Analysis of Functions of one Variable

Kenneth Kuttler klkuttler@gmail.com

March 7, 2024

Contents

Chapter 1

Introduction

The difference between advanced calculus and calculus is that all the theorems are proved completely and the role of plane geometry is minimized. Instead, the notion of completeness is of preeminent importance. Formal manipulations are of no significance at all unless they aid in showing something significant. Routine skills involving elementary functions and integration techniques are supposed to be mastered and have no place in advanced calculus which deals with the fundamental issues related to existence and meaning. This is a subject which places calculus as part of mathematics and involves proofs and definitions, not algorithms and busy work. Roughly speaking, it is nineteenth century calculus rather than eighteenth century calculus. This book is **not intended** to be a first course in Calculus. It fails to discuss many of the important techniques and applications of calculus in favor of these more theoretical considerations.

An orderly development of the elementary functions independent of geometry is included but I assume the reader is familiar enough with these functions to use them in problems which illustrate some of the ideas presented. I have placed the construction of the real numbers at the end to conform with the historical development of analysis. Completeness of the real line was used as an axiom and all the classical major theorems proved long before Dedekind and Cantor showed how to construct the real numbers from the rational numbers. However, this could be presented earlier.

There is also a brief discussion of complex analysis of functions of a complex variable and a few other somewhat unusual topics like the generalized Riemann integral which ties in very well with the nineteenth century ideas although it dates from the 1950's. Probably the generalized Riemann integral should be included much more than it is in presentations of Calculus.

Chapter 2

The Real and Complex Numbers

2.1 Real and Rational Numbers

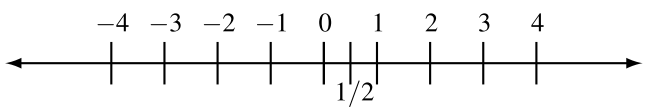
To begin with, consider the real numbers, denoted by $\mathbb{R}$, as a line extending infinitely far in both directions. In this book, the notation, $\equiv$ indicates something is being defined. Thus the integers are defined as

$$\mathbb{Z} \equiv \{\cdots - 1, 0, 1, \cdots\},$$

the natural numbers, $\mathbb{N} \equiv \{1, 2, \cdots\}$ and the rational numbers, defined as the numbers which are the quotient of two integers.

$$\mathbb{Q} \equiv \left\{\frac{m}{n} \text{ such that } m, n \in \mathbb{Z}, n \neq 0\right\}$$

are each subsets of $\mathbb{R}$ as indicated in the following picture.



As shown in the picture, $\frac{1}{2}$ is half way between the number 0 and the number, 1. By analogy, you can see where to place all the other rational numbers. It is assumed that $\mathbb{R}$ has the following algebra properties, listed here as a collection of assertions called axioms. These properties will not be proved which is why they are called axioms rather than theorems. In general, axioms are statements which are regarded as true. Often these are things which are "self evident" either from experience or from some sort of intuition but this does not have to be the case. We always assume $0 \neq 1$ because if not, you would end up with $x = x1 = x0 = 0$ for all x and we are not interested in such a stupid thing.

Axiom 2.1.1 $x + y = y + x$, *(commutative law for addition)*

Axiom 2.1.2 $x + 0 = x$, *(additive identity).*

Axiom 2.1.3 *For each* $x \in \mathbb{R}$, *there exists* $-x \in \mathbb{R}$ *such that* $x + (-x) = 0$, *(existence of additive inverse).*

Axiom 2.1.4 $(x + y) + z = x + (y + z)$, *(associative law for addition).*

Axiom 2.1.5 $xy = yx$, *(commutative law for multiplication).*

Axiom 2.1.6 $(xy)z = x(yz)$, *(associative law for multiplication).*

Axiom 2.1.7 $1x = x$, *(multiplicative identity).*

Axiom 2.1.8 *For each* $x \neq 0$, *there exists* x^{-1} *such that* $xx^{-1} = 1$.*(existence of multiplicative inverse).*

Axiom 2.1.9 $x(y + z) = xy + xz$.*(distributive law).*

These axioms are known as the field axioms and any set (there are many others besides $\mathbb{R}$) which has two such operations satisfying the above axioms is called a field. Division and subtraction are defined in the usual way by $x - y \equiv x + (-y)$ and $x/y \equiv x\left(y^{-1}\right)$. It is assumed that the reader is completely familiar with these axioms in the sense that he or she can do the usual algebraic manipulations taught in high school and junior high algebra courses. The axioms listed above are just a careful statement of exactly what is necessary to make the usual algebraic manipulations valid. A word of advice regarding division and subtraction is in order here. Whenever you feel a little confused about an algebraic expression which involves division or subtraction, think of division as multiplication by the multiplicative inverse as just indicated and think of subtraction as addition of the additive inverse. Thus, when you see x/y, think $x\left(y^{-1}\right)$ and when you see $x - y$, think $x + (-y)$. In many cases the source of confusion will disappear almost magically. The reason for this is that subtraction and division do not satisfy the associative law. This means there is a natural ambiguity in an expression like $6 - 3 - 4$. Do you mean $(6-3)-4 = -1$ or $6 - (3-4) = 6 - (-1) = 7$? It makes a difference doesn't it? However, the so called binary operations of addition and multiplication are associative and so no such confusion will occur. It is conventional to simply do the operations in order of appearance reading from left to right. Thus, if you see $6-3-4$, you would normally interpret it as the first of the above alternatives. This is no problem for English speakers, but what if you grew up speaking Hebrew or Arabic in which you read from right to left?

In the first part of the following theorem, the claim is made that the additive inverse and the multiplicative inverse are unique. This means that for a given number, only one number has the property that it is an additive inverse and that, given a nonzero number, only one number has the property that it is a multiplicative inverse. The significance of this is that if you are wondering if a given number is the additive inverse of a given number, all you have to do is to check and see if it acts like one.

Theorem 2.1.10 *The above axioms imply the following.*

1. *The multiplicative inverse and additive inverses are unique.*
2. $0x = 0,\ -(-x) = x,$
3. $(-1)(-1) = 1,\ (-1)x = -x$
4. *If* $xy = 0$ *then either* $x = 0$ *or* $y = 0$.

