R\sin t} dt$$

$$\begin{aligned}&\le \int_{\delta}^{\pi-\delta} e^{-R\sin\delta}dt+\int_{\pi-\delta}^{\pi} e^{-R\sin t}dt+\int_{0}^{\delta} e^{-R\sin t}dt\\ &\le e^{-R\sin\delta}\pi+\varepsilon\end{aligned}$$

whenever δ is small enough. Letting δ be this small, it follows that

$$\lim_{R\to\infty}\left|\int_0^{\pi} e^{iRe^{it}}dt\right|\le\varepsilon$$

and since ε is arbitrary, this shows the integral over the top semicircle converges to 0. Therefore, for some function $e(r)$ which converges to zero as $r\to 0$,

$$e(r)=\int_{\text{top semicircle}}\frac{e^{iz}}{z}dz-\pi+\int_r^R\frac{e^{ix}}{x}dx+\int_{-R}^{-r}\frac{e^{ix}}{x}dx$$

Letting $r\to 0$, we see

$$\pi=\int_{\text{top semicircle}}\frac{e^{iz}}{z}dz+\int_{-R}^{R}\frac{e^{ix}}{x}dx$$

and so, taking $R\to\infty$,

$$\pi=\lim_{R\to\infty}\int_{-R}^{R}\frac{e^{ix}}{x}dx=2\lim_{R\to\infty}\int_0^R\frac{\sin x}{x},$$

showing that $\frac{\pi}{2}=\int_0^{\infty}\frac{\sin x}{x}dx$ with the above interpretation of the integral.

Sometimes we don't blow up the curves and take limits. Sometimes the problem of interest reduces directly to a complex integral over a closed curve. Here is an example of this.

Example 28.4 *The integral is*

$$\int_0^{\pi}\frac{\cos\theta}{2+\cos\theta}d\theta$$

This integrand is even and so we may write it as

$$\frac{1}{2}\int_{-\pi}^{\pi}\frac{\cos\theta}{2+\cos\theta}d\theta.$$

For z on the unit circle, $z=e^{i\theta}$, $\overline{z}=\frac{1}{z}$ and therefore, $\cos\theta=\frac{1}{2}\left(z+\frac{1}{z}\right)$. Thus $dz=ie^{i\theta}d\theta$ and so $d\theta=\frac{dz}{iz}$. Note that we are proceeding formally in order to get a complex integral which reduces to the one of interest. It follows that a complex integral which reduces to the one we want is

$$\frac{1}{2i}\int_{\gamma}\frac{\frac{1}{2}\left(z+\frac{1}{z}\right)}{2+\frac{1}{2}\left(z+\frac{1}{z}\right)}\frac{dz}{z}=\frac{1}{2i}\int_{\gamma}\frac{z^2+1}{z\left(4z+z^2+1\right)}dz$$

where γ is the unit circle. Now the integrand has poles of order 1 at those points where $z\left(4z+z^2+1\right)=0$. These points are

$$0,-2+\sqrt{3},-2-\sqrt{3}.$$

Only the first two are inside the unit circle. It is also clear the function has simple poles at these points. Therefore,

$$\operatorname{Res}(f,0)=\lim_{z\to 0}z\left(\frac{z^2+1}{z\left(4z+z^2+1\right)}\right)=1.$$

$$\operatorname{Res}\left(f, -2+\sqrt{3}\right) =$$

$$\lim_{z\to -2+\sqrt{3}} \left(z - \left(-2+\sqrt{3}\right)\right) \frac{z^2+1}{z\left(4z+z^2+1\right)} = -\frac{2}{3}\sqrt{3}.$$

It follows

$$\begin{aligned} \int_0^{\pi} \frac{\cos\theta}{2+\cos\theta} d\theta &= \frac{1}{2i} \int_{\gamma} \frac{z^2+1}{z\left(4z+z^2+1\right)} dz \\ &= \frac{1}{2i} 2\pi i \left(1 - \frac{2}{3}\sqrt{3}\right) \\ &= \pi \left(1 - \frac{2}{3}\sqrt{3}\right). \end{aligned}$$

Other rational functions of the trig functions will work out by this method also.

Sometimes we have to be clever about which version of an analytic function that reduces to a real function we should use. The following is such an example.

Example 28.5 *The integral here is*

$$\int_0^{\infty} \frac{\ln x}{1+x^4} dx.$$

We would like to use the same curve we used in the integral involving $\frac{\sin x}{x}$ but this will create problems with the log since the usual version of the log is not defined on the negative real axis. This does not need to concern us however. We simply use another branch of the logarithm. We leave out the ray from 0 along the negative y axis and use Theorem 26.4 to define $L(z)$ on this set. Thus $L(z) = \ln|z| + i\arg_1(z)$ where $\arg_1(z)$ will be the angle, θ, between $-\frac{\pi}{2}$ and $\frac{3\pi}{2}$ such that $z = |z| e^{i\theta}$. Now the only singularities contained in this curve are

$$\frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}, -\frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}$$

and the integrand, f has simple poles at these points. Thus using the same procedure as in the other examples,

$$\operatorname{Res}\left(f, \frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}\right) =$$

$$\frac{1}{32}\sqrt{2}\pi - \frac{1}{32}i\sqrt{2}\pi$$

and

$$\operatorname{Res}\left(f, \frac{-1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}\right) =$$

$$\frac{3}{32}\sqrt{2}\pi + \frac{3}{32}i\sqrt{2}\pi.$$

We need to consider the integral along the small semicircle of radius r. This reduces to

$$\int_{\pi}^{0} \frac{\ln|r| + it}{1+\left(re^{it}\right)^4} \left(rie^{it}\right) dt$$

which clearly converges to zero as $r \to 0$ because $r \ln r \to 0$. Therefore, taking the limit as $r \to 0$,

$$\int_{\text{large semicircle}} \frac{L(z)}{1+z^4} dz + \lim_{r \to 0+} \int_{-R}^{-r} \frac{\ln(-t) + i\pi}{1+t^4} dt +$$

$$\lim_{r \to 0+} \int_r^R \frac{\ln t}{1+t^4} dt = 2\pi i \left(\frac{3}{32}\sqrt{2}\pi + \frac{3}{32} i\sqrt{2}\pi + \frac{1}{32}\sqrt{2}\pi - \frac{1}{32} i\sqrt{2}\pi \right).$$

Observing that $\int_{\text{large semicircle}} \frac{L(z)}{1+z^4} dz \to 0$ as $R \to \infty$, we may write

$$e(R) + 2 \lim_{r \to 0+} \int_r^R \frac{\ln t}{1+t^4} dt + i\pi \int_{-\infty}^0 \frac{1}{1+t^4} dt = \left(-\frac{1}{8} + \frac{1}{4} i \right) \pi^2 \sqrt{2}$$

where $e(R) \to 0$ as $R \to \infty$. From an earlier example this becomes

$$e(R) + 2 \lim_{r \to 0+} \int_r^R \frac{\ln t}{1+t^4} dt + i\pi \left(\frac{\sqrt{2}}{4} \pi \right) = \left(-\frac{1}{8} + \frac{1}{4} i \right) \pi^2 \sqrt{2}.$$

Now letting $r \to 0+$ and $R \to \infty$, we see

$$\begin{aligned} 2 \int_0^\infty \frac{\ln t}{1+t^4} dt &= \left(-\frac{1}{8} + \frac{1}{4} i \right) \pi^2 \sqrt{2} - i\pi \left(\frac{\sqrt{2}}{4} \pi \right) \\ &= -\frac{1}{8} \sqrt{2} \pi^2, \end{aligned}$$

and so

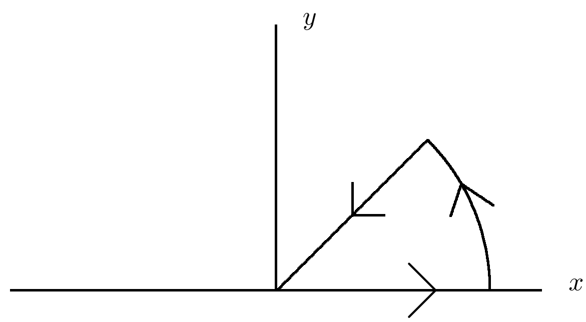
$$\int_0^\infty \frac{\ln t}{1+t^4} dt = -\frac{1}{16} \sqrt{2} \pi^2,$$

which is probably not the first thing you would thing of. You might try to imagine how this could be obtained using elementary techniques.

Example 28.6 *The Fresnel integrals are*

$$\int_0^\infty \cos x^2 dx, \ \int_0^\infty \sin x^2 dx.$$

To evaluate these integrals we will consider $f(z) = e^{iz^2}$ on the curve which goes from the origin to the point r on the x axis and from this point to the point $r\left(\frac{1+i}{\sqrt{2}}\right)$ along a circle of radius r, and from there back to the origin as illustrated in the following picture.



Thus the curve we integrate over is shaped like a slice of pie. Denote by γ_r the curved part. Since f is analytic,

$$
\begin{aligned}
0 &= \int_{\gamma_r} e^{iz^2} dz + \int_0^r e^{ix^2} dx - \int_0^r e^{i\left(t\left(\frac{1+i}{\sqrt{2}}\right)\right)^2} \left(\frac{1+i}{\sqrt{2}}\right) dt \\
&= \int_{\gamma_r} e^{iz^2} dz + \int_0^r e^{ix^2} dx - \int_0^r e^{-t^2} \left(\frac{1+i}{\sqrt{2}}\right) dt \\
&= \int_{\gamma_r} e^{iz^2} dz + \int_0^r e^{ix^2} dx - \frac{\sqrt{\pi}}{2} \left(\frac{1+i}{\sqrt{2}}\right) + e(r)
\end{aligned}
$$

where $e(r) \to 0$ as $r \to \infty$. Here we used the fact that $\int_0^\infty e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$. Now we need to examine the first of these integrals.

$$
\begin{aligned}
\left| \int_{\gamma_r} e^{iz^2} dz \right| &= \left| \int_0^{\frac{\pi}{4}} e^{i\left(re^{it}\right)^2} rie^{it} dt \right| \\
&\leq r \int_0^{\frac{\pi}{4}} e^{-r^2 \sin 2t} dt \\
&= \frac{r}{2} \int_0^1 \frac{e^{-r^2 u}}{\sqrt{1-u^2}} du
\end{aligned}
$$

$$
\leq \frac{r}{2} \int_0^{r^{-(3/2)}} \frac{1}{\sqrt{1-u^2}} du + \frac{r}{2} \left(\int_0^1 \frac{1}{\sqrt{1-u^2}} \right) e^{-\left(r^{1/2}\right)}
$$

which converges to zero as $r \to \infty$. Therefore, taking the limit as $r \to \infty$,

$$
\frac{\sqrt{\pi}}{2} \left(\frac{1+i}{\sqrt{2}}\right) = \int_0^\infty e^{ix^2} dx
$$

and so we can now find the Fresnel integrals

$$
\int_0^\infty \sin x^2 dx = \frac{\sqrt{\pi}}{2\sqrt{2}} = \int_0^\infty \cos x^2 dx.
$$

The next example illustrates the technique of integrating around a branch point.

Example 28.7 $\int_0^\infty \frac{x^{p-1}}{1+x} dx,\ p \in (0,1)$.

Since the exponent of x in the numerator is larger than -1. The integral does converge. However, the techniques of real analysis don't tell us what it converges to. The contour we will use is as follows: From $(\varepsilon, 0)$ to $(r, 0)$ along the x axis and then from $(r, 0)$ to $(r, 0)$ counter clockwise along the circle of radius r, then from $(r, 0)$ to $(\varepsilon, 0)$ along the x axis and from $(\varepsilon, 0)$ to $(\varepsilon, 0)$, clockwise along the circle of radius ε. You should draw a picture of this contour. The interesting thing about this is that we cannot define z^{p-1} all the way around 0. Therefore, we use a branch of z^{p-1} corresponding to the branch of the logarithm obtained by deleting the positive x axis. Thus

$$z^{p-1} = e^{(\ln|z|+iA(z))(p-1)}$$

where $z = |z| e^{iA(z)}$ and $A(z) \in (0, 2\pi)$. Along the integral which goes in the positive direction on the x axis, we will let $A(z) = 0$ while on the one which goes in the negative direction, we take $A(z) = 2\pi$. This is the appropriate choice obtained by replacing the line from $(\varepsilon, 0)$ to $(r, 0)$ with two lines having a small gap and then taking a limit as the gap closes. We leave it as an exercise to verify that the two integrals taken along the circles of radius ε and r converge to 0 as $\varepsilon \to 0$ and as $r \to \infty$. Therefore, taking the limit,

$$\int_0^\infty \frac{x^{p-1}}{1+x} dx + \int_\infty^0 \frac{x^{p-1}}{1+x} \left(e^{2\pi i(p-1)}\right) dx = 2\pi i \operatorname{Res}(f, -1) .$$

Calculating the residue of the integrand at -1, and simplifying the above expression, we obtain

$$\left(1 - e^{2\pi i(p-1)}\right) \int_0^\infty \frac{x^{p-1}}{1+x} dx = 2\pi i e^{(p-1)i\pi}.$$

Upon simplification we see that

$$\int_0^\infty \frac{x^{p-1}}{1+x} dx = \frac{\pi}{\sin p\pi}.$$

The following example is one of the most interesting. By an auspicious choice of the contour it is possible to obtain a very interesting formula for $\cot \pi z$ known as the Mittag Leffler expansion of $\cot \pi z$.