Proof: Suppose then that x is a real number and that $x + y = 0 = x + z$. It is necessary to verify $y = z$. From the above axioms, there exists an additive inverse, $-x$ for x. Therefore,

$$-x + 0 = (-x) + (x + y) = (-x) + (x + z)$$

and so by the associative law for addition,

$$((-x) + x) + y = ((-x) + x) + z$$

which implies $0 + y = 0 + z$. Now by the definition of the additive identity, this implies $y = z$. You should prove the multiplicative inverse is unique.

Consider 2. It is desired to verify $0x = 0$. From the definition of the additive identity and the distributive law it follows that

$$0x = (0 + 0)x = 0x + 0x.$$

From the existence of the additive inverse and the associative law it follows

$$\begin{aligned} 0 &= (-0x) + 0x = (-0x) + (0x + 0x) \\ &= ((-0x) + 0x) + 0x = 0 + 0x = 0x \end{aligned}$$

To verify the second claim in 2., it suffices to show x acts like the additive inverse of $-x$ in order to conclude that $-(-x) = x$. This is because it has just been shown that additive inverses are unique. By the definition of additive inverse, $x + (-x) = 0$ and so $x = -(-x)$ as claimed.

To demonstrate 3., $(-1)(1 + (-1)) = (-1)0 = 0$ and so using the definition of the multiplicative identity, and the distributive law, $(-1) + (-1)(-1) = 0$. It follows from 1. and 2. that $1 = -(-1) = (-1)(-1)$. To verify $(-1)x = -x$, use 2. and the distributive law to write

$$x + (-1)x = x(1 + (-1)) = x0 = 0.$$

Therefore, by the uniqueness of the additive inverse proved in 1., it follows $(-1)x = -x$ as claimed.

To verify 4., suppose $x \neq 0$. Then x^{-1} exists by the axiom about the existence of multiplicative inverses. Therefore, by 2. and the associative law for multiplication,

$$y = \left(x^{-1}x\right)y = x^{-1}(xy) = x^{-1}0 = 0.$$

This proves 4. ■

Recall the notion of something raised to an integer power. Thus $y^2 = y \times y$ and $b^{-3} = \frac{1}{b^3}$ etc.

Also, there are a few **conventions** related to the order in which operations are performed. Exponents are always done before multiplication. Thus $xy^2 = x\left(y^2\right)$ and is not equal to $(xy)^2$. Division or multiplication is always done before addition or subtraction. Thus $x - y(z + w) = x - [y(z + w)]$ and is not equal to $(x - y)(z + w)$. Parentheses are done before anything else. Be very careful of such things since they are a source of mistakes. When you have doubts, insert parentheses to resolve the ambiguities.

Also recall summation notation.

Definition 2.1.11 *Let $x_1, x_2, \cdots, x_m$ be numbers. Then $\sum_{j=1}^{m} x_j \equiv x_1 + x_2 + \cdots + x_m$. Thus this symbol, $\sum_{j=1}^{m} x_j$ means to take the numbers, $x_1, x_2, \cdots, x_m$ and add them all together. Note the use of the j as a generic variable which takes values from 1 up to m. This notation will be used whenever there are things which can be added, not just numbers.*

As an example of the use of this notation, you should verify the following.

Example 2.1.12 $\sum_{k=1}^{6}(2k+1) = 48$.

Be sure you understand why $\sum_{k=1}^{m+1} x_k = \sum_{k=1}^{m} x_k + x_{m+1}$. As a slight generalization of this notation, $\sum_{j=k}^{m} x_j \equiv x_k + \cdots + x_m$. It is also possible to change the variable of summation. $\sum_{j=1}^{m} x_j = x_1 + x_2 + \cdots + x_m$ while if r is an integer, the notation requires $\sum_{j=1+r}^{m+r} x_{j-r} = x_1 + x_2 + \cdots + x_m$ and so $\sum_{j=1}^{m} x_j = \sum_{j=1+r}^{m+r} x_{j-r}$.

Summation notation will be used throughout the book whenever it is convenient to do so.

2.2 Exercises

1. Consider the expression $x+y(x+y)-x(y-x) \equiv f(x,y)$. Find $f(-1,2)$.

2. Show $-(ab)=(-a)b$.

3. Show on the number line the effect of multiplying a number by -1.

4. Add the fractions $\frac{x}{x^2-1}+\frac{x-1}{x+1}$.

5. Find a formula for $(x+y)^2$, $(x+y)^3$, and $(x+y)^4$. Based on what you observe for these, give a formula for $(x+y)^8$.

6. When is it true that $(x+y)^n = x^n + y^n$?

7. Find the error in the following argument. Let $x=y=1$. Then $xy=y^2$ and so $xy-x^2=y^2-x^2$. Therefore, $x(y-x)=(y-x)(y+x)$. Dividing both sides by $(y-x)$ yields $x=x+y$. Now substituting in what these variables equal yields $1=1+1$.

8. Find the error in the following argument. $\sqrt{x^2+1}=x+1$ and so letting $x=2$, $\sqrt{5}=3$. Therefore, $5=9$.

9. Find the error in the following. Let $x=1$ and $y=2$. Then $\frac{1}{3}=\frac{1}{x+y}=\frac{1}{x}+\frac{1}{y}=1+\frac{1}{2}=\frac{3}{2}$. Then cross multiplying, yields $2=9$.

10. Find the error in the following argument. Let $x=3$ and $y=1$. Then $1=3-2=3-(3-1)=x-y(x-y)=(x-y)(x-y)=2^2=4$.

11. Find the error in the following. $\frac{xy+y}{x}=y+y=2y$. Now let $x=2$ and $y=2$ to obtain $3=4$.

12. Show the rational numbers satisfy the field axioms. You may assume the associative, commutative, and distributive laws hold for the integers.

13. Show that for n a positive integer, $\sum_{k=0}^{n}(a+bk)=\sum_{k=0}^{n}(a+b(n-k))$. Explain why
$$2\sum_{k=0}^{n}(a+bk)=\sum_{k=0}^{n}2a+bn=(n+1)(2a+bn)$$
and so $\sum_{k=0}^{n}(a+bk)=(n+1)\frac{a+(a+bn)}{2}$.