Example 28.8 *We let γ_N be the contour which goes from $-N - \frac{1}{2} - Ni$ horizontally to $N + \frac{1}{2} - Ni$ and from there, vertically to $N + \frac{1}{2} + Ni$ and then horizontally to $-N - \frac{1}{2} + Ni$ and finally vertically to $-N - \frac{1}{2} - Ni$. Thus the contour is a large rectangle and the direction of integration is in the counter clockwise direction. We will look at the following integral.*

$$I_N \equiv \int_{\gamma_N} \frac{\pi \cos \pi z}{\sin \pi z \left(\alpha^2 - z^2\right)} dz$$

where $\alpha \in \mathbb{R}$ is not an integer. This will be used to verify the formula of Mittag Leffler,

$$\frac{1}{\alpha^2} + \sum_{n=1}^{\infty} \frac{2}{\alpha^2 - n^2} = \frac{\pi \cot \pi\alpha}{\alpha}. \tag{28.1}$$

We leave it as an exercise to verify that $\cot \pi z$ is bounded on this contour and that therefore, $I_N \to 0$ as $N \to \infty$. Now we compute the residues of the integrand at $\pm\alpha$ and at n where $|n| < N + \frac{1}{2}$ for n an integer. These are the only singularities of the integrand in this contour and therefore, we can evaluate I_N by using these. We leave it as an exercise to calculate these residues and find that the residue at $\pm\alpha$ is

$$\frac{-\pi \cos \pi\alpha}{2\alpha \sin \pi\alpha}$$

while the residue at n is

$$\frac{1}{\alpha^2 - n^2}.$$

Therefore,

$$0 = \lim_{N\to\infty} I_N = \lim_{N\to\infty} 2\pi i \left[\sum_{n=-N}^{N} \frac{1}{\alpha^2 - n^2} - \frac{\pi \cot \pi\alpha}{\alpha} \right]$$

which establishes the following formula of Mittag Leffler.

$$\lim_{N\to\infty} \sum_{n=-N}^{N} \frac{1}{\alpha^2 - n^2} = \frac{\pi \cot \pi\alpha}{\alpha}.$$

Writing this in a slightly nicer form, we obtain (28.1).

28.1 The argument principle and Rouche's theorem

This technique of evaluating integrals by computing the residues also leads to the proof of a theorem referred to as the argument principle.

Definition 28.9 *We say a function defined on U, an open set, is meromorphic if its only singularities are poles, isolated singularities, a, for which*

$$\lim_{z\to a} |f(z)| = \infty.$$

Theorem 28.10 *(argument principle) Let f be meromorphic in U and let its poles be $\{p_1, \cdots, p_m\}$ and its zeros be $\{z_1, \cdots, z_n\}$. Let z_k be a zero of order r_k and let p_k be a pole of order l_k. Let $\gamma : [a,b] \to U$ be a continuous simple closed curve having bounded variation for which the inside of $\gamma([a,b])$ contains all the poles and zeros of f and is contained in U. Also let $n(\gamma, z) = 1$ for all z contained in the inside of $\gamma([a,b])$. Then*

$$\frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} dz = \sum_{k=1}^{n} r_k - \sum_{k=1}^{m} l_k$$

Proof: This theorem follows from computing the residues of f'/f. It has residues at poles and zeros. See Problem 4.

With the argument principle, we can prove Rouche's theorem . In the argument principle, we will denote by Z_f the quantity $\sum_{k=1}^{m} r_k$ and by P_f the quantity $\sum_{k=1}^{n} l_k$. Thus Z_f is the number of zeros of f counted according to the order of the zero with a similar definition holding for P_f.

$$\frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} dz = Z_f - P_f$$

Theorem 28.11 *(Rouche's theorem) Let f, g be meromorphic in U and let Z_f and P_f denote respectively the numbers of zeros and poles of f counted according to order. Let Z_g and P_g be defined similarly. Let $\gamma : [a,b] \to U$ be a simple closed continuous curve having bounded variation such that all poles and zeros of both f and g are inside $\gamma([a,b])$. Also let $n(\gamma, z) = 1$ for every z inside $\gamma([a,b])$. Also suppose that for $z \in \gamma([a,b])$*

$$|f(z) + g(z)| < |f(z)| + |g(z)|.$$

Then

$$Z_f - P_f = Z_g - P_g.$$

Proof: We see from the hypotheses that

$$\left|1+\frac{f(z)}{g(z)}\right|<1+\left|\frac{f(z)}{g(z)}\right|$$

which shows that for all $z \in \gamma([a,b])$,

$$\frac{f(z)}{g(z)} \in \mathbb{C} \setminus [0,\infty).$$

Letting l denote a branch of the logarithm defined on $\mathbb{C} \setminus [0,\infty)$, it follows that $l\left(\frac{f(z)}{g(z)}\right)$ is a primitive for the function, $\frac{(f/g)'}{(f/g)}$. Therefore, by the argument principle,

$$\begin{aligned} 0 &= \frac{1}{2\pi i}\int_\gamma \frac{(f/g)'}{(f/g)}dz = \frac{1}{2\pi i}\int_\gamma \left(\frac{f'}{f}-\frac{g'}{g}\right)dz \\ &= Z_f - P_f - (Z_g - P_g). \end{aligned}$$

This proves the theorem.

28.2 Exercises

1. In Example 28.2 we found the integral of a rational function of a certain sort. The technique used in this example typically works for rational functions of the form $\frac{f(x)}{g(x)}$ where $\deg(g(x)) \geq \deg f(x) + 2$ provided the rational function has no poles on the real axis. State and prove a theorem based on these observations.

2. Fill in the missing details of Example 28.8 about $I_N \to 0$. Note how important it was that the contour was chosen just right for this to happen. Also verify the claims about the residues.

3. Suppose f has a pole of order m at $z = a$. Define $g(z)$ by

$$g(z) = (z-a)^m f(z).$$

Show

$$\operatorname{Res}(f,a) = \frac{1}{(m-1)!}g^{(m-1)}(a).$$

Hint: Use the Laurent series.

4. Give a proof of Theorem 28.10. **Hint:** Let p be a pole. Show that near p, a pole of order m,

$$\frac{f'(z)}{f(z)} = \frac{-m+\sum_{k=1}^\infty b_k (z-p)^k}{(z-p)+\sum_{k=2}^\infty c_k (z-p)^k}$$

Show that $\operatorname{Res}(f,p) = -m$. Carry out a similar procedure for the zeros.

5. Use Rouche's theorem to prove the fundamental theorem of algebra which says that if $p(z) = z^n + a_{n-1}z^{n-1}\cdots + a_1 z + a_0$, then p has n zeros in $\mathbb{C}$. **Hint:** Let $q(z) = -z^n$ and let γ be a large circle, $\gamma(t) = re^{it}$ for r sufficiently large.

6. Consider the two polynomials $z^5 + 3z^2 - 1$ and $z^5 + 3z^2$. Show that on $|z| = 1$, we have the conditions for Rouche's theorem holding. Now use Rouche's theorem to verify that $z^5 + 3z^2 - 1$ must have two zeros in $|z| < 1$.

7. Consider the polynomial, $z^{11} + 7z^5 + 3z^2 - 17$. Use Rouche's theorem to find a bound on the zeros of this polynomial. In other words, find r such that if z is a zero of the polynomial, $|z| < r$. Try to make r fairly small if possible.

8. Verify that $\int_0^\infty e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$. **Hint:** Use polar coordinates.

9. Use the contour described in Example 28.2 to compute the exact values of the following improper integrals.

   (a) $\int_{-\infty}^\infty \frac{x}{(x^2+4x+13)^2} dx$

   (b) $\int_0^\infty \frac{x^2}{(x^2+a^2)^2} dx$

   (c) $\int_{-\infty}^\infty \frac{dx}{(x^2+a^2)(x^2+b^2)}, a, b > 0$

10. Evaluate the following improper integrals.

    (a) $\int_0^\infty \frac{\cos ax}{(x^2+b^2)^2} dx$

    (b) $\int_0^\infty \frac{x \sin x}{(x^2+a^2)^2} dx$

11. Find the Cauchy principle value of the integral

$$\int_{-\infty}^\infty \frac{\sin x}{(x^2+1)(x-1)} dx$$

defined as

$$\lim_{\varepsilon \to 0+} \left(\int_{-\infty}^{1-\varepsilon} \frac{\sin x}{(x^2+1)(x-1)} dx + \int_{1+\varepsilon}^\infty \frac{\sin x}{(x^2+1)(x-1)} dx \right).$$

12. Find a formula for the integral $\int_{-\infty}^\infty \frac{dx}{(1+x^2)^{n+1}}$ where n is a nonnegative integer.

13. Using the contour of Example 28.3 find $\int_{-\infty}^\infty \frac{\sin^2 x}{x^2} dx$.

14. If $m < n$ for m and n integers, show

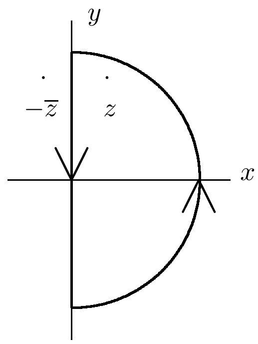
$$\int_0^\infty \frac{x^{2m}}{1+x^{2n}} dx = \frac{\pi}{n} \frac{1}{\sin\left(\frac{2m+1}{2n}\pi\right)}.$$

15. Find $\int_{-\infty}^\infty \frac{1}{(1+x^4)^2} dx$.

16. Find $\int_0^\infty \frac{\ln(x)}{1+x^2} dx = 0$

28.3 The Poisson formulas and the Hilbert transform

In this section we consider various applications of the above ideas by focussing on the contour, γ_R shown below, which represents a semicircle of radius R in the right half plane the direction of integration indicated by the arrows.



We will suppose that f is analytic in a region containing the right half plane and use the Cauchy integral formula to write

$$f(z) = \frac{1}{2\pi i}\int_{\gamma_R} \frac{f(w)}{w-z}dw,\ 0 = \frac{1}{2\pi i}\int_{\gamma_R} \frac{f(w)}{w+\overline{z}}dw,$$

the second integral equaling zero because the integrand is analytic as indicated in the picture. Therefore, multiplying the second integral by α and subtracting from the first we obtain

$$f(z) = \frac{1}{2\pi i}\int_{\gamma_R} f(w)\left(\frac{w+\overline{z}-\alpha w+\alpha z}{(w-z)(w+\overline{z})}\right)dw. \tag{28.2}$$

We would like to have the integrals over the semicircular part of the contour converge to zero as $R \to \infty$. This requires some sort of growth condition on f. Let

$$M(R) = \max\left\{\left|f\left(\mathrm{Re}^{it}\right)\right| : t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right\}.$$

We leave it as an exercise to verify that when

$$\lim_{R\to\infty} \frac{M(R)}{R} = 0 \text{ for } \alpha = 1 \tag{28.3}$$

and

$$\lim_{R\to\infty} M(R) = 0 \text{ for } \alpha \neq 1, \tag{28.4}$$

then this condition that the integrals over the curved part of γ_R converge to zero is satisfied. We assume this takes place in what follows. Taking the limit as $R \to \infty$

$$f(z) = \frac{-1}{2\pi}\int_{-\infty}^{\infty} f(i\xi)\left(\frac{i\xi+\overline{z}-\alpha i\xi+\alpha z}{(i\xi-z)(i\xi+\overline{z})}\right)d\xi \tag{28.5}$$

the negative sign occurring because the direction of integration along the y axis is negative. If $\alpha = 1$ and $z = x+iy$, this reduces to

$$f(z) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(i\xi)\left(\frac{x}{|z-i\xi|^2}\right)d\xi, \tag{28.6}$$

which is called the Poisson formula for a half plane.. If we assume $M(R) \to 0$, and take $\alpha = -1$, (28.5) reduces to

$$\frac{i}{\pi}\int_{-\infty}^{\infty} f(i\xi)\left(\frac{\xi-y}{|z-i\xi|^2}\right)d\xi. \tag{28.7}$$

Of course we can consider real and imaginary parts of f in these formulas. Let

$$f(i\xi) = u(\xi) + iv(\xi).$$

From (28.6) we obtain upon taking the real part,

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} u(\xi)\left(\frac{x}{|z-i\xi|^2}\right)d\xi. \tag{28.8}$$

Taking real and imaginary parts in (28.7) gives the following.

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} v(\xi)\left(\frac{y-\xi}{|z-i\xi|^2}\right)d\xi, \tag{28.9}$$

$$v(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} u(\xi)\left(\frac{\xi-y}{|z-i\xi|^2}\right)d\xi. \tag{28.10}$$

These are called the conjugate Poisson formulas because knowledge of the imaginary part on the y axis leads to knowledge of the real part for $\operatorname{Re} z > 0$ while knowledge of the real part on the imaginary axis leads to knowledge of the real part on $\operatorname{Re} z > 0$.

We obtain the Hilbert transform by formally letting $z = iy$ in the conjugate Poisson formulas and picking $x = 0$. Letting $u(0,y) = u(y)$ and $v(0,y) = v(y)$, we obtain, at least formally

$$\begin{aligned} u(y) &= \frac{1}{\pi}\int_{-\infty}^{\infty} v(\xi)\left(\frac{1}{y-\xi}\right)d\xi, \\ v(y) &= -\frac{1}{\pi}\int_{-\infty}^{\infty} u(\xi)\left(\frac{1}{y-\xi}\right)d\xi. \end{aligned}$$

Of course there are major problems in writing these integrals due to the integrand possessing a nonintegrable singularity at y. There is a large theory connected with the meaning of such integrals as these known as the theory of singular integrals. Here we evaluate these integrals by taking a contour which goes around the singularity and then taking a limit to obtain a principle value integral.

The case when $\alpha = 0$ in (28.5) yields

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(i\xi)}{(z-i\xi)}d\xi. \tag{28.11}$$

We will use this formula in considering the problem of finding the inverse Laplace transform.

We say a function, f, defined on $(0,\infty)$ is of exponential type if

$$|f(t)| < Ae^{at} \tag{28.12}$$

for some constants A and a. For such a function we can define the Laplace transform as follows.

$$F(s) \equiv \int_0^{\infty} f(t)\, e^{-st}dt \equiv Lf. \tag{28.13}$$

We leave it as an exercise to show that this integral makes sense for all $\operatorname{Re} s > a$ and that the function so defined is analytic on $\operatorname{Re} z > a$. Using the estimate, (28.12), we obtain that for $\operatorname{Re} s > a$,

$$|F(s)| \leq \left|\frac{A}{s-a}\right|. \tag{28.14}$$

We will show that if $f(t)$ is given by the formula,

$$e^{-(a+\varepsilon)t} f(t) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi t} F(i\xi + a + \varepsilon)\, d\xi,$$

then $Lf = F$ for all s large enough.

$$L\left(e^{-(a+\varepsilon)t} f(t)\right) = \frac{1}{2\pi} \int_0^{\infty} e^{-st} \int_{-\infty}^{\infty} e^{i\xi t} F(i\xi + a + \varepsilon)\, d\xi dt$$

Now if

$$\int_{-\infty}^{\infty} |F(i\xi + a + \varepsilon)|\, d\xi < \infty, \tag{28.15}$$

we can use Fubini's theorem to interchange the order of integration. Unfortunately, we do not know this. The best we have is the estimate (28.14). However, this is a very crude estimate and often (28.15) will hold. Therefore, we shall assume whatever we need in order to continue with the symbol pushing and interchange the order of integration to obtain with the aid of (28.11) the following:

$$\begin{aligned} L\left(e^{-(a+\varepsilon)t} f(t)\right) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_0^{\infty} e^{-(s-i\xi)t} dt\right) F(i\xi + a + \varepsilon)\, d \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{F(i\xi + a + \varepsilon)}{s - i\xi} d\xi \\ &= F(s + a + \varepsilon) \end{aligned}$$

for all $s > 0$. (The reason for fussing with $\xi + a + \varepsilon$ rather than just ξ is so the function, $\xi \to F(\xi + a + \varepsilon)$ will be analytic on $\operatorname{Re} \xi > -\varepsilon$, a region containing the right half plane allowing us to use (28.11).) Now with this information, we may verify that $L(f)(s) = F(s)$ for all $s > a$. We just showed

$$\int_0^{\infty} e^{-wt} e^{-(a+\varepsilon)t} f(t)\, dt = F(w + a + \varepsilon)$$

whenever $\operatorname{Re} w > 0$. Let $s = w + a + \varepsilon$. Then $L(f)(s) = F(s)$ whenever $\operatorname{Re} s > a + \varepsilon$. Since ε is arbitrary, this verifies $L(f)(s) = F(s)$ for all $s > a$. It follows that if we are given $F(s)$ which is analytic for $\operatorname{Re} s > a$ and we want to find f such that $L(f) = F$, we should pick $c > a$ and define

$$e^{-ct} f(t) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi t} F(i\xi + c)\, d\xi.$$

Changing the variable, to let $s = i\xi + c$, we may write this as

$$f(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{st} F(s)\, ds, \tag{28.16}$$

and we know from the above argument that we can expect this procedure to work if things are not too pathological. This integral is called the Bromwich integral for the inversion of the Laplace transform. The function $f(t)$ is the inverse Laplace transform.

We illustrate this procedure with a simple example. Suppose $F(s) = \frac{s}{(s^2+1)^2}$. In this case, F is analytic for $\operatorname{Re} s > 0$. Let $c = 1$ and integrate over a contour which goes from $c - iR$ vertically to $c + iR$ and then follows a semicircle in the counter clockwise direction back to $c - iR$. Clearly the integrals over the curved portion of the contour converge to 0 as $R \to \infty$. There are two residues of this function, one at i and one at $-i$. At both of these points the poles are of order two and so we find the residue at i by

$$\begin{aligned} \operatorname{Res}(f, i) &= \lim_{s \to i} \frac{d}{ds} \left(\frac{e^{ts} s (s - i)^2}{(s^2 + 1)^2}\right) \\ &= \frac{-ite^{it}}{4} \end{aligned}$$

and the residue at $-i$ is

$$\begin{aligned}\operatorname{Res}(f,-i) &= \lim_{s\to -i}\frac{d}{ds}\left(\frac{e^{ts}s\,(s+i)^2}{(s^2+1)^2}\right)\\ &= \frac{ite^{-it}}{4}\end{aligned}$$

Now evaluating the contour integral and taking $R\to\infty$, we find that the integral in (28.16) equals

$$2\pi i\left(\frac{ite^{-it}}{4}+\frac{-ite^{it}}{4}\right)=i\pi t\sin t$$

and therefore,

$$f(t)=\frac{1}{2}t\sin t.$$

You should verify that this actually works giving $L(f)=\frac{s}{(s^2+1)^2}$.

28.4 Exercises

1. Verify that the integrals over the curved part of γ_R in (28.2) converge to zero when (28.3) and (28.4) are satisfied.

2. Obtain similar formulas to (28.8) for the imaginary part in the case where $\alpha=1$ and formulas (28.9) - (28.10) in the case where $\alpha=-1$. Observe that these formulas give an explicit formula for $f(z)$ if either the real or the imaginary parts of f are known along the line $x=0$.

3. Verify that the formula for the Laplace transform, (28.13) makes sense for all $s>a$ and that F is analytic for $\operatorname{Re} z>a$.

4. Find inverse Laplace transforms for the functions,
$\frac{a}{s^2+a^2}, \frac{a}{s^2(s^2+a^2)}, \frac{1}{s^7}, \frac{s}{(s^2+a^2)^2}$.

5. Consider the analytic function e^{-z}. Show it satisfies the necessary conditions in order to apply formula (28.6). Use this to verify the formulas,
$$\begin{aligned}e^{-x}\cos y &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x\cos\xi}{x^2+(y-\xi)^2}d\xi,\\ e^{-x}\sin y &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x\sin\xi}{x^2+(y-\xi)^2}d\xi.\end{aligned}$$

6. The Poisson formula gives
$$u(x,y)=\frac{1}{\pi}\int_{-\infty}^{\infty}u(0,\xi)\left(\frac{x}{x^2+(y-\xi)^2}\right)d\xi$$
whenever u is the real part of a function analytic in the right half plane which has a suitable growth condition. Show that this implies
$$1=\frac{1}{\pi}\int_{-\infty}^{\infty}\left(\frac{x}{x^2+(y-\xi)^2}\right)d\xi.$$

7. Now consider an arbitrary continuous function, $u(\xi)$ and define

$$u(x,y) \equiv \frac{1}{\pi}\int_{-\infty}^{\infty} u(\xi)\left(\frac{x}{x^2+(y-\xi)^2}\right)d\xi.$$

Verify that for $u(x,y)$ given by this formula,

$$\lim_{x\to 0+} |u(x,y)-u(y)| = 0,$$

and that u is a harmonic function, $u_{xx}+u_{yy}=0$, on $x>0$. Therefore, this integral yields a solution to the Dirichlet problem on the half plane which is to find a harmonic function which assumes given boundary values.

8. To what extent can we relax the assumption that $\xi \to u(\xi)$ is continuous?

28.5 Infinite products

In this section we give an introduction to the topic of infinite products and apply the theory to the Gamma function. To begin with we give a definition of what is meant by an infinite product.

Definition 28.12 $\prod_{n=1}^{\infty}(1+u_n) \equiv \lim_{n\to\infty}\prod_{k=1}^{n}(1+u_k)$ *whenever this limit exists. If* $u_n = u_n(z)$ *for* $z \in H$, *we say the infinite product converges uniformly on* H *if the partial products,* $\prod_{k=1}^{n}(1+u_k(z))$ *converge uniformly on* H.

Lemma 28.13 *Let* $P_N \equiv \prod_{k=1}^{N}(1+u_k)$ *and let* $Q_N \equiv \prod_{k=1}^{N}(1+|u_k|)$. *Then*

$$Q_N \leq \exp\left(\sum_{k=1}^{N}|u_k|\right),\ |P_N - 1| \leq Q_N - 1$$

Proof: To verify the first inequality,

$$Q_N = \prod_{k=1}^{N}(1+|u_k|) \leq \prod_{k=1}^{N} e^{|u_k|} = \exp\left(\sum_{k=1}^{N}|u_k|\right).$$

The second claim is obvious if $N=1$. Consider $N=2$.

$$\begin{aligned}|(1+u_1)(1+u_2)-1| &= |u_2+u_1+u_1u_2| \\ &\leq 1+|u_1|+|u_2|+|u_1||u_2|-1 \\ &= (1+|u_1|)(1+|u_2|)-1\end{aligned}$$

Continuing this way the desired inequality follows.

The main theorem is the following.

Theorem 28.14 *Let* $H \subseteq \mathbb{C}$ *and suppose that* $\sum_{n=1}^{\infty}|u_n(z)|$ *converges uniformly on* H. *Then*

$$P(z) \equiv \prod_{n=1}^{\infty}(1+u_n(z))$$

converges uniformly on H. *If* $(n_1, n_2, \cdots)$ *is any permutation of* $(1, 2, \cdots)$, *then for all* $z \in H$,

$$P(z) = \prod_{k=1}^{\infty}(1+u_{n_k}(z))$$

and P *has a zero at* z_0 *if and only if* $u_n(z_0) = -1$ *for some* n.

Proof: We use Lemma 28.13 to write for $m < n$, and all $z \in H$,

$$
\begin{aligned}
&\left|\prod_{k=1}^{n}(1+u_k(z)) - \prod_{k=1}^{m}(1+u_k(z))\right| \\
\leq\ & \left|\prod_{k=1}^{m}(1+|u_k(z)|)\right| \left|\prod_{k=m+1}^{n}(1+u_k(z)) - 1\right| \\
\leq\ & \exp\left(\sum_{k=1}^{\infty}|u_k(z)|\right) \left|\prod_{k=m+1}^{n}(1+|u_k(z)|) - 1\right| \\
\leq\ & C\left(\exp\left(\sum_{k=m+1}^{\infty}|u_k(z)|\right) - 1\right) \\
\leq\ & C\left(e^{\varepsilon} - 1\right)
\end{aligned}
$$

whenever m is large enough. This shows the partial products form a uniformly Cauchy sequence and hence converge uniformly on H. This verifies the first part of the theorem.

Next we need to verify the part about taking the product in different orders. Suppose then that $(n_1, n_2, \cdots)$ is a permutation of the list, $(1, 2, \cdots)$ and choose M large enough that for all $z \in H$,

$$
\left|\prod_{k=1}^{\infty}(1+u_k(z)) - \prod_{k=1}^{M}(1+u_k(z))\right| < \varepsilon.
$$

Then for all N sufficiently large, $\{n_1, n_2, \cdots, n_N\} \supseteq \{1, 2, \cdots, M\}$. Then for N this large, we use Lemma 28.13 to obtain

$$
\begin{aligned}
&\left|\prod_{k=1}^{M}(1+u_k(z)) - \prod_{k=1}^{N}(1+u_{n_k}(z))\right| \leq \\
&\left|\prod_{k=1}^{M}(1+u_k(z))\right| \left|1 - \prod_{k\leq N, n_k > M}(1+u_{n_k}(z))\right| \\
\leq\ & \left|\prod_{k=1}^{M}(1+u_k(z))\right| \left|\prod_{k\leq N, n_k > M}(1+|u_{n_k}(z)|) - 1\right| \\
\leq\ & \left|\prod_{k=1}^{M}(1+u_k(z))\right| \left|\prod_{l=M}^{\infty}(1+|u_l(z)|) - 1\right| \\
\leq\ & \left|\prod_{k=1}^{M}(1+u_k(z))\right| \left(\exp\left(\sum_{l=M}^{\infty}|u_l(z)|\right) - 1\right) \\
\leq\ & \left|\prod_{k=1}^{M}(1+u_k(z))\right| (\exp\varepsilon - 1) \qquad (28.17) \\
\leq\ & \left|\prod_{k=1}^{\infty}(1+|u_k(z)|)\right| (\exp\varepsilon - 1) \qquad (28.18)
\end{aligned}
$$

whenever M is large enough. Therefore, this shows, using (28.18) that

$$
\left|\prod_{k=1}^{N}(1+u_{n_k}(z)) - \prod_{k=1}^{\infty}(1+u_k(z))\right| \leq
$$

$$\left|\prod_{k=1}^{N}\left(1+u_{n_k}\left(z\right)\right)-\prod_{k=1}^{M}\left(1+u_k\left(z\right)\right)\right|+$$

$$\left|\prod_{k=1}^{M}\left(1+u_k\left(z\right)\right)-\prod_{k=1}^{\infty}\left(1+u_k\left(z\right)\right)\right|$$

$$\leq\varepsilon+\left(\left|\prod_{k=1}^{\infty}\left(1+\left|u_k\left(z\right)\right|\right)\right|+\varepsilon\right)\left(\exp\varepsilon-1\right)$$

which verifies the claim about convergence of the permuted products.

It remains to verify the assertion about the points, z_0, where $P(z_0)=0$. Obviously, if $u_n(z_0)=-1$, then $P(z_0)=0$. Suppose then that $P(z_0)=0$. Letting $n_k=k$ and using (28.17), we may take the limit as $N\to\infty$ to obtain

$$\left|\prod_{k=1}^{M}\left(1+u_k\left(z_0\right)\right)\right|=$$

$$\left|\prod_{k=1}^{M}\left(1+u_k\left(z_0\right)\right)-\prod_{k=1}^{\infty}\left(1+u_k\left(z_0\right)\right)\right|$$

$$\leq\quad\left|\prod_{k=1}^{M}\left(1+u_k\left(z_0\right)\right)\right|\left(\exp\varepsilon-1\right).$$

If ε is chosen small enough in this inequality, we see this implies $\prod_{k=1}^{M}\left(1+u_k\left(z\right)\right)=0$ and therefore, $u_k(z_0)=-1$ for some $k\leq M$. This proves the theorem.

Now we present the Weierstrass product formula. This formula tells how to factor analytic functions into an infinite product. It is a very interesting and useful theorem. First we need to give a definition of the elementary factors.

Definition 28.15 *Let $E_0(z)\equiv 1-z$ and for $p\geq 1$,*

$$E_p\left(z\right)\equiv\left(1-z\right)\exp\left(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p}\right)$$

The fundamental factors satisfy an important estimate which is stated next.