2.3 Set Notation

A set is just a collection of things called elements. Often these are also referred to as points in calculus. For example $\{1,2,3,8\}$ would be a set consisting of the elements 1,2,3, and 8. To indicate that 3 is an element of $\{1,2,3,8\}$, it is customary to write $3\in\{1,2,3,8\}$. $9\notin\{1,2,3,8\}$ means 9 is not an element of $\{1,2,3,8\}$. Sometimes a rule specifies a set. For example you could specify a set as all integers larger than 2. This would be written as $S=\{x\in\mathbb{Z}:x>2\}$. This notation says: the set of all integers, x, such that $x>2$.

If A and B are sets with the property that every element of A is an element of B, then A is a subset of B. For example, $\{1,2,3,8\}$ is a subset of $\{1,2,3,4,5,8\}$, in symbols,

$\{1,2,3,8\} \subseteq \{1,2,3,4,5,8\}$. The same statement about the two sets may also be written as $\{1,2,3,4,5,8\} \supseteq \{1,2,3,8\}$.

The union of two sets is the set consisting of everything which is contained in at least one of the sets, A or B. As an example of the union of two sets, $\{1,2,3,8\} \cup \{3,4,7,8\} = \{1,2,3,4,7,8\}$ because these numbers are those which are in at least one of the two sets. In general

$$A \cup B \equiv \{x : x \in A \text{ or } x \in B\}.$$

Be sure you understand that something which is in both A and B is in the union. It is not an exclusive or.

The intersection of two sets, A and B consists of everything which is in both of the sets. Thus $\{1,2,3,8\} \cap \{3,4,7,8\} = \{3,8\}$ because 3 and 8 are those elements the two sets have in common. In general,

$$A \cap B \equiv \{x : x \in A \text{ and } x \in B\}.$$

When with real numbers, $[a,b]$ denotes the set of real numbers x, such that $a \leq x \leq b$ and $[a,b)$ denotes the set of real numbers such that $a \leq x < b$. (a,b) consists of the set of real numbers, x such that $a < x < b$ and $(a,b]$ indicates the set of numbers, x such that $a < x \leq b$. $[a,\infty)$ means the set of all numbers, x such that $x \geq a$ and $(-\infty,a]$ means the set of all real numbers which are less than or equal to a. These sorts of sets of real numbers are called intervals. The two points, a and b are called endpoints of the interval. Other intervals such as $(-\infty,b)$ are defined by analogy to what was just explained. In general, the curved parenthesis indicates the end point it sits next to is not included while the square parenthesis indicates this end point is included. The reason that there will always be a curved parenthesis next to ∞ or $-\infty$ is that these are not real numbers. Therefore, they cannot be included in any set of real numbers. It is assumed that the reader is already familiar with order which is discussed in the next section more carefully. The emphasis here is on the geometric significance of these intervals. That is $[a,b)$ consists of all points of the number line which are to the right of a possibly equaling a and to the left of b. In the above description, I have used the usual description of this set in terms of order.

A special set which needs to be given a name is the empty set also called the null set, denoted by $\emptyset$. Thus $\emptyset$ is defined as the set which has no elements in it. Mathematicians like to say the empty set is a subset of every set. The reason they say this is that if it were not so, there would have to exist a set A, such that $\emptyset$ has something in it which is not in A. However, $\emptyset$ has nothing in it and so the least intellectual discomfort is achieved by saying $\emptyset \subseteq A$.

If A and B are two sets, $A \setminus B$ denotes the set of things which are in A but not in B. Thus

$$A \setminus B \equiv \{x \in A : x \notin B\}.$$

Set notation is used whenever convenient.

2.4 Order

The real numbers also have an order defined on them. This order may be defined by reference to the positive real numbers, those to the right of 0 on the number line, denoted by $\mathbb{R}^+$ which is assumed to satisfy the following axioms.

Axiom 2.4.1 *The sum of two positive real numbers is positive.*

Axiom 2.4.2 *The product of two positive real numbers is positive.*

Axiom 2.4.3 *For a given real number x one and only one of the following alternatives holds. Either x is positive, $x = 0$, or $-x$ is positive.*

Definition 2.4.4 *$x < y$ exactly when $y + (-x) \equiv y - x \in \mathbb{R}^+$. In the usual way, $x < y$ is the same as $y > x$ and $x \leq y$ means either $x < y$ or $x = y$. The symbol $\geq$ is defined similarly.*

Theorem 2.4.5 *The following hold for the order defined as above.*

1. *If $x < y$ and $y < z$ then $x < z$ (Transitive law).*
2. *If $x < y$ then $x + z < y + z$ (addition to an inequality).*
3. *If $x \leq 0$ and $y \leq 0$, then $xy \geq 0$.*
4. *If $x > 0$ then $x^{-1} > 0$.*
5. *If $x < 0$ then $x^{-1} < 0$.*
6. *If $x < y$ then $xz < yz$ if $z > 0$, (multiplication of an inequality).*
7. *If $x < y$ and $z < 0$, then $xz > zy$ (multiplication of an inequality).*
8. *Each of the above holds with $>$ and $<$ replaced by $\geq$ and $\leq$ respectively except for 4 and 5 in which we must also stipulate that $x \neq 0$.*
9. *For any x and y, exactly one of the following must hold. Either $x = y$, $x < y$, or $x > y$ (trichotomy).*
10. *$xy > 0$ if and only if both x, y are positive or both $-x, -y$ are positive. Thus $xy = 0$ means x, y have the same sign.*

Proof: First consider 1, the transitive law. Suppose $x < y$ and $y < z$. Why is $x < z$? In other words, why is $z - x \in \mathbb{R}^+$? It is because $z - x = (z - y) + (y - x)$ and both $z - y, y - x \in \mathbb{R}^+$. Thus by 2.4.1 above, $z - x \in \mathbb{R}^+$ and so $z > x$.

Next consider 2, addition to an inequality. If $x < y$ why is $x + z < y + z$? it is because

$$\begin{aligned} (y+z)+-(x+z) &= (y+z)+(-1)(x+z) \\ &= y+(-1)x+z+(-1)z \\ &= y-x \in \mathbb{R}^+. \end{aligned}$$

Next consider 3. If $x \leq 0$ and $y \leq 0$, why is $xy \geq 0$? First note there is nothing to show if either x or y equal 0 so assume this is not the case. By 2.4.3 $-x > 0$ and $-y > 0$. Therefore, by 2.4.2 and what was proved about $-x = (-1)x$,

$$(-x)(-y) = (-1)^2 xy \in \mathbb{R}^+.$$

Is $(-1)^2 = 1$? If so the claim is proved. But $-(-1) = (-1)^2$ and $-(-1) = 1$ because $-1 + 1 = 0$.

Next consider 4. If $x > 0$ why is $x^{-1} > 0$? By 2.4.3 either $x^{-1} = 0$ or $-x^{-1} \in \mathbb{R}^+$. It can't happen that $x^{-1} = 0$ because then you would have to have $1 = 0x$ and as was shown earlier, $0x = 0$. Therefore, consider the possibility that $-x^{-1} \in \mathbb{R}^+$. This can't work either because then you would have

$$(-1)x^{-1}x = (-1)(1) = -1$$

and it would follow from 2.4.2 that $-1 \in \mathbb{R}^+$. But this is impossible because if $x \in \mathbb{R}^+$, then if $-1 \in \mathbb{R}$, $(-1)x = -x \in \mathbb{R}^+$ and contradicts 2.4.3 which states that either $-x$ or x is in $\mathbb{R}^+$ but not both.

Next consider 5. If $x < 0$, why is $x^{-1} < 0$? As before, $x^{-1} \neq 0$. If $x^{-1} > 0$, then as before,

$$-x\left(x^{-1}\right) = -1 \in \mathbb{R}^+$$

which was just shown not to occur.

Next consider 6. If $x < y$ why is $xz < yz$ if $z > 0$? This follows because

$$yz - xz = z(y - x) \in \mathbb{R}^+$$

since both z and $y - x \in \mathbb{R}^+$.

Next consider 7. If $x < y$ and $z < 0$, why is $xz > zy$? This follows because

$$zx - zy = z(x - y) \in \mathbb{R}^+$$

by what was proved in 3.

The next two claims are obvious and left for you.

Now suppose $xy > 0$. If $-x > 0$ and $y > 0$, then $-xy > 0$ contrary to $xy > 0$. It is similar if $x > 0$. Thus if $xy > 0$ either both x, y are positive or both $-x, -y$ are positive. In the second case, we say both x, y are negative. If both x, y are positive, then $xy > 0$ by the order axioms. If $-x, -y$ both positive, then $xy = (-1)^2 xy = (-x)(-y) > 0$. ■

Note that trichotomy could be stated by saying $x \leq y$ or $y \leq x$.

Definition 2.4.6 $|x| \equiv \begin{cases} x \text{ if } x \geq 0, \\ -x \text{ if } x < 0. \end{cases}$

Note that $|x|$ can be thought of as the distance between x and 0.

Theorem 2.4.7 $|xy| = |x|\,|y|$.

Proof: You can verify this by checking all available cases. Do so. ■

Theorem 2.4.8 *The following inequalities hold.*

$$|x + y| \leq |x| + |y|, \quad ||x| - |y|| \leq |x - y|.$$

Either of these inequalities may be called the triangle inequality.

Proof: First note that if $a, b \in \mathbb{R}^+ \cup \{0\}$ then $a \leq b$ if and only if $a^2 \leq b^2$. Here is why. Suppose $a \leq b$. Then by the properties of order proved above, $a^2 \leq ab \leq b^2$ because $b^2 - ab = b(b - a) \in \mathbb{R}^+ \cup \{0\}$. Next suppose $a^2 \leq b^2$. If both $a, b = 0$ there is nothing to show. Assume then they are not both 0. Then

$$b^2 - a^2 = (b + a)(b - a) \in \mathbb{R}^+.$$

By the above theorem on order, $(a+b)^{-1} \in \mathbb{R}^+$ and so using the associative law,

$$(a+b)^{-1}((b+a)(b-a)) = (b-a) \in \mathbb{R}^+$$

Now

$$\begin{aligned} |x+y|^2 &= (x+y)^2 = x^2 + 2xy + y^2 \\ &\leq |x|^2 + |y|^2 + 2|x||y| = (|x| + |y|)^2 \end{aligned}$$

and so the first of the inequalities follows. Note I used $xy \leq |xy| = |x||y|$ which follows from the definition.

To verify the other form of the triangle inequality, $x = x - y + y$ so $|x| \leq |x-y| + |y|$ and so $|x| - |y| \leq |x-y| = |y-x|$. Now repeat the argument replacing the roles of x and y to conclude $|y| - |x| \leq |y-x|$ .Therefore, $||y| - |x|| \leq |y-x|$. ■

Example 2.4.9 *Solve the inequality* $2x + 4 \leq x - 8$

Subtract $2x$ from both sides to yield $4 \leq -x - 8$. Next add 8 to both sides to get $12 \leq -x$. Then multiply both sides by (-1) to obtain $x \leq -12$. Alternatively, subtract x from both sides to get $x + 4 \leq -8$. Then subtract 4 from both sides to obtain $x \leq -12$.

Example 2.4.10 *Solve the inequality* $(x+1)(2x-3) \geq 0$.

If this is to hold, either both of the factors, $x+1$ and $2x-3$ are nonnegative or they are both non-positive. The first case yields $x + 1 \geq 0$ and $2x - 3 \geq 0$ so $x \geq -1$ and $x \geq \frac{3}{2}$ yielding $x \geq \frac{3}{2}$. The second case yields $x + 1 \leq 0$ and $2x - 3 \leq 0$ which implies $x \leq -1$ and $x \leq \frac{3}{2}$. Therefore, the solution to this inequality is $x \leq -1$ or $x \geq \frac{3}{2}$.

Example 2.4.11 *Solve the inequality* $(x)(x+2) \geq -4$

Here the problem is to find x such that $x^2 + 2x + 4 \geq 0$. However, $x^2 + 2x + 4 = (x+1)^2 + 3 \geq 0$ for all x. Therefore, the solution to this problem is all $x \in \mathbb{R}$.

Example 2.4.12 *Solve the inequality* $2x + 4 \leq x - 8$

This is written as $(-\infty, -12]$.

Example 2.4.13 *Solve the inequality* $(x+1)(2x-3) \geq 0$.