Lemma 28.16 *For all $|z|\leq 1$ and $p=0,1,2,\cdots$,*

$$\left|1-E_p\left(z\right)\right|\leq\left|z\right|^{p+1}.$$

Proof: If $p=0$ this is obvious. Suppose therefore, that $p\geq 1$.

$$E_p'\left(z\right)=-\exp\left(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p}\right)+$$

$$\left(1-z\right)\exp\left(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p}\right)\left(1+z+\cdots+z^{p-1}\right)$$

and so, since $(1-z)\left(1+z+\cdots+z^{p-1}\right)=1-z^p$,

$$E_p'(z)=-z^p\exp\left(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p}\right)$$

which shows that E_p' has a zero of order p at 0. Thus, from the equation just derived,

$$E_p'(z)=-z^p\sum_{k=0}^{\infty}a_kz^k$$

where each $a_k\geq 0$ and $a_0=1$. This last assertion about the sign of the a_k follows easily from differentiating the function $f(z)=\exp\left(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p}\right)$ and evaluating the derivatives at $z=0$. A primitive for $E_p'(z)$ is of the form $-\sum_{k=0}^{\infty}a_k\frac{z^{k+1+p}}{k+p+1}$ and so integrating from 0 to z along $\gamma(0,z)$ we see that

$$E_p(z)-E_p(0)=$$

$$\begin{aligned}E_p(z)-1 &= -\sum_{k=0}^{\infty}a_k\frac{z^{k+p+1}}{k+p+1}\\ &= -z^{p+1}\sum_{k=0}^{\infty}a_k\frac{z^k}{k+p+1}\end{aligned}$$

which shows that $(E_p(z)-1)/z^{p+1}$ has a removable singularity at $z=0$.

Now from the formula for $E_p(z)$,

$$E_p(z)-1=(1-z)\exp\left(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p}\right)-1$$

and so

$$E_p(1)-1=-1=-\sum_{k=0}^{\infty}a_k\frac{1}{k+p+1}$$

Since each $a_k\geq 0$, we see that for $|z|=1$,

$$\frac{|1-E_p(z)|}{|z^{p+1}|}\leq\sum_{k=1}^{\infty}a_k\frac{1}{k+p+1}=1.$$

Now by the maximum modulus theorem,

$$|1-E_p(z)|\leq|z|^{p+1}$$

for all $|z|\leq 1$. This proves the lemma.

Theorem 28.17 *Let z_n be a sequence of nonzero complex numbers which have no limit point in $\mathbb{C}$ and suppose there exist, p_n, nonnegative integers such that*

$$\sum_{n=1}^{\infty}\left(\frac{r}{|z_n|}\right)^{1+p_n}<\infty \tag{28.19}$$

for all $r \in \mathbb{R}$. Then

$$P(z) \equiv \prod_{n=1}^{\infty} E_{p_n}\left(\frac{z}{z_n}\right)$$

is analytic on $\mathbb{C}$ and has a zero at each point, z_n and at no others. If w occurs m times in $\{z_n\}$, then P has a zero of order m at w.

Proof: The series

$$\sum_{n=1}^{\infty} \left|\frac{z}{z_n}\right|^{1+p_n}$$

converges uniformly on any compact set because if $|z| \leq r$, then

$$\left|\left(\frac{z}{z_n}\right)^{1+p_n}\right| \leq \left(\frac{r}{|z_n|}\right)^{1+p_n}$$

and so we may apply the Weierstrass M test to obtain the uniform convergence of $\sum_{n=1}^{\infty}\left(\frac{z}{z_n}\right)^{1+p_n}$ on $|z| < r$. Also,

$$\left|E_{p_n}\left(\frac{z}{z_n}\right) - 1\right| \leq \left(\frac{|z|}{|z_n|}\right)^{p_n+1}$$

by Lemma 28.16 whenever n is large enough because the hypothesis that $\{z_n\}$ has no limit point requires that $\lim_{n\to\infty}|z_n| = \infty$. Therefore, by Theorem 28.14,

$$P(z) \equiv \prod_{n=1}^{\infty} E_{p_n}\left(\frac{z}{z_n}\right)$$

converges uniformly on compact subsets of $\mathbb{C}$. Letting $P_n(z)$ denote the nth partial product for $P(z)$, we have for $|z| < r$

$$P_n(z) = \frac{1}{2\pi i}\int_{\gamma_r} \frac{P_n(w)}{w - z}dw$$

where $\gamma_r(t) \equiv re^{it}$, $t \in [0, 2\pi]$. By the uniform convergence of P_n to P on compact sets, it follows the same formula holds for P in place of P_n showing that P is analytic in $B(0, r)$. Since r is arbitrary, we see that P is analytic on all of $\mathbb{C}$.

Now we ask where the zeros of P are. By Theorem 28.14, the zeros occur at exactly those points, z, where

$$E_{p_n}\left(\frac{z}{z_n}\right) - 1 = -1.$$

In that theorem $E_{p_n}\left(\frac{z}{z_n}\right) - 1$ plays the role of $u_n(z)$. Thus we need $E_{p_n}\left(\frac{z}{z_n}\right) = 0$ for some n. However, this occurs exactly when $\frac{z}{z_n} = 1$ so the zeros of P are the points $\{z_n\}$.

If w occurs m times in the sequence, $\{z_n\}$, we let $n_1, \cdots, n_m$ be those indices at which w occurs. Then we choose a permutation of $(1, 2, \cdots)$ which starts with the list $(n_1, \cdots, n_m)$. By Theorem 28.14,

$$P(z) = \prod_{k=1}^{\infty} E_{p_{n_k}}\left(\frac{z}{z_{n_k}}\right) = \left(1 - \frac{z}{w}\right)^m g(z)$$

where g is an analytic function which is not equal to zero at w. It follows from this that P has a zero of order m at w. This proves the theorem.

The next theorem is the Weierstrass factorization theorem which can be used to factor a given function, f, rather than only deciding convergence questions.

Theorem 28.18 *Let f be analytic on $\mathbb{C}$, $f(0) \neq 0$, and let the zeros of f be $\{z_k\}$, listed according to order. (Thus if z is a zero of order m, it will be listed m times in the list, $\{z_k\}$.) Then there exists an entire function, g and a sequence of nonnegative integers, p_n such that*

$$f(z) = e^{g(z)} \prod_{n=1}^{\infty} E_{p_n}\left(\frac{z}{z_n}\right). \tag{28.20}$$

Note that $e^{g(z)} \neq 0$ for any z and this is the interesting thing about this function.

Proof: We know $\{z_n\}$ cannot have a limit point because if there were a limit point of this sequence, it would follow from Theorem 26.1 that $f(z) = 0$ for all z, contradicting the hypothesis that $f(0) \neq 0$. Hence $\lim_{n\to\infty} |z_n| = \infty$ and so

$$\sum_{n=1}^{\infty} \left(\frac{r}{|z_n|}\right)^{1+n-1} = \sum_{n=1}^{\infty} \left(\frac{r}{|z_n|}\right)^{n} < \infty$$

by the root test. Therefore, by Theorem 28.17 we may write

$$P(z) = \prod_{n=1}^{\infty} E_{p_n}\left(\frac{z}{z_n}\right)$$

a function analytic on $\mathbb{C}$ by picking $p_n = n-1$ or perhaps some other choice. (We know $p_n = n-1$ works but we do not know this is the only choice that might work.) Then f/P has only removable singularities in $\mathbb{C}$ and no zeros thanks to Theorem 28.17. Thus, letting $h(z) = f(z)/P(z)$, we know from Corollary 25.12 that h'/h has a primitive, $\widetilde{g}$. Then

$$\left(he^{-\widetilde{g}}\right)' = 0$$

and so

$$h(z) = e^{a+ib} e^{\widetilde{g}(z)}$$

for some constants, a, b. Therefore, letting $g(z) = \widetilde{g}(z) + a + ib$, we see that $h(z) = e^{g(z)}$ and thus (28.20) holds. This proves the theorem.

Corollary 28.19 *Let f be analytic on $\mathbb{C}$, f has a zero of order m at 0, and let the other zeros of f be $\{z_k\}$, listed according to order. (Thus if z is a zero of order l, it will be listed l times in the list, $\{z_k\}$.) Then there exists an entire function, g and a sequence of nonnegative integers, p_n such that*

$$f(z) = z^m e^{g(z)} \prod_{n=1}^{\infty} E_{p_n}\left(\frac{z}{z_n}\right).$$

Proof: Since f has a zero of order m at 0, it follows from Theorem 26.1 that $\{z_k\}$ cannot have a limit point in $\mathbb{C}$ and so we may apply Theorem 28.18 to the function, $f(z)/z^m$ which has a removable singularity at 0. This proves the corollary.

28.6 Exercises

1. Show $\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right)=\frac{1}{2}$. **Hint:** Take the ln of the partial product and then exploit the telescoping series.

2. Suppose $P(z)=\prod_{k=1}^{\infty}f_k(z)\neq 0$ for all $z\in U$, an open set, that convergence is uniform on compact subsets of U, and f_k is analytic on U. Show
$$P'(z)=\sum_{k=1}^{\infty}f_k'(z)\prod_{n\neq k}f_n(z).$$
Hint: Use a branch of the logarithm, defined near $P(z)$ and logarithmic differentiation.

3. Show that $\frac{\sin \pi z}{\pi z}$ has a removable singularity at $z=0$ and so there exists an analytic function, q, defined on $\mathbb{C}$ such that $\frac{\sin \pi z}{\pi z}=q(z)$ and $q(0)=1$. Using the Weierstrass product formula, show that
$$\begin{aligned} q(z) &= e^{g(z)}\prod_{k\in\mathbb{Z},k\neq 0}\left(1-\frac{z}{k}\right)e^{\frac{z}{k}} \\ &= e^{g(z)}\prod_{k=1}^{\infty}\left(1-\frac{z^2}{k^2}\right)\end{aligned}$$
for some analytic function, $g(z)$ and that we may take $g(0)=0$.

4. $\uparrow$ Use Problem 2 along with Problem 3 to show that
$$\frac{\cos \pi z}{z}-\frac{\sin \pi z}{\pi z^2}=e^{g(z)}g'(z)\prod_{k=1}^{\infty}\left(1-\frac{z^2}{k^2}\right)-$$
$$2ze^{g(z)}\sum_{n=1}^{\infty}\frac{1}{n^2}\prod_{k\neq n}\left(1-\frac{z^2}{k^2}\right).$$
Now divide this by $q(z)$ on both sides to show
$$\pi\cot \pi z-\frac{1}{z}=g'(z)+2z\sum_{n=1}^{\infty}\frac{1}{z^2-n^2}.$$
Use the Mittag Leffler expansion for the $\cot \pi z$ to conclude from this that $g'(z)=0$ and hence, $g(z)=0$ so that
$$\frac{\sin \pi z}{\pi z}=\prod_{k=1}^{\infty}\left(1-\frac{z^2}{k^2}\right).$$

5. $\uparrow$ In the formula for the product expansion of $\frac{\sin \pi z}{\pi z}$, let $z=\frac{1}{2}$ to obtain a formula for $\frac{\pi}{2}$ called Wallis's formula. Is this formula you have come up with a good way to calculate π?

6. This and the next collection of problems are dealing with the gamma function. Show that
$$\left|\left(1+\frac{z}{n}\right)e^{\frac{-z}{n}}-1\right|\leq\frac{C(z)}{n^2}$$
and therefore,
$$\sum_{n=1}^{\infty}\left|\left(1+\frac{z}{n}\right)e^{\frac{-z}{n}}-1\right|<\infty$$
with the convergence uniform on compact sets.

7. $\uparrow$ Show $\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)e^{\frac{-z}{n}}$ converges to an analytic function on $\mathbb{C}$ which has zeros only at the negative integers and that therefore,
$$\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)^{-1}e^{\frac{z}{n}}$$
is a meromorphic function (Analytic except for poles) having simple poles at the negative integers.

8. $\uparrow$Show there exists γ such that if
$$\Gamma(z)\equiv\frac{e^{-\gamma z}}{z}\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)^{-1}e^{\frac{z}{n}},$$
then $\Gamma(1)=1$. **Hint:** $\prod_{n=1}^{\infty}(1+n)e^{-1/n}=c=e^{\gamma}$.

9. $\uparrow$Now show that
$$\gamma=\lim_{n\to\infty}\left[\sum_{k=1}^{n}\frac{1}{k}-\ln n\right]$$
Hint: Show $\gamma=\sum_{n=1}^{\infty}\left[\ln\left(1+\frac{1}{n}\right)-\frac{1}{n}\right]=\sum_{n=1}^{\infty}\left[\ln(1+n)-\ln n-\frac{1}{n}\right]$.

10. $\uparrow$Justify the following argument leading to Gauss's formula
$$\begin{aligned}\Gamma(z) &= \lim_{n\to\infty}\left(\prod_{k=1}^{n}\left(\frac{k}{k+z}\right)e^{\frac{z}{k}}\right)\frac{e^{-\gamma z}}{z}\\ &= \lim_{n\to\infty}\left(\frac{n!}{(1+z)(2+z)\cdots(n+z)}e^{z\left(\sum_{k=1}^{n}\frac{1}{k}\right)}\right)\frac{e^{-\gamma z}}{z}\\ &= \lim_{n\to\infty}\frac{n!}{(1+z)(2+z)\cdots(n+z)}e^{z\left(\sum_{k=1}^{n}\frac{1}{k}\right)}e^{-z\left[\sum_{k=1}^{n}\frac{1}{k}-\ln n\right]}\\ &= \lim_{n\to\infty}\frac{n!n^{z}}{(1+z)(2+z)\cdots(n+z)}.\end{aligned}$$

11. $\uparrow$ Verify from the Gauss formula above that $\Gamma(z+1)=\Gamma(z)z$ and that for n a nonnegative integer, $\Gamma(n+1)=n!$.

12. $\uparrow$ The usual definition of the gamma function for positive x is
$$\Gamma_1(x)\equiv\int_0^{\infty}e^{-t}t^{x-1}dt.$$
Show $\left(1-\frac{t}{n}\right)^n\leq e^{-t}$ for $t\in[0,n]$. Then show
$$\int_0^n\left(1-\frac{t}{n}\right)^n t^{x-1}dt=\frac{n!n^x}{x(x+1)\cdots(x+n)}.$$
Use the first part and the dominated convergence theorem or heuristics if you have not studied this theorem to conclude that
$$\Gamma_1(x)=\lim_{n\to\infty}\frac{n!n^x}{x(x+1)\cdots(x+n)}=\Gamma(x).$$
Hint: To show $\left(1-\frac{t}{n}\right)^n\leq e^{-t}$ for $t\in[0,n]$, verify this is equivalent to showing $(1-u)^n\leq e^{-nu}$ for $u\in[0,1]$.

13. $\uparrow$Show $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$. whenever $\operatorname{Re} z > 0$. **Hint:** You have already shown that this is true for positive real numbers. Verify this formula for $\operatorname{Re} z$ yields an analytic function.

14. $\uparrow$Show $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Then find $\Gamma\left(\frac{5}{2}\right)$.

The Riemann mapping theorem

We know from the open mapping theorem that analytic functions map regions to other regions or else to single points. In this chapter we prove the remarkable Riemann mapping theorem which states that for every simply connected region, U there exists an analytic function, f such that $f(U) = B(0,1)$ and in addition to this, f is one to one. The proof involves several ideas which have been developed up to now. We also need the following important theorem, a case of Montel's theorem.

Theorem 29.1 *Let U be an open set in $\mathbb{C}$ and let $\mathcal{F}$ denote a set of analytic functions mapping U to $B(0,M)$. Then there exists a sequence of functions from $\mathcal{F}$, $\{f_n\}_{n=1}^{\infty}$ and an analytic function, f such that $f_n^{(k)}$ converges uniformly to $f^{(k)}$ on every compact subset of U.*

Proof: First we note there exists a sequence of compact sets, K_n such that $K_n \subseteq \operatorname{int} K_{n+1} \subseteq U$ for all n where here $\operatorname{int} K$ denotes the interior of the set K, the union of all open sets contained in K and $\cup_{n=1}^{\infty} K_n = U$. We leave it as an exercise to verify that $\overline{B(0,n)} \cap \left\{z \in U : \operatorname{dist}\left(z, U^C\right) \leq \frac{1}{n}\right\}$ works for K_n. Then there exist positive numbers, δ_n such that if $z \in K_n$, then $\overline{B(z,\delta_n)} \subseteq \operatorname{int} K_{n+1}$. Now denote by $\mathcal{F}_n$ the set of restrictions of functions of $\mathcal{F}$ to K_n. Then let $z \in K_n$ and let $\gamma(t) \equiv z + \delta_n e^{it}, t \in [0, 2\pi]$. It follows that for $z_1 \in B(z,\delta_n)$, and $f \in \mathcal{F}$,

$$\begin{aligned} |f(z) - f(z_1)| &= \left|\frac{1}{2\pi i}\int_\gamma f(w)\left(\frac{1}{w-z} - \frac{1}{w-z_1}\right)dw\right| \\ &\leq \frac{1}{2\pi}\left|\int_\gamma f(w)\frac{z-z_1}{(w-z)(w-z_1)}dw\right| \end{aligned}$$

Letting $|z_1 - z| < \frac{\delta_n}{2}$, we can estimate this and write

$$\begin{aligned} |f(z) - f(z_1)| &\leq \frac{M}{2\pi}2\pi\delta_n\frac{|z-z_1|}{\delta_n^2/2} \\ &\leq 2M\frac{|z-z_1|}{\delta_n}. \end{aligned}$$

It follows that $\mathcal{F}_n$ is equicontinuous and uniformly bounded so by the Arzela Ascoli theorem there exists a sequence, $\{f_{nk}\}_{k=1}^{\infty} \subseteq \mathcal{F}$ which converges uniformly on K_n. Let $\{f_{1k}\}_{k=1}^{\infty}$ converge uniformly on K_1. Then use the Arzela Ascoli theorem applied to this sequence to get a subsequence, denoted by $\{f_{2k}\}_{k=1}^{\infty}$ which also converges uniformly on K_2. Continue in this way to obtain $\{f_{nk}\}_{k=1}^{\infty}$ which converges uniformly on $K_1, \cdots, K_n$. Now the sequence $\{f_{nn}\}_{n=m}^{\infty}$ is a subsequence of $\{f_{mk}\}_{k=1}^{\infty}$ and so it converges uniformly on K_m for all m. Denoting f_{nn} by f_n for short, this is the sequence of functions promised by the theorem. It is clear $\{f_n\}_{n=1}^{\infty}$ converges uniformly on every compact subset of U because every such set is contained in K_m for all m large enough. Let $f(z)$ be the point to which $f_n(z)$ converges. Then f is a continuous function defined on U. We need to verify f is analytic. But, letting $T \subseteq U$,

$$\int_{\partial T} f(z)\,dz = \lim_{n\to\infty}\int_{\partial T} f_n(z)\,dz = 0.$$

Therefore, by Morera's theorem we see that f is analytic. As for the uniform convergence of the derivatives of f, this follows from the Cauchy integral formula. For $z \in K_n$, and $\gamma(t) \equiv z + \delta_n e^{it}, t \in [0, 2\pi]$,

$$\begin{aligned}
|f'(z) - f_k'(z)| &\leq \frac{1}{2\pi} \left| \int_\gamma \frac{f_k(w) - f(w)}{(w-z)^2} dw \right| \\
&\leq ||f_k - f|| \frac{1}{2\pi} 2\pi \delta_n \frac{1}{\delta_n^2} \\
&= ||f_k - f|| \frac{1}{\delta_n},
\end{aligned}$$

where here $||f_k - f|| \equiv \sup \{|f_k(z) - f(z)| : z \in K_n\}$. Thus we get uniform convergence of the derivatives. The consideration of the other derivatives is similar.

Since the family, $\mathcal{F}$ satisfies the conclusion of Theorem 29.1 it is known as a normal family of functions.

The following result is about a certain class of so called fractional linear transformations,

Lemma 29.2 *For $\alpha \in B(0,1)$, let*

$$\phi_\alpha(z) \equiv \frac{z - \alpha}{1 - \overline{\alpha} z}.$$

Then ϕ_α maps $B(0,1)$ one to one and onto $B(0,1)$, $\phi_\alpha^{-1} = \phi_{-\alpha}$, and

$$\phi_\alpha'(\alpha) = \frac{1}{1 - |\alpha|^2}.$$

Proof: First we show $\phi_\alpha(z) \in B(0,1)$ whenever $z \in B(0,1)$. If this is not so, there exists $z \in B(0,1)$ such that

$$|z - \alpha|^2 \geq |1 - \overline{\alpha} z|^2.$$

However, this requires

$$|z|^2 + |\alpha|^2 > 1 + |\alpha|^2 |z|^2$$

and so

$$|z|^2 \left(1 - |\alpha|^2\right) > 1 - |\alpha|^2$$

contradicting $|z| < 1$.

It remains to verify ϕ_α is one to one and onto with the given formula for ϕ_α^{-1}. But it is easy to verify $\phi_\alpha\left(\phi_{-\alpha}(w)\right) = w$. Therefore, ϕ_α is onto and one to one. To verify the formula for ϕ_α', just differentiate the function. Thus,

$$\phi_\alpha'(z) = (z - \alpha)(-1)(1 - \overline{\alpha} z)^{-2}(-\overline{\alpha}) + (1 - \overline{\alpha} z)^{-1}$$

and the formula for the derivative follows.

The next lemma, known as Schwarz's lemma is interesting for its own sake but will be an important part of the proof of the Riemann mapping theorem.

Lemma 29.3 *Suppose $F : B(0,1) \to B(0,1)$, F is analytic, and $F(0) = 0$. Then for all $z \in B(0,1)$,*

$$|F(z)| \leq |z|, \tag{29.1}$$

and

$$|F'(0)| \leq 1. \tag{29.2}$$

If equality holds in (29.2) then there exists $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ and

$$F(z) = \lambda z. \tag{29.3}$$

Proof: We know $F(z) = zG(z)$ where G is analytic. Then letting $|z| < r < 1$, the maximum modulus theorem implies

$$|G(z)| \leq \sup \frac{|F(re^{it})|}{r} \leq \frac{1}{r}.$$

Therefore, letting $r \to 1$ we get

$$|G(z)| \leq 1 \tag{29.4}$$

It follows that (29.1) holds. Since $F'(0) = G(0)$, (29.4) implies (29.2). If equality holds in (29.2), then from the maximum modulus theorem, we see that G achieves its maximum at an interior point and is consequently equal to a constant, $\lambda, |\lambda| = 1$. Thus $F(z) = z\lambda$ which shows (29.3). This proves the lemma.

Definition 29.4 *We say a region, U has the square root property if whenever $f, \frac{1}{f} : U \to \mathbb{C}$ are both analytic, it follows there exists $\phi : U \to \mathbb{C}$ such that ϕ is analytic and $f(z) = \phi^2(z)$.*

The next theorem will turn out to be equivalent to the Riemann mapping theorem.

Theorem 29.5 *Let $U \neq \mathbb{C}$ for U a region and suppose U has the square root property. Then there exists $h : U \to B(0,1)$ such that h is one to one, onto, and analytic.*

Proof: We define $\mathcal{F}$ to be the set of functions, f such that $f : U \to B(0,1)$ is one to one and analytic. We will show $\mathcal{F}$ is nonempty. Then we will show there is a function in $\mathcal{F}$, h, such that for some fixed $z_0 \in U$, $|h'(z_0)| \geq |\psi'(z_0)|$ for all $\psi \in \mathcal{F}$. When we have done this, we show h is actually onto. This will prove the theorem.

Now we begin by showing $\mathcal{F}$ is nonempty. Since $U \neq \mathbb{C}$ it follows there exists $\xi \notin U$. Then letting $f(z) \equiv z - \xi$, it follows f and $\frac{1}{f}$ are both analytic on U. Since U has the square root property, there exists $\phi : U \to \mathbb{C}$ such that $\phi^2(z) = f(z)$ for all $z \in U$. By the open mapping theorem, there exists a such that for some $r < |a|$,

$$B(a,r) \subseteq \phi(U).$$

It follows that if $z \in U$, then $\phi(z) \notin B(-a,r)$ because if this were to occur for some $z_1 \in U$, then $-\phi(z_1) \in B(a,r)$ and so there exists $z_2 \in B(a,r)$ such that

$$-\phi(z_1) = \phi(z_2).$$

Squaring both sides, it follows that $z_1 - \xi = z_2 - \xi$ and so $z_1 = z_2$. Therefore, we would have $\phi(z_2) = 0$ and so $0 \in B(a,r)$ contrary to the construction in which $r < |a|$. Now let

$$\psi(z) \equiv \frac{r}{\phi(z) + a}.$$

ψ is well defined because we just verified the denominator is nonzero. It also follows that $|\psi(z)| \leq 1$ because if not,

$$r > |\phi(z) + a|$$

for some $z \in U$, contradicting what was just shown about $\phi(U) \cap B(-a,r) = \emptyset$. Therefore, we have shown that $\mathcal{F} \neq \emptyset$.

For $z_0 \in U$ fixed, let

$$\eta \equiv \sup \left\{ |\psi'(z_0)| : \psi \in \mathcal{F} \right\}.$$

Thus $\eta > 0$ because $\psi'(z_0) \neq 0$ for ψ defined above. By Theorem 29.1, there exists a sequence, $\{\psi_n\}$, of functions in $\mathcal{F}$ and an analytic function, h, such that

$$\left|\psi_n'(z_0)\right| \to \eta(z_0) \tag{29.5}$$

and

$$\psi_n \to h, \psi_n' \to h', \tag{29.6}$$

uniformly on all compact sets of U. It follows

$$|h'(z_0)| = \lim_{n\to\infty} \left|\psi_n'(z_0)\right| = \eta$$

and for all $z \in U$,

$$|h(z)| = \lim_{n\to\infty} |\psi_n(z)| \leq 1.$$

We need to verify that h is one to one. Suppose $h(z_1) = \alpha$ and $z_2 \in U$. We must verify that $h(z_2) \neq \alpha$. We choose $r > 0$ such that $h - \alpha$ has no zeros on $\partial B(z_2, r)$, $\overline{B(z_2, r)} \subseteq U$, and

$$\overline{B(z_2, r)} \cap \overline{B(z_1, r)} = \emptyset.$$

We can do this because, the zeros of $h - \alpha$ are isolated since h is not constant due to the fact that $h'(z_0) = \eta \neq 0$. Let $\psi_n(z_1) = \alpha_n$. Thus $\psi_n - \alpha_n$ has a zero at z_1 and since ψ_n is one to one, it has no zeros in $\overline{B(z_2, r)}$. Thus by Theorem 26.6, the theorem on counting zeros, for $\gamma(t) \equiv z_2 + re^{it}, t \in [0, 2\pi]$,

$$\begin{aligned} 0 &= \lim_{n\to\infty} \frac{1}{2\pi i} \int_\gamma \frac{\psi_n'(w)}{\psi_n(w) - \alpha_n} dw \\ &= \frac{1}{2\pi i} \int_\gamma \frac{h'(w)}{h(w) - \alpha} dw, \end{aligned}$$

which shows that $h - \alpha$ has no zeros in $B(z_2, r)$. This shows that h is one to one since $z_2 \neq z_1$ was arbitrary. Therefore, $h \in \mathcal{F}$. This completes the second step of the proof. It only remains to verify that h is onto.

To show h is onto, we use the fractional linear transformation of Lemma 29.2. Suppose h is not onto. Then there exists $\alpha \in B(0, 1) \setminus h(U)$. Then $0 \notin \phi_\alpha \circ h$ because $\alpha \notin h(U)$. Therefore, since U has the square root property, there exists g, an analytic function defined on U such that

$$g^2 = \phi_\alpha \circ h.$$

The function g is one to one because if $g(z_1) = g(z_2)$, then we could square both sides and conclude that

$$\phi_\alpha \circ h(z_1) = \phi_\alpha \circ h(z_2)$$

and since ϕ_α and h are one to one, this shows $z_1 = z_2$. It follows that $g \in \mathcal{F}$ also. Now let $\psi \equiv \phi_{g(z_0)} \circ g$. Thus $\psi(z_0) = 0$. If we define $s(w) \equiv w^2$, then using Lemma 29.2, in particular, the description of $\phi_\alpha^{-1} = \phi_{-\alpha}$, we obtain

$$g = \phi_{-g(z_0)} \circ \psi$$

and therefore,

$$\begin{aligned} h(z) &= \phi_{-\alpha}\left(g^2(z)\right) \\ &= \left(\phi_{-\alpha} \circ s \circ \phi_{-g(z_0)} \circ \psi\right)(z) \\ &= (F \circ \psi)(z) \end{aligned}$$

Now $F(0) = \phi_\alpha^{-1}\left(\phi_{g(z_0)}^{-2}(0)\right) = \phi_\alpha^{-1}\left(g^2(z_0)\right) = h(z_0)$.