This was worked earlier and $x \leq -1$ or $x \geq \frac{3}{2}$ was the answer. In terms of set notation this is denoted by $(-\infty, -1] \cup [\frac{3}{2}, \infty)$.

Example 2.4.14 *Solve the equation* $|x-1| = 2$

This will be true when $x - 1 = 2$ or when $x - 1 = -2$. Therefore, there are two solutions to this problem, $x = 3$ or $x = -1$.

Example 2.4.15 *Solve the inequality* $|2x-1| < 2$

From the number line, it is necessary to have $2x - 1$ between -2 and 2 because the inequality says that the distance from $2x - 1$ to 0 is less than 2. Therefore, $-2 < 2x - 1 < 2$ and so $-1/2 < x < 3/2$. In other words, $-1/2 < x$ and $x < 3/2$.

Example 2.4.16 *Solve the inequality* $|2x-1|>2$.

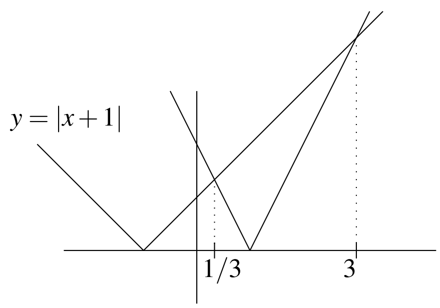
This happens if $2x-1>2$ or if $2x-1<-2$. Thus the solution is $x>3/2$ or $x<-1/2$. Written in terms of intervals this is $\left(\frac{3}{2},\infty\right)\cup\left(-\infty,-\frac{1}{2}\right)$.

Example 2.4.17 *Solve* $|x+1|=|2x-2|$

There are two ways this can happen. It could be the case that $x+1=2x-2$ in which case $x=3$ or alternatively, $x+1=2-2x$ in which case $x=1/3$.

Example 2.4.18 *Solve* $|x+1|\leq|2x-2|$

In order to keep track of what is happening, it is a very good idea to graph the two relations, $y=|x+1|$ and $y=|2x-2|$ on the same set of coordinate axes. This is not a hard job. $|x+1|=x+1$ when $x>-1$ and $|x+1|=-1-x$ when $x\leq-1$. Therefore, it is not hard to draw its graph. Similar considerations apply to the other relation. Functions and their graphs are discussed formally later but I assume the reader has seen these things. The result is



Equality holds exactly when $x=3$ or $x=\frac{1}{3}$ as in the preceding example. Consider x between $\frac{1}{3}$ and 3. You can see these values of x do not solve the inequality. For example $x=1$ does not work. Therefore, $\left(\frac{1}{3},3\right)$ must be excluded. The values of x larger than 3 do not produce equality so either $|x+1|<|2x-2|$ for these points or $|2x-2|<|x+1|$ for these points. Checking examples, you see the first of the two cases is the one which holds. Therefore, $[3,\infty)$ is included. Similar reasoning obtains $(-\infty,\frac{1}{3}]$. It follows the solution set to this inequality is $(-\infty,\frac{1}{3}]\cup[3,\infty)$.

Example 2.4.19 *Suppose* $\varepsilon>0$ *is a given positive number. Obtain a number,* $\delta>0$, *such that if* $|x-1|<\delta$, *then* $\left|x^2-1\right|<\varepsilon$.

First of all, note $\left|x^2-1\right|=|x-1||x+1|\leq(|x|+1)|x-1|$. Now if $|x-1|<1$, it follows $|x|<2$ and so for $|x-1|<1$, $\left|x^2-1\right|<3|x-1|$.Now let $\delta=\min\left(1,\frac{\varepsilon}{3}\right)$. This notation means to take the minimum of the two numbers, 1 and $\frac{\varepsilon}{3}$. Then if $|x-1|<\delta$, $\left|x^2-1\right|<3|x-1|<3\frac{\varepsilon}{3}=\varepsilon$.

2.5 Exercises

1. Solve $(3x+2)(x-3) \leq 0$.
2. Solve $(3x+2)(x-3) > 0$.
3. Solve $\frac{x+2}{3x-2} < 0$.
4. Solve$\frac{x+1}{x+3} < 1$.
5. Solve $(x-1)(2x+1) \leq 2$.
6. Solve $(x-1)(2x+1) > 2$.
7. Solve $x^2 - 2x \leq 0$.
8. Solve $(x+2)(x-2)^2 \leq 0$.
9. Solve $\frac{3x-4}{x^2+2x+2} \geq 0$.
10. Solve $\frac{3x+9}{x^2+2x+1} \geq 1$.
11. Solve $\frac{x^2+2x+1}{3x+7} < 1$.
12. Solve $|x+1| = |2x-3|$.
13. Solve $|3x+1| < 8$. Give your answer in terms of intervals on the real line.
14. Sketch on the number line the solution to the inequality $|x-3| > 2$.
15. Sketch on the number line the solution to the inequality $|x-3| < 2$.
16. Show $|x| = \sqrt{x^2}$.
17. Solve $|x+2| < |3x-3|$.
18. Tell when equality holds in the triangle inequality.
19. Solve $|x+2| \leq 8 + |2x-4|$.
20. Solve $(x+1)(2x-2)x \geq 0$.
21. Solve $\frac{x+3}{2x+1} > 1$.
22. Solve $\frac{x+2}{3x+1} > 2$.
23. Describe the set of numbers, a such that there is no solution to $|x+1| = 4 - |x+a|$.
24. Suppose $0 < a < b$. Show $a^{-1} > b^{-1}$.
25. Show that if $|x-6| < 1$, then $|x| < 7$.
26. Suppose $|x-8| < 2$. How large can $|x-5|$ be?
27. Obtain a number, $\delta > 0$, such that if $|x-1| < \delta$, then $\left|x^2 - 1\right| < 1/10$.
28. Obtain a number, $\delta > 0$, such that if $|x-4| < \delta$, then $|\sqrt{x} - 2| < 1/10$.
29. Suppose $\varepsilon > 0$ is a given positive number. Obtain a number, $\delta > 0$, such that if $|x-1| < \delta$, then $|\sqrt{x} - 1| < \varepsilon$. **Hint:** This δ will depend in some way on ε. You need to tell how.