There are two cases to consider. First suppose that $h(z_0) \neq 0$. Then define

$$G \equiv \phi_{h(z_0)} \circ F.$$

Then $G : B(0,1) \to B(0,1)$ and $G(0) = 0$. Therefore by the Schwarz lemma, Lemma 29.3,

$$|G'(0)| = \left|\left(\frac{1}{1 - |h(z_0)|^2}\right) F'(0)\right| \leq 1$$

which implies $|F'(0)| < 1$. In the case where $h(z_0) = 0$, we note that because of the function, s, in the definition of F, F is not one to one and so we cannot have $F(z) = \lambda z$ for some $|\lambda| = 1$. Therefore, by the Schwarz lemma applied to F, we see $|F'(0)| < 1$. Therefore,

$$\begin{aligned} \eta &= |h'(z_0)| = |F'(\psi(z_0))| \left|\psi'(z_0)\right| \\ &= |F'(0)| \left|\psi'(z_0)\right| < \left|\psi'(z_0)\right|, \end{aligned}$$

contradicting the definition of η. Therefore, h must be onto and this proves the theorem.

We now give a simple lemma which will yield the usual form of the Riemann mapping theorem.

Lemma 29.6 *Let U be a simply connected region with $U \neq \mathbb{C}$. Then U has the square root property.*

Proof: Let f and $\frac{1}{f}$ both be analytic on U. Then $\frac{f'}{f}$ is analytic on U so by Corollary 25.12, there exists $\widetilde{F}$, analytic on U such that $\widetilde{F}' = \frac{f'}{f}$ on U. Then $\left(fe^{-\widetilde{F}}\right)' = 0$ and so $f(z) = Ce^{\widetilde{F}} = e^{a+ib}e^{\widetilde{F}}$. Now let $F = \widetilde{F} + a + ib$. Then F is still a primitive of f'/f and we have $f(z) = e^{F(z)}$. Now let $\phi(z) \equiv e^{\frac{1}{2}F(z)}$. Then ϕ is the desired square root and so U has the square root property.

Corollary 29.7 *(Riemann mapping theorem) Let U be a simply connected region with $U \neq \mathbb{C}$ and let $a \in U$. Then there exists a function, $f : U \to B(0,1)$ such that f is one to one, analytic, and onto with $f(a) = 0$. Furthermore, f^{-1} is also analytic.*

Proof: From Theorem 29.5 and Lemma 29.6 there exists a function, $g : U \to B(0,1)$ which is one to one, onto, and analytic. We need to show that there exists a function, f, which does what g does but in addition, $f(a) = 0$. We can do so by letting $f = \phi_{g(a)} \circ g$ if $g(a) \neq 0$. The assertion that f^{-1} is analytic follows from the open mapping theorem.

29.1 Exercises

1. Prove that in Theorem 29.1 it suffices to assume $\mathcal{F}$ is uniformly bounded on each compact subset of U.

2. Verify the conclusion of Theorem 29.1 involving the higher order derivatives.

3. What if $U = \mathbb{C}$? Does there exist an analytic function, f mapping U one to one and onto $B(0,1)$? Explain why or why not. Was $U \neq \mathbb{C}$ used in the proof of the Riemann mapping theorem?

4. Verify that $|\phi_\alpha(z)| = 1$ if $|z| = 1$. Apply the maximum modulus theorem to conclude that $|\phi_\alpha(z)| \leq 1$ for all $|z| < 1$.

5. Suppose that $|f(z)| \leq 1$ for $|z| = 1$ and $f(\alpha) = 0$ for $|\alpha| < 1$. Show that $|f(z)| \leq |\phi_\alpha(z)|$ for all $z \in B(0,1)$. **Hint:** Consider $\frac{f(z)(1-\overline{\alpha}z)}{z-\alpha}$ which has a removable singularity at α. Show the modulus of this function is bounded by 1 on $|z| = 1$. Then apply the maximum modulus theorem.

Approximation of analytic functions

Consider the function, $\frac{1}{z} = f(z)$ for z defined on $U \equiv B(0,1) \setminus \{0\}$. Clearly f is analytic on U. Suppose we could approximate f uniformly by polynomials on $\overline{\text{ann}\left(0, \frac{1}{2}, \frac{3}{4}\right)}$, a compact subset of U. Then, there would exist a suitable polynomial $p(z)$, such that $\left|\frac{1}{2\pi i}\int_\gamma f(z) - p(z)\, dz\right| < \frac{1}{10}$ where here γ is a circle of radius $\frac{2}{3}$. However, this is impossible because $\frac{1}{2\pi i}\int_\gamma f(z)\, dz = 1$ while $\frac{1}{2\pi i}\int_\gamma p(z)\, dz = 0$. This shows we cannot expect to be able to uniformly approximate analytic functions on compact sets using polynomials. It turns out we will be able to approximate by rational functions. The following lemma is the one of the key results which will allow us to verify a theorem on approximation. We will use the notation

$$||f - g||_{K,\infty} \equiv \sup\{|f(z) - g(z)| : z \in K\}$$

which describes the manner in which the approximation is measured.

Lemma 30.1 *Let R be a rational function which has a pole only at $a \in V$, a component of $\mathbb{C} \setminus K$ where K is a compact set. Suppose $b \in \overline{V}$. Then for $\varepsilon > 0$ given, there exists a rational function, Q, having a pole only at b such that*

$$||R - Q||_{K,\infty} < \varepsilon. \tag{30.1}$$

If it happens that V is unbounded, then there exists a polynomial, P such that

$$||R - P||_{K,\infty} < \varepsilon. \tag{30.2}$$

Proof: We say $b \in V$ satisfies P if for all $\varepsilon > 0$ there exists a rational function, Q_b, having a pole only at b such that

$$||R - Q_b||_{K,\infty}.$$

Now we define a set,

$$S \equiv \{b \in V : b \text{ satisfies } P\}.$$

We observe that $S \neq \emptyset$ because $a \in S$.

We now show that S is open. Suppose $b_1 \in S$. Then there exists a $\delta > 0$ such that

$$\left|\frac{b_1 - b}{z - b}\right| < \frac{1}{2} \tag{30.3}$$

for all $z \in K$ whenever $b \in B(b_1, \delta)$. If not, there would exist a sequence $b_n \to b$ for which $\left|\frac{b_1 - b_n}{\text{dist}(b_n, K)}\right| \geq \frac{1}{2}$. Then taking the limit and using the fact that $\text{dist}(b_n, K) \to \text{dist}(b, K) > 0$, (why?) we obtain a contradiction.

Since b_1 satisfies P, there exists a rational function Q_{b_1} with the desired properties. We will show we can approximate Q_{b_1} with Q_b thus yielding an approximation to R by the use of the triangle inequality,

$$||R - Q_{b_1}||_{K,\infty} + ||Q_{b_1} - Q_b||_{K,\infty} \geq ||R - Q_b||_{K,\infty}.$$

Since Q_{b_1} has poles only at b_1, it follows it is a sum of functions of the form $\frac{\alpha_n}{(z-b_1)^n}$. Therefore, it suffices to assume Q_{b_1} is of the special form

$$Q_{b_1}(z) = \frac{1}{(z-b_1)^n}.$$

However,

$$\begin{aligned} \frac{1}{(z-b_1)^n} &= \frac{1}{(z-b)^n \left(1 - \frac{b_1 - b}{z-b}\right)^n} \\ &= \frac{1}{(z-b)^n} \sum_{k=0}^{\infty} A_k \left(\frac{b_1-b}{z-b}\right)^k. \end{aligned} \tag{30.4}$$

We leave it as an exercise to find A_k and to verify using the Weierstrass M test that this series converges absolutely and uniformly on K because of the estimate (30.3) which holds for all $z \in K$. Therefore, a suitable partial sum can be made as close as desired to $\frac{1}{(z-b_1)^n}$. This shows that b satisfies P whenever b is close enough to b_1 verifying that S is open.

Next we show that S is closed in V. Let $b_n \in S$ and suppose $b_n \to b \in V$. Then for all n large enough,

$$\frac{1}{2} \operatorname{dist}(b, K) \geq |b_n - b|$$

and so we obtain the following for all n large enough.

$$\left|\frac{b-b_n}{z-b_n}\right| < \frac{1}{2},$$

for all $z \in K$. Now a repeat of the above argument in (30.4) with b_n playing the role of b_1 shows that $b \in S$. Since S is both open and closed in V it follows that, since $S \neq \emptyset$, we must have $S = V$. Otherwise V would fail to be connected.

Now let $b \in \partial V$. Then a repeat of the argument that was just given to verify that S is closed shows that b satisfies P and proves (30.1).

It remains to consider the case where V is unbounded. Since $S = V$, pick $b \in V = S$ large enough that

$$\left|\frac{z}{b}\right| < \frac{1}{2} \tag{30.5}$$

for all $z \in K$. As before, it suffices to assume that Q_b is of the form

$$Q_b(z) = \frac{1}{(z-b)^n}$$

Then we leave it as an exercise to verify that, thanks to (30.5),

$$\frac{1}{(z-b)^n} = \frac{(-1)^n}{b^n} \sum_{k=0}^{\infty} A_k \left(\frac{z}{b}\right)^k \tag{30.6}$$

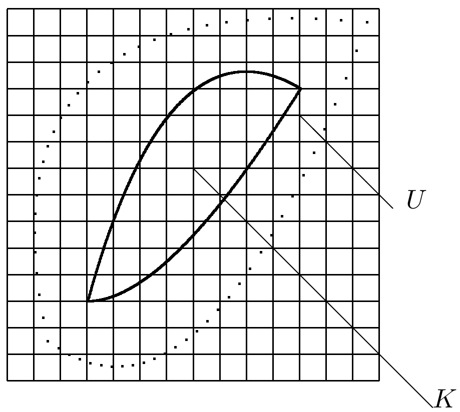
with the convergence uniform on K. Therefore, we may approximate R uniformly by a polynomial consisting of a partial sum of the above infinite sum.

The next theorem is interesting for its own sake. It gives the existence, under certain conditions, of a contour for which the Cauchy integral formula holds.

Theorem 30.2 *Let $K \subseteq U$ where K is compact and U is open. Then there exist linear mappings, $\gamma_k : [0,1] \to U \setminus K$ such that for all $z \in K$,*

$$f(z) = \frac{1}{2\pi i} \sum_{k=1}^{p} \int_{\gamma_k} \frac{f(w)}{w-z} dw. \tag{30.7}$$

Proof: Tile $\mathbb{R}^2 = \mathbb{C}$ with little squares having diameters less than δ where $0 < \delta \leq \operatorname{dist}\left(K, U^C\right)$ (see Problem 3). Now let $\{R_j\}_{j=1}^m$ denote those squares that have nonempty intersection with K. For example, see the following picture.



Let $\left\{v_j^k\right\}_{k=1}^4$ denote the four vertices of R_j where v_j^1 is the lower left, v_j^2 the lower right, v_j^3 the upper right and v_j^4 the upper left. Let $\gamma_j^k : [0,1] \to U$ be defined as

$$\begin{aligned} \gamma_j^k(t) &\equiv v_j^k + t\left(v_j^{k+1} - v_j^k\right) \text{ if } k < 4, \\ \gamma_j^4(t) &\equiv v_j^4 + t\left(v_j^1 - v_j^4\right) \text{ if } k = 4. \end{aligned}$$

Define

$$\int_{\partial R_j} g(w)\, dw \equiv \sum_{k=1}^{4} \int_{\gamma_j^k} g(w)\, dw.$$

Thus we integrate over the boundary of the square in the counter clockwise direction. Let $\left\{\gamma_j\right\}_{j=1}^p$ denote the curves, γ_j^k which have the property that $\gamma_j^k([0,1]) \cap K = \emptyset$.

Claim: $\sum_{j=1}^m \int_{\partial R_j} g(w)\, dw = \sum_{j=1}^p \int_{\gamma_j} g(w)\, dw$.

Proof of the claim: If $\gamma_j^k([0,1]) \cap K \neq \emptyset$, then for some $r \neq j$,

$$\gamma_r^l([0,1]) = \gamma_j^k([0,1])$$

but $\gamma_r^l = -\gamma_j^k$(The directions are opposite.). Hence, in the sum on the left, the only possibly nonzero contributions come from those curves, γ_j^k for which $\gamma_j^k([0,1]) \cap K = \emptyset$ and this proves the claim.

Now let $z \in K$ and suppose z is in the interior of R_s, one of these squares which intersect K. Then by the Cauchy integral formula,

$$f(z) = \frac{1}{2\pi i} \int_{\partial R_s} \frac{f(w)}{w-z} dw,$$

and if $j \neq s$,

$$0 = \frac{1}{2\pi i} \int_{\partial R_j} \frac{f(w)}{w - z} dw.$$

Therefore,

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \sum_{j=1}^{m} \int_{\partial R_j} \frac{f(w)}{w - z} dw \\ &= \frac{1}{2\pi i} \sum_{j=1}^{p} \int_{\gamma_j} \frac{f(w)}{w - z} dw. \end{aligned}$$

This proves (30.7) in the case where z is in the interior of some R_s. The general case follows from using the continuity of the functions, $f(z)$ and

$$z \to \frac{1}{2\pi i} \sum_{j=1}^{p} \int_{\gamma_j} \frac{f(w)}{w - z} dw.$$

This proves the theorem.

30.1 Runge's theorem

With the above preparation we are ready to prove the very remarkable Runge theorem which says that we can approximate analytic functions on arbitrary compact sets with rational functions which have a certain nice form. Actually, the theorem we will present first is a variant of Runge's theorem because it focuses on a single compact set.

Theorem 30.3 *Let K be a compact subset of an open set, U and let $\{b_j\}$ be a set which consists of one point from the closure of each bounded component of $\mathbb{C} \setminus K$. Let f be analytic on U. Then for each $\varepsilon > 0$, there exists a rational function, Q whose poles are all contained in the set, $\{b_j\}$ such that*

$$||Q - f||_{K,\infty} < \varepsilon. \tag{30.8}$$

Proof: By Theorem 30.2 there are curves, γ_k described there such that for all $z \in K$,

$$f(z) = \frac{1}{2\pi i} \sum_{k=1}^{p} \int_{\gamma_k} \frac{f(w)}{w - z} dw. \tag{30.9}$$

Defining $g(w, z) \equiv \frac{f(w)}{w-z}$ for $(w, z) \in \cup_{k=1}^{p} \gamma_k([0, 1]) \times K$, we see that g is uniformly continuous and so there exists a $\delta > 0$ such that if $||\mathcal{P}|| < \delta$, then for all $z \in K$,

$$\left| f(z) - \frac{1}{2\pi i} \sum_{k=1}^{p} \sum_{j=1}^{n} \frac{f(\gamma_k(\tau_j))(\gamma_k(t_i) - \gamma_k(t_{i-1}))}{\gamma_k(\tau_j) - z} \right| < \frac{\varepsilon}{2}.$$

The complicated expression is obtained by replacing each integral in (30.9) with a Riemann sum. Simplifying the appearance of this, it follows there exists a rational function of the form

$$R(z) = \sum_{k=1}^{M} \frac{A_k}{w_k - z}$$

where the w_k are elements of components of $\mathbb{C} \setminus K$ and A_k are complex numbers such that

$$||R - f||_{K,\infty} < \frac{\varepsilon}{2}.$$

Consider the rational function, $R_k(z) \equiv \frac{A_k}{w_k - z}$ where $w_k \in V_j$, one of the components of $\mathbb{C} \setminus K$, the given point of $\overline{V_j}$ being b_j or else V_j is unbounded. By Lemma 30.1, there exists a function, Q_k which is either a rational function having its only pole at b_j or a polynomial, depending on whether V_j is bounded, such that

$$||R_k - Q_k||_{K,\infty} < \frac{\varepsilon}{2M}.$$

Letting $Q(z) \equiv \sum_{k=1}^{M} Q_k(z)$,

$$||R - Q||_{K,\infty} < \frac{\varepsilon}{2}.$$

It follows

$$||f - Q||_{K,\infty} \leq ||f - R||_{K,\infty} + ||R - Q||_{K,\infty} < \varepsilon.$$

This proves the theorem.

Runge's theorem concerns the case where the given points are contained in $\mathbb{C} \setminus U$ for U an open set rather than a compact set. Note that here there could be uncountably many components of $\mathbb{C} \setminus U$ because the components are no longer open sets. An easy example of this phenomenon in one dimension is where $U = [0,1] \setminus P$ for P the Cantor set. Then you can show that $\mathbb{R} \setminus U$ has uncountably many components. Nevertheless, Runge's theorem will follow from Theorem 30.3 with the aid of the following interesting lemma.

Lemma 30.4 *Let U be an open set in $\mathbb{C}$. Then there exists a sequence of compact sets, $\{K_n\}$ such that*

$$U = \cup_{k=1}^{\infty} K_n, \ \cdots, K_n \subseteq \operatorname{int} K_{n+1} \cdots, \tag{30.10}$$

and for any $K \subseteq U$,

$$K \subseteq K_n, \tag{30.11}$$

for all n sufficiently large, and every component of $\widehat{\mathbb{C}} \setminus K_n$ contains a component of $\widehat{\mathbb{C}} \setminus U$.

Proof: Let

$$V_n \equiv \{z : |z| > n\} \cup \bigcup_{z \notin U} B\left(z, \frac{1}{n}\right).$$

Thus $\{z : |z| > n\}$ contains the point, ∞. Now let

$$K_n \equiv \widehat{\mathbb{C}} \setminus V_n = \mathbb{C} \setminus V_n \subseteq U.$$

We leave it as an exercise to verify that (30.10) and (30.11) hold. It remains to show that every component of $\widehat{\mathbb{C}} \setminus K_n$ contains a component of $\widehat{\mathbb{C}} \setminus U$. Let D be a component of $\widehat{\mathbb{C}} \setminus K_n \equiv V_n$.

If $\infty \notin D$, then D contains no point of $\{z : |z| > n\}$ because this set is connected and D is a component. (If it did contain a point of this set, it would have to contain the whole set..) Therefore, $D \subseteq \bigcup_{z \notin U} B\left(z, \frac{1}{n}\right)$ and so D contains some point of $B\left(z, \frac{1}{n}\right)$ for some $z \notin U$. Therefore, since this ball is connected, it follows D must contain the whole ball and consequently D contains some point of U^C. (The point z at the center of the ball will do.) Since D contains $z \notin U$, it must contain the component, H_z, determined by this point. The reason for this is that

$$H_z \subseteq \widehat{\mathbb{C}} \setminus U \subseteq \widehat{\mathbb{C}} \setminus K_n$$

and H_z is connected. Therefore, H_z can only have points in one component of $\widehat{\mathbb{C}} \setminus K_n$. Since it has a point in D, it must therefore, be totally contained in D. This verifies the desired condition in the case where $\infty \notin D$.

Now suppose that $\infty \in D$. We know that $\infty \notin U$ because U is given to be a set in $\mathbb{C}$. Letting H_∞ denote the component of $\widehat{\mathbb{C}} \setminus U$ determined by ∞, it follows from similar reasoning to the above that $H_\infty \subseteq D$ and this proves the lemma.

Theorem 30.5 *(Runge) Let U be an open set, and let A be a set which has one point in each bounded component of $\widehat{\mathbb{C}} \setminus U$ and let f be analytic on U. Then there exists a sequence of rational functions, $\{R_n\}$ having poles only in A such that R_n converges uniformly to f on compact subsets of U.*

Proof: Let K_n be the compact sets of Lemma 30.4 where each component of $\widehat{\mathbb{C}} \setminus K_n$ contains a component of $\widehat{\mathbb{C}} \setminus U$. It follows each bounded component of $\widehat{\mathbb{C}} \setminus K_n$ contains a point of A. Therefore, by Theorem 30.3 there exists R_n a rational function with poles only in A such that

$$||R_n - f||_{K_n,\infty} < \frac{1}{n}.$$

It follows, since a given compact set, K is a subset of K_n for all n large enough, that $R_n \to f$ uniformly on K. This proves the theorem.

Corollary 30.6 *Let U be simply connected and f is analytic on U. Then there exists a sequence of polynomials, $\{p_n\}$ such that $p_n \to f$ uniformly on compact sets of U.*

Proof: By definition of what is meant by simply connected, $\widehat{\mathbb{C}} \setminus U$ is connected and so there are no bounded components of $\widehat{\mathbb{C}} \setminus U$. Therefore, $A = \emptyset$ and it follows that R_n in the above theorem must be a polynomial since it is rational and has no poles.

30.2 Exercises

1. Let K be any nonempty set in $\mathbb{C}$ and define

$$\text{dist}(z, K) \equiv \inf \{|z - w| : w \in K\}.$$

   Show that $z \to \text{dist}(z, K)$ is a continuous function.

2. Verify the series in (30.4) converges absolutely on K and find A_k. Also do the same for (30.6). **Hint:** You know that for $|z| < 1$, $\frac{1}{1-z} = \sum_{k=0}^{\infty} z^k$. Differentiate both sides as many times as needed to obtain a formula for A_k. Then apply the Weierstrass M test and the ratio test.

3. In Theorem 30.2 we had a compact set, K contained in an open set U and we used the fact that

$$\text{dist}\left(K, U^C\right) \equiv \inf\left\{|z - w| : w \in U^C, z \in K\right\} > 0.$$

   Prove this.

4. For $U = [0, 1] \setminus P$ for P the Cantor set, show that $\mathbb{R} \setminus U$ has uncountably many components. **Hint:** Show that the component of $\mathbb{R} \setminus U$ determined by $p \in P$, is just the single point, p and then show P is uncountable.

5. In the proof of Lemma 30.4, verify that (30.10) and (30.11) are satisfied for the given choice of K_n.

The Hausdorff Maximal theorem

The purpose of this appendix is to prove the equivalence between the axiom of choice, the Hausdorff maximal theorem, and the well-ordering principle. The Hausdorff maximal theorem and the well-ordering principle are very useful but a little hard to believe; so, it may be surprising that they are equivalent to the axiom of choice. First we give a proof that the axiom of choice implies the Hausdorff maximal theorem, a remarkable theorem about partially ordered sets.

We say that a nonempty set is partially ordered if there exists a partial order, $\prec$, satisfying

$$x \prec x$$

and

$$\text{if } x \prec y \text{ and } y \prec z \text{ then } x \prec z.$$

An example of a partially ordered set is the set of all subsets of a given set and $\prec \equiv \subseteq$. Note that we can not conclude that any two elements in a partially ordered set are related. In other words, just because $x,\ y$ are in the partially ordered set, it does not follow that either $x \prec y$ or $y \prec x$. We call a subset of a partially ordered set, $\mathcal{C}$, a chain if $x,\ y \in \mathcal{C}$ implies that either $x \prec y$ or $y \prec x$. If either $x \prec y$ or $y \prec x$ we say that x and y are comparable. A chain is also called a totally ordered set. We say $\mathcal{C}$ is a maximal chain if whenever $\widetilde{\mathcal{C}}$ is a chain containing $\mathcal{C}$, it follows the two chains are equal. In other words $\mathcal{C}$ is a maximal chain if there is no strictly larger chain.

Lemma A.1 *Let $\mathcal{F}$ be a nonempty partially ordered set with partial order $\prec$. Then assuming the axiom of choice, there exists a maximal chain in $\mathcal{F}$.*

Proof: Let $\mathcal{X}$ be the set of all chains from $\mathcal{F}$. For $\mathcal{C} \in \mathcal{X}$, let

$$S_{\mathcal{C}} = \{x \in \mathcal{F} \text{ such that } \mathcal{C} \cup \{x\} \text{ is a chain strictly larger than } \mathcal{C}\}.$$

If $S_{\mathcal{C}} = \emptyset$ for any $\mathcal{C}$, then $\mathcal{C}$ is maximal and we are done. Thus, assume $S_{\mathcal{C}} \neq \emptyset$ for all $\mathcal{C} \in \mathcal{X}$. Let $f(\mathcal{C}) \in S_{\mathcal{C}}$. (This is where the axiom of choice is being used.) Let

$$g(\mathcal{C}) = \mathcal{C} \cup \{f(\mathcal{C})\}.$$

Thus $g(\mathcal{C}) \supsetneq \mathcal{C}$ and $g(\mathcal{C}) \setminus \mathcal{C} = \{f(\mathcal{C})\} = \{\text{a single element of } \mathcal{F}\}$. We call a subset $\mathcal{T}$ of $\mathcal{X}$ a tower if

$$\emptyset \in \mathcal{T},$$

$$\mathcal{C} \in \mathcal{T} \text{ implies } g(\mathcal{C}) \in \mathcal{T},$$

and if $\mathcal{S} \subseteq \mathcal{T}$ is totally ordered with respect to set inclusion, then

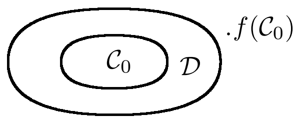
$$\cup \mathcal{S} \in \mathcal{T}.$$

Here $\mathcal{S}$ is a chain with respect to set inclusion whose elements are chains.

Note that $\mathcal{X}$ is a tower. Let $\mathcal{T}_0$ be the intersection of all towers. Thus, $\mathcal{T}_0$ is a tower, the smallest tower. We wish to show that any two sets in $\mathcal{T}_0$ are comparable in the sense of set inclusion so that $\mathcal{T}_0$ is actually a chain. Let $\mathcal{C}_0$ be a set of $\mathcal{T}_0$ which is comparable to every set of $\mathcal{T}_0$. Such sets exist, $\emptyset$ being an example. Let

$$\mathcal{B} \equiv \{\mathcal{D} \in \mathcal{T}_0 : \mathcal{D} \supsetneq \mathcal{C}_0 \text{ and } f(\mathcal{C}_0) \notin \mathcal{D}\}.$$

The picture represents sets of $\mathcal{B}$. As illustrated in the picture, $\mathcal{D}$ is a set of $\mathcal{B}$ when $\mathcal{D}$ is larger than $\mathcal{C}_0$ but fails to be comparable to $g(\mathcal{C}_0)$. Thus there would be more than one chain ascending from $\mathcal{C}_0$ if $\mathcal{B} \neq \emptyset$, rather like a tree growing upward in more than one direction from a fork in the trunk. We will show this can't take place for any such $\mathcal{C}_0$ by showing $\mathcal{B} = \emptyset$.



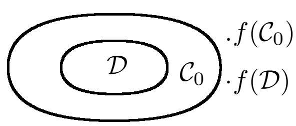
This will be accomplished by showing $\widetilde{\mathcal{T}_0} \equiv \mathcal{T}_0 \setminus \mathcal{B}$ is a tower. Since $\mathcal{T}_0$ is the smallest tower, this will require that $\widetilde{\mathcal{T}_0} = \mathcal{T}_0$ and so $\mathcal{B} = \emptyset$.

Claim: $\widetilde{\mathcal{T}_0}$ is a tower and so $\mathcal{B} = \emptyset$.

Proof of the claim: It is clear that $\emptyset \in \widetilde{\mathcal{T}_0}$ because for $\emptyset$ to be contained in $\mathcal{B}$ it would be required to properly contain $\mathcal{C}_0$ which is not possible. Suppose $\mathcal{D} \in \widetilde{\mathcal{T}_0}$. We need to verify $g(\mathcal{D}) \in \widetilde{\mathcal{T}_0}$.