2.6 The Binomial Theorem

Consider the following problem: You have the integers $S_n = \{1, 2, \cdots, n\}$ and k is an integer no larger than n. How many ways are there to fill k slots with these integers starting from left to right if whenever an integer from S_n has been used, it cannot be re used in any succeeding slot?

$$\overbrace{\_\_\_\_, \_\_\_\_, \_\_\_\_, \_\_\_\_, \cdots, \_\_\_\_}^{k \text{ of these slots}}$$

This number is known as permutations of n things taken k at a time and is denoted by $P(n,k)$. It is easy to figure it out. There are n choices for the first slot. For each choice for the fist slot, there remain $n-1$ choices for the second slot. Thus there are $n(n-1)$ ways to fill the first two slots. Now there remain $n-2$ ways to fill the third. Thus there are $n(n-1)(n-2)$ ways to fill the first three slots. Continuing this way, you see there are

$$P(n,k) = n(n-1)(n-2)\cdots(n-k+1)$$

ways to do this.

Now define for k a positive integer, $k! \equiv k(k-1)(k-2)\cdots 1$, $0! \equiv 1$. This is called k factorial. Thus $P(k,k) = k!$ and you should verify that $P(n,k) = \frac{n!}{(n-k)!}$. Now consider the number of ways of selecting a set of k different numbers from S_n. For each set of k numbers there are $P(k,k) = k!$ ways of listing these numbers in order. Therefore, denoting by $\left(\begin{array}{c} n \\ k \end{array}\right)$ the number of ways of selecting a set of k numbers from S_n, it must be the case that

$$\left(\begin{array}{c} n \\ k \end{array}\right) k! = P(n,k) = \frac{n!}{(n-k)!}$$

Therefore, $\left(\begin{array}{c} n \\ k \end{array}\right) = \frac{n!}{k!(n-k)!}$. How many ways are there to select no numbers from S_n? Obviously one way. Note the above formula gives the right answer in this case as well as in all other cases due to the definition which says $0! = 1$.

Now consider the problem of writing a formula for $(x+y)^n$ where n is a positive integer. Imagine writing it like this:

$$\overbrace{(x+y)(x+y)\cdots(x+y)}^{n \text{ times}}$$

Then you know the result will be sums of terms of the form $a_k x^k y^{n-k}$. What is a_k? In other words, how many ways can you pick x from k of the factors above and y from the other $n-k$. There are n factors so the number of ways to do it is $\left(\begin{array}{c} n \\ k \end{array}\right)$. Therefore, a_k is the above formula and so this proves the following important theorem known as the binomial theorem.

Theorem 2.6.1 *The following formula holds for any n a positive integer.*

$$(x+y)^n = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array}\right) x^k y^{n-k}.$$

2.7 Well Ordering and Archimedean Property

Definition 2.7.1 *A set is well ordered if every nonempty subset S, contains a smallest element z having the property that $z \leq x$ for all $x \in S$.*

Axiom 2.7.2 *Any set of integers larger than a given number is well ordered.*

In particular, the natural numbers defined as $\mathbb{N} \equiv \{1, 2, \cdots\}$ is well ordered.

The above axiom implies the principle of mathematical induction.

Theorem 2.7.3 *(Mathematical induction) A set $S \subseteq \mathbb{Z}$, having the property that $a \in S$ and $n+1 \in S$ whenever $n \in S$ contains all integers $x \in \mathbb{Z}$ such that $x \geq a$.*

Proof: Let $T \equiv ([a,\infty) \cap \mathbb{Z}) \setminus S$. Thus T consists of all integers larger than or equal to a which are not in S. The theorem will be proved if $T = \emptyset$. If $T \neq \emptyset$ then by the well ordering principle, there would have to exist a smallest element of T, denoted as b. It must be the case that $b > a$ since by definition, $a \notin T$. Then the integer, $b-1 \geq a$ and $b-1 \notin S$ because if $b-1 \in S$, then $b-1+1 = b \in S$ by the assumed property of S. Therefore, $b-1 \in ([a,\infty) \cap \mathbb{Z}) \setminus S = T$ which contradicts the choice of b as the smallest element of T.

($b-1$ is smaller.) Since a contradiction is obtained by assuming $T \neq \emptyset$, it must be the case that $T = \emptyset$ and this says that everything in $[a,\infty) \cap \mathbb{Z}$ is also in S. ■

Mathematical induction is a very useful device for proving theorems about the integers.

Example 2.7.4 *Prove by induction that* $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

By inspection, if $n = 1$ then the formula is true. The sum yields 1 and so does the formula on the right. Suppose this formula is valid for some $n \geq 1$ where n is an integer. Then

$$\sum_{k=1}^{n+1} k^2 = \sum_{k=1}^{n} k^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2.$$

The step going from the first to the second equality is based on the assumption that the formula is true for n. This is called the induction hypothesis. Now simplify the expression in the second line,

$$\frac{n(n+1)(2n+1)}{6} + (n+1)^2.$$

This equals $(n+1)\left(\frac{n(2n+1)}{6} + (n+1)\right)$ and

$$\frac{n(2n+1)}{6} + (n+1) = \frac{6(n+1) + 2n^2 + n}{6} = \frac{(n+2)(2n+3)}{6}$$

Therefore,

$$\sum_{k=1}^{n+1} k^2 = \frac{(n+1)(n+2)(2n+3)}{6} = \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6},$$

showing the formula holds for $n+1$ whenever it holds for n. This proves the formula by mathematical induction.

Example 2.7.5 *Show that for all* $n \in \mathbb{N}$, $\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} < \frac{1}{\sqrt{2n+1}}$.

If $n = 1$ this reduces to the statement that $\frac{1}{2} < \frac{1}{\sqrt{3}}$ which is obviously true. Suppose then that the inequality holds for n. Then

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{2n+1}} \frac{2n+1}{2n+2} = \frac{\sqrt{2n+1}}{2n+2}.$$

The theorem will be proved if this last expression is less than $\frac{1}{\sqrt{2n+3}}$. This happens if and only if

$$\left(\frac{1}{\sqrt{2n+3}}\right)^2 = \frac{1}{2n+3} > \frac{2n+1}{(2n+2)^2}$$

which occurs if and only if $(2n+2)^2 > (2n+3)(2n+1)$ and this is clearly true which may be seen from expanding both sides. This proves the inequality.