Case 1: $f(\mathcal{D}) \in \mathcal{C}_0$. If $\mathcal{D} \subseteq \mathcal{C}_0$, then since both $\mathcal{D}$ and $\{f(\mathcal{D})\}$ are contained in $\mathcal{C}_0$, it follows $g(\mathcal{D}) \subseteq \mathcal{C}_0$ and so $g(\mathcal{D}) \notin \mathcal{B}$. On the other hand, if $\mathcal{D} \supsetneq \mathcal{C}_0$, then since $\mathcal{D} \in \widetilde{\mathcal{T}_0}$, we know $f(\mathcal{C}_0) \in \mathcal{D}$ and so $g(\mathcal{D})$ also contains $f(\mathcal{C}_0)$ implying $g(\mathcal{D}) \notin \mathcal{B}$. These are the only two cases to consider because we are given that $\mathcal{C}_0$ is comparable to every set of $\mathcal{T}_0$.

Case 2: $f(\mathcal{D}) \notin \mathcal{C}_0$. If $\mathcal{D} \subsetneq \mathcal{C}_0$ then we can't have $f(\mathcal{D}) \notin \mathcal{C}_0$ because if this were so, $g(\mathcal{D}\ )$ would not compare to $\mathcal{C}_0$.



Hence if $f(\mathcal{D}) \notin \mathcal{C}_0$, then $\mathcal{D} \supseteq \mathcal{C}_0$. If $\mathcal{D} = \mathcal{C}_0$, then $f(\mathcal{D}) = f(\mathcal{C}_0) \in g(\mathcal{D})$ so $g(\mathcal{D}) \notin \mathcal{B}$. Therefore, assume $\mathcal{D} \supsetneq \mathcal{C}_0$. Then, since $\mathcal{D}$ is in $\widetilde{\mathcal{T}_0}$, $f(\mathcal{C}_0) \in \mathcal{D}$ and so $f(\mathcal{C}_0) \in g(\mathcal{D})$. Therefore, $g(\mathcal{D}) \in \widetilde{\mathcal{T}_0}$.

Now suppose $\mathcal{S}$ is a totally ordered subset of $\widetilde{\mathcal{T}_0}$ with respect to set inclusion. Then if every element of $\mathcal{S}$ is contained in $\mathcal{C}_0$, so is $\cup\mathcal{S}$ and so $\cup\mathcal{S} \in \widetilde{\mathcal{T}_0}$. If, on the other hand, some chain from $\mathcal{S}$, $\mathcal{C}$, contains $\mathcal{C}_0$ properly, then since $\mathcal{C} \notin \mathcal{B}$, $f(\mathcal{C}_0) \in \mathcal{C} \subseteq \cup\mathcal{S}$ showing that $\cup\mathcal{S} \notin \mathcal{B}$ also. This has proved $\widetilde{\mathcal{T}_0}$ is a tower and since $\mathcal{T}_0$ is the smallest tower, it follows $\widetilde{\mathcal{T}_0} = \mathcal{T}_0$. We have shown roughly that no splitting into more than one ascending chain can occur at any $\mathcal{C}_0$ which is comparable to every set of $\mathcal{T}_0$. Next we will show that every element of $\mathcal{T}_0$ has the property that it is comparable to all other elements of $\mathcal{T}_0$. We will do so by showing that these elements which possess this property form a tower.

Define $\mathcal{T}_1$ to be the set of all elements of $\mathcal{T}_0$ which are comparable to every element of $\mathcal{T}_0$. (Recall the elements of $\mathcal{T}_0$ are chains from the original partial order.)

Claim: $\mathcal{T}_1$ is a tower.

Proof of the claim: It is clear that $\emptyset \in \mathcal{T}_1$ because $\emptyset$ is a subset of every set. Suppose $\mathcal{C}_0 \in \mathcal{T}_1$. We need to verify that $g(\mathcal{C}_0) \in \mathcal{T}_1$. Let $\mathcal{D} \in \mathcal{T}_0$ (Thus $\mathcal{D} \subseteq \mathcal{C}_0$ or else $\mathcal{D} \supsetneq \mathcal{C}_0$.)and consider $g(\mathcal{C}_0) \equiv \mathcal{C}_0 \cup \{f(\mathcal{C}_0)\}$. If $\mathcal{D} \subseteq \mathcal{C}_0$, then $\mathcal{D} \subseteq g(\mathcal{C}_0)$ so $g(\mathcal{C}_0)$ is comparable to $\mathcal{D}$. If $\mathcal{D} \supsetneq \mathcal{C}_0$, then $\mathcal{D} \supseteq g(\mathcal{C}_0)$ by what was just shown ($\mathcal{B} = \emptyset$). Hence $g(\mathcal{C}_0)$ is comparable to $\mathcal{D}$. Since $\mathcal{D}$ was arbitrary, it follows $g(\mathcal{C}_0)$ is comparable to every set of $\mathcal{T}_0$. Now suppose $\mathcal{S}$ is a chain of elements of $\mathcal{T}_1$ and let $\mathcal{D}$ be an element of $\mathcal{T}_0$. If every element in the chain, $\mathcal{S}$ is contained in $\mathcal{D}$, then $\cup\mathcal{S}$ is also contained in $\mathcal{D}$. On the other hand, if some set, $\mathcal{C}$, from $\mathcal{S}$ contains $\mathcal{D}$ properly, then $\cup\mathcal{S}$ also contains $\mathcal{D}$. Thus $\cup\mathcal{S} \in \mathcal{T}_1$ since it is comparable to every $\mathcal{D} \in \mathcal{T}_0$.

This shows $\mathcal{T}_1$ is a tower and proves therefore, that $\mathcal{T}_0 = \mathcal{T}_1$. Thus every set of $\mathcal{T}_0$ compares with every other set of $\mathcal{T}_0$ showing $\mathcal{T}_0$ is a chain in addition to being a tower.

Now $\cup\mathcal{T}_0, g(\cup\mathcal{T}_0) \in \mathcal{T}_0$. Hence, because $g(\cup\mathcal{T}_0)$ is an element of $\mathcal{T}_0$, and $\mathcal{T}_0$ is a chain of these, it follows $g(\cup\mathcal{T}_0) \subseteq \cup\mathcal{T}_0$. Thus

$$\cup\mathcal{T}_0 \supseteq g(\cup\mathcal{T}_0) \supsetneq \cup\mathcal{T}_0,$$

a contradiction. Hence there must exist a maximal chain after all. This proves the lemma.

If X is a nonempty set, we say $\leq$ is an order on X if

$$x \leq x,$$

and if $x,\ y \in X$, then

$$\text{either } x \leq y \text{ or } y \leq x$$

and

$$\text{if } x \leq y \text{ and } y \leq z \text{ then } x \leq z.$$

We say that $\leq$ is a well order and say that $(X, \leq)$ is a well-ordered set if every nonempty subset of X has a smallest element. More precisely, if $S \neq \emptyset$ and $S \subseteq X$ then there exists an $x \in S$ such that $x \leq y$ for all $y \in S$. A familiar example of a well-ordered set is the natural numbers.

Lemma A.2 *The Hausdorff maximal principle implies every nonempty set can be well-ordered.*

Proof: Let X be a nonempty set and let $a \in X$. Then $\{a\}$ is a well-ordered subset of X. Let

$$\mathcal{F} = \{S \subseteq X : \text{ there exists a well order for } S\}.$$

Thus $\mathcal{F} \neq \emptyset$. We will say that for $S_1,\ S_2 \in \mathcal{F}$, $S_1 \prec S_2$ if $S_1 \subseteq S_2$ and there exists a well order for S_2, $\leq_2$ such that

$$(S_2, \leq_2) \text{ is well-ordered}$$

and if

$$y \in S_2 \setminus S_1 \text{ then } x \leq_2 y \text{ for all } x \in S_1,$$

and if $\leq_1$is the well order of S_1 then the two orders are consistent on S_1. Then we observe that $\prec$ is a partial order on $\mathcal{F}$. By the Hausdorff maximal principle, we let $\mathcal{C}$ be a maximal chain in $\mathcal{F}$ and let

$$X_\infty \equiv \cup\mathcal{C}.$$

We also define an order, $\leq$, on X_∞ as follows. If $x,\ y$ are elements of X_∞, we pick $S \in \mathcal{C}$ such that $x,\ y$ are both in S. Then if $\leq_S$ is the order on S, we let $x \leq y$ if and only if $x \leq_S y$. This definition is well defined because of the definition of the order, $\prec$. Now let U be any nonempty subset of X_∞. Then $S \cap U \neq \emptyset$ for some $S \in \mathcal{C}$. Because of the definition of $\leq$, if $y \in S_2 \setminus S_1,\ S_i \in \mathcal{C}$, then $x \leq y$ for all $x \in S_1$. Thus, if $y \in X_\infty \setminus S$ then $x \leq y$ for all $x \in S$ and so the smallest element of $S \cap U$ exists and is the smallest element in U. Therefore X_∞ is well-ordered. Now suppose there exists $z \in X \setminus X_\infty$. Define the following order, $\leq_1$, on $X_\infty \cup \{z\}$.

$$x \leq_1 y \text{ if and only if } x \leq y \text{ whenever } x, y \in X_\infty$$

$$x \leq_1 z \text{ whenever } x \in X_\infty.$$

Then let

$$\widetilde{\mathcal{C}} = \{S \in \mathcal{C} \text{ or } X_\infty \cup \{z\}\}.$$

Then $\widetilde{\mathcal{C}}$ is a strictly larger chain than $\mathcal{C}$ contradicting maximality of $\mathcal{C}$. Thus $X \setminus X_\infty = \emptyset$ and this shows X is well-ordered by $\leq$. This proves the lemma.

With these two lemmas we can now state the main result.

Theorem A.3 *The following are equivalent.*

The axiom of choice

The Hausdorff maximal principle

The well-ordering principle.

Proof: It only remains to prove that the well-ordering principle implies the axiom of choice. Let I be a nonempty set and let X_i be a nonempty set for each $i \in I$. Let $X = \cup\{X_i : i \in I\}$ and well order X. Let $f(i)$ be the smallest element of X_i. Then

$$f \in \prod_{i \in I} X_i.$$

A.1 Exercises

1. Zorn's lemma states that in a nonempty partially ordered set, if every chain has an upper bound, there exists a maximal element, x in the partially ordered set. When we say x is maximal, we mean that if $x \prec y$, it follows $y = x$. Show Zorn's lemma is equivalent to the Hausdorff maximal theorem.

2. Let X be a vector space. We say $Y \subseteq X$ is a Hamel basis if every element of X can be written in a unique way as a finite linear combination of elements in Y. Show that every vector space has a Hamel basis and that if Y, Y_1 are two Hamel bases of X, then there exists a one to one and onto map from Y to Y_1.

3. $\uparrow$ Using the Baire category theorem of Chapter 14 show that any Hamel basis of a Banach space is either finite or uncountable.

4. $\uparrow$ Consider the vector space of all polynomials defined on $[0,1]$. Does there exist a norm, $||\cdot||$ defined on these polynomials such that with this norm, the vector space of polynomials becomes a Banach space (complete normed vector space)?

Bibliography

[1] **Adams R.** *Sobolev Spaces*, Academic Press, New York, San Francisco, London, 1975.

[2] **Apostol T.** *Calculus Volume II Second editioin,* Wiley *1969.*

[3] **Apostol, T. M.** *Mathematical Analysis.* Addison Wesley Publishing Co., 1974.

[4] **Bruckner A. , Bruckner J., and Thomson B.,** *Real Analysis* Prentice Hall 1997.

[5] **Birkoff G.** and **MacLane S.** *A Survey of Modern Algebra,* Macmillan 1977.

[6] **Conway J. B.** *Functions of one Complex variable Second edition,* Springer Verlag 1978.

[7] **Deimling K.** *Nonlinear Functional Analysis,* Springer-Verlag, 1985.

[8] **Dontchev A.L.** The Graves theorem Revisited, *Journal of Convex Analysis,* Vol. 3, 1996, No.1, 45-53.

[9] **Dunford N.** and **Schwartz J.T.** *Linear Operators,* Interscience Publishers, a division of John Wiley and Sons, New York, part 1 1958, part 2 1963, part 3 1971.

[10] **Edwards C.** *Advanced Calculus of Several Variables*, Dover 1994.

[11] **Evans L.C.** and **Gariepy,** *Measure Theory and Fine Properties of Functions,* CRC Press, 1992.

[12] **Federer H.,** *Geometric Measure Theory,* Springer-Verlag, New York, 1969.

[13] **Gaskill H.** and **Narayanaswami P.** *Elements of Real Analysis*, Prentice Hall 1998.

[14] **Hardy G.** *A Course Of Pure Mathematics, Tenth edition* Cambridge University Press 1992.

[15] **Hewitt E.** and **Stromberg K.** *Real and Abstract Analysis,* Springer-Verlag, New York, 1965.

[16] **Hoffman K. and Kunze R**. *Linear Algebra* Prentice Hall 1971.

[17] **Ince, E.L.**, *Ordinary Differential Equations,* Dover 1956

[18] **Jones F.,** *Lebesgue Integration on Euclidean Space,* Jones and Bartlett 1993.

[19] **Kuttler K.L.** *Modern Analysis* CRC Press 1998.

[20] **Levinson, N. and Redheffer, R**. *Complex Variables,* Holden Day, Inc. 1970

[21] **Nobel B. and Daniel J**. *Applied Linear Algebra,* Prentice Hall 1977.

[22] **Ray W.O.** *Real Analysis,* Prentice-Hall, 1988.

[23] **Rudin W.** *Principles of Mathematical Analysis*, McGraw Hill, 1976.

[24] **Rudin W.** *Real and Complex Analysis,* third edition, McGraw-Hill, 1987.

[25] **Rudin W.** *Functional Analysis,* second edition, McGraw-Hill, 1991.

[26] **Silverman R**. *Complex Analysis with Applications*, Dover, 1974.

[27] **Spiegel, M.R.** *Complex Variables* Schaum's outline series, McGraw Hill, 1998.

[28] **Wade W.R.** *An Introduction to Analysis,* Prentice Hall, 1995.

[29] **Yosida K.** *Functional Analysis,* Springer-Verlag, New York, 1978.

Index