Lets review the process just used. If S is the set of integers at least as large as 1 for which the formula holds, the first step was to show $1 \in S$ and then that whenever $n \in S$, it follows $n+1 \in S$. Therefore, by the principle of mathematical induction, S contains $[1,\infty) \cap \mathbb{Z}$, all positive integers. In doing an inductive proof of this sort, the set, S is normally not mentioned. One just verifies the steps above. First show the thing is true for some $a \in \mathbb{Z}$ and then verify that whenever it is true for m it follows it is also true for $m+1$. When this has been done, the theorem has been proved for all $m \geq a$.

Definition 2.7.6 *The Archimedean property states that whenever $x \in \mathbb{R}$, and $a > 0$, there exists $n \in \mathbb{N}$ such that $na > x$.*

This is not hard to believe. Just look at the number line. Imagine the intervals

$$[0,a), [a,2a), [2a,3a), \cdots.$$

If $x < 0$, you could consider a and it would be larger than x. If $x \geq 0$, surely, it is reasonable to suppose that x would be on one of these intervals, say $[pa, (p+1)a)$. This Archimedean property is quite important because it shows every fixed real number is smaller than some integer. It also can be used to verify a very important property of the rational numbers.

Axiom 2.7.7 *$\mathbb{R}$ has the Archimedean property.*

Theorem 2.7.8 *Suppose $x < y$ and $y - x > 1$. Then there exists an integer, $l \in \mathbb{Z}$, such that $x < l < y$. If x is an integer, there is no integer y satisfying $x < y < x+1$.*

Proof: Let x be the smallest positive integer. Not surprisingly, $x = 1$ but this can be proved. If $x < 1$ then $x^2 < x$ contradicting the assertion that x is the smallest natural number. Therefore, 1 is the smallest natural number. This shows there is no integer y, satisfying $x < y < x+1$ since otherwise, you could subtract x and conclude $0 < y - x < 1$ for some integer $y - x$.

Now suppose $y - x > 1$ and let $S \equiv \{w \in \mathbb{N} : w \geq y\}$. The set S is nonempty by the Archimedean property. Let k be the smallest element of S. Therefore, $k - 1 < y$. Either $k - 1 \leq x$ or $k - 1 > x$. If $k - 1 \leq x$, then

$$y - x \leq y - (k-1) = \overbrace{y - k}^{\leq 0} + 1 \leq 1$$

contrary to the assumption that $y - x > 1$. Therefore, $x < k - 1 < y$ and this proves the theorem with $l = k - 1$. ■

It is the next theorem which gives the density of the rational numbers. This means that for any real number, there exists a rational number arbitrarily close to it.

Theorem 2.7.9 *If $x < y$ then there exists a rational number r such that $x < r < y$.*

Proof: Let $n \in \mathbb{N}$ be large enough that $n(y - x) > 1$. Thus $(y - x)$ added to itself n times is larger than 1. Therefore,

$$n(y - x) = ny + n(-x) = ny - nx > 1.$$

It follows from Theorem 2.7.8 there exists $m \in \mathbb{Z}$ such that $nx < m < ny$ and so take $r = m/n$. ■

Definition 2.7.10 *A set $S \subseteq \mathbb{R}$ is dense in $\mathbb{R}$ if whenever $a < b$, $S \cap (a,b) \neq \emptyset$.*

Thus the above theorem says $\mathbb{Q}$ is "dense" in $\mathbb{R}$.

You probably saw the process of division in elementary school. Even though you saw it at a young age it is very profound and quite difficult to understand. Suppose you want to do the following problem $\frac{79}{22}$. What did you do? You likely did a process of long division which gave the following result. $\frac{79}{22} = 3$ with remainder 13. This meant $79 = 3(22) + 13$.You were given two numbers, 79 and 22 and you wrote the first as some multiple of the second added to a third number which was smaller than the second number. Can this always be done? The answer is in the next theorem and depends here on the Archimedean property of the real numbers.

Theorem 2.7.11 *Suppose* $0 < a$ *and let* $b \geq 0$. *Then there exists a unique integer* p *and real number* r *such that* $0 \leq r < a$ *and* $b = pa + r$.

Proof: Let $S \equiv \{n \in \mathbb{N} : an > b\}$. By the Archimedean property this set is nonempty. Let $p+1$ be the smallest element of S. Then $pa \leq b$ because $p+1$ is the smallest in S. Therefore, $r \equiv b - pa \geq 0$. If $r \geq a$ then $b - pa \geq a$ and so $b \geq (p+1)a$ contradicting $p+1 \in S$. Therefore, $r < a$ as desired.

To verify uniqueness of p and r, suppose p_i and r_i, $i = 1,2$, both work and $r_2 > r_1$. Then a little algebra shows $p_1 - p_2 = \frac{r_2 - r_1}{a} \in (0,1)$. Thus $p_1 - p_2$ is an integer between 0 and 1, contradicting Theorem 2.7.8. The case that $r_1 > r_2$ cannot occur either by similar reasoning. Thus $r_1 = r_2$ and it follows that $p_1 = p_2$. ■

This theorem is called the Euclidean algorithm when a and b are integers. In this case, you would have r is an integer because it equals an integer.

2.8 Arithmetic of Integers

Here we consider some very important algebraic notions including the Euclidean algorithm just mentioned and issues related to whether two numbers are relatively prime, prime numbers and so forth. The following definition describes what is meant by a prime number and also what is meant by the word "divides".

Definition 2.8.1 *The number* a *divides the number* b *if, in Theorem 2.7.11,* $r = 0$. *That is, there is zero remainder. The notation for this is* $a|b$, *read* a *divides* b *and* a *is called a factor of* b. *A prime number is one which has the property that the only numbers which divide it are itself and 1 and it is at least 2. The greatest common divisor of two positive integers* m,n *is that number* p *which has the property that* p *divides both* m *and* n *and also if* q *divides both* m *and* n, *then* q *divides* p. *Two integers are relatively prime if their greatest common divisor is one. The greatest common divisor of* m *and* n *is denoted as* (m,n).

There is a phenomenal and amazing theorem which relates the greatest common divisor to the smallest number in a certain set. Suppose m,n are two positive integers. Then if x,y are integers, so is $xm + yn$. Consider all integers which are of this form. Some are positive such as $1m + 1n$ and some are not. The set S in the following theorem consists of exactly those integers of this form which are positive. Then the greatest common divisor of m and n will be the smallest number in S. This is what the following theorem says.

Theorem 2.8.2 *Let* m,n *be two positive integers and define*

$$S \equiv \{xm + yn \in \mathbb{N} : x,y \in \mathbb{Z}\}.$$

Then the smallest number in S *is the greatest common divisor, denoted by* (m,n).

Proof: First note that both m and n are in S so it is a nonempty set of positive integers. By well ordering, there is a smallest element of S, called $p = x_0 m + y_0 n$. Either p divides m or it does not. If p does not divide m, then by Theorem 2.7.11, $m = pq + r$ where $0 < r < p$. Thus $m = (x_0 m + y_0 n)q + r$ and so, solving for r,

$$r = m(1 - x_0) + (-y_0 q)n \in S.$$

However, this is a contradiction because p was the smallest element of S. Thus $p|m$. Similarly $p|n$.

Now suppose q divides both m and n. Then $m = qx$ and $n = qy$ for integers, x and y. Therefore,

$$p = mx_0 + ny_0 = x_0 qx + y_0 qy = q(x_0 x + y_0 y)$$

showing $q|p$. Therefore, $p = (m,n)$. ■

This amazing theorem will now be used to prove a fundamental property of prime numbers which leads to the fundamental theorem of arithmetic, the major theorem which says every integer can be factored as a product of primes.

Theorem 2.8.3 *If p is a prime and $p|ab$ then either $p|a$ or $p|b$.*

Proof: Suppose p does not divide a. Then since p is prime, the only factors of p are 1 and p so follows $(p,a) = 1$ and therefore, there exists integers, x and y such that $1 = ax + yp$. Multiplying this equation by b yields $b = abx + ybp$. Since $p|ab$, $ab = pz$ for some integer z. Therefore, $b = abx + ybp = pzx + ybp = p(xz + yb)$ and this shows p divides b. ■

Theorem 2.8.4 *(Fundamental theorem of arithmetic) Let $a \in \mathbb{N} \setminus \{1\}$. Then $a = \prod_{i=1}^{n} p_i$ where p_i are all prime numbers. Furthermore, this prime factorization is unique except for the order of the factors.*

Proof: If a equals a prime number, the prime factorization clearly exists. In particular the prime factorization exists for the prime number 2. Assume this theorem is true for all $a \leq n-1$. If n is a prime, then it has a prime factorization. On the other hand, if n is not a prime, then there exist two integers k and m such that $n = km$ where each of k and m are less than n. Therefore, each of these is no larger than $n-1$ and consequently, each has a prime factorization. Thus so does n. It remains to argue the prime factorization is unique except for order of the factors.

Suppose $\prod_{i=1}^{n} p_i = \prod_{j=1}^{m} q_j$ where the p_i and q_j are all prime, there is no way to reorder the q_k such that $m = n$ and $p_i = q_i$ for all i, and $n + m$ is the smallest positive integer such that this happens. Then by Theorem 2.8.3, $p_1 | q_j$ for some j. Since these are prime numbers this requires $p_1 = q_j$. Reordering if necessary it can be assumed that $q_j = q_1$. Then dividing both sides by $p_1 = q_1, \prod_{i=1}^{n-1} p_{i+1} = \prod_{j=1}^{m-1} q_{j+1}$. Since $n + m$ was as small as possible for the theorem to fail, it follows that $n - 1 = m - 1$ and the prime numbers, $q_2, \cdots, q_m$ can be reordered in such a way that $p_k = q_k$ for all $k = 2, \cdots, n$. Hence $p_i = q_i$ for all i because it was already argued that $p_1 = q_1$, and this results in a contradiction. ■

2.9 Exercises

1. By Theorem 2.7.9 it follows that for $a < b$, there exists a rational number between a and b. Show there exists an integer k such that $a < \frac{k}{2^m} < b$ for some k, m integers.

2. Show there is no smallest number in $(0,1)$. Recall $(0,1)$ means the real numbers which are strictly larger than 0 and smaller than 1.

3. Show there is no smallest number in $\mathbb{Q} \cap (0,1)$.

4. Show that if $S \subseteq \mathbb{R}$ and S is well ordered with respect to the usual order on $\mathbb{R}$ then S cannot be dense in $\mathbb{R}$.

5. Prove by induction that $\sum_{k=1}^{n} k^3 = \frac{1}{4}n^4 + \frac{1}{2}n^3 + \frac{1}{4}n^2$.

6. It is a fine thing to be able to prove a theorem by induction but it is even better to be able to come up with a theorem to prove in the first place. Derive a formula for $\sum_{k=1}^{n} k^4$ in the following way. Look for a formula in the form $An^5 + Bn^4 + Cn^3 + Dn^2 + En + F$. Then try to find the constants A, B, C, D, E, and F such that things work out right. In doing this, show

$$(n+1)^4 =$$

$$\left(A(n+1)^5 + B(n+1)^4 + C(n+1)^3 + D(n+1)^2 + E(n+1) + F\right)$$

$$-An^5 + Bn^4 + Cn^3 + Dn^2 + En + F$$

and so some progress can be made by matching the coefficients. When you get your answer, prove it is valid by induction.

7. Prove by induction that whenever $n \geq 2, \sum_{k=1}^{n} \frac{1}{\sqrt{k}} > \sqrt{n}$.

8. If $r \neq 0$, show by induction that $\sum_{k=1}^{n} ar^k = a\frac{r^{n+1}}{r-1} - a\frac{r}{r-1}$.

9. Prove by induction that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

10. Let a and d be real numbers. Find a formula for $\sum_{k=1}^{n} (a + kd)$ and then prove your result by induction.

11. Consider the geometric series, $\sum_{k=1}^{n} ar^{k-1}$. Prove by induction that if $r \neq 1$, then $\sum_{k=1}^{n} ar^{k-1} = \frac{a - ar^n}{1-r}$.

12. This problem is a continuation of Problem 11. You put money in the bank and it accrues interest at the rate of r per payment period. These terms need a little explanation. If the payment period is one month, and you started with \$100 then the amount at the end of one month would equal $100(1+r) = 100 + 100r$. In this the second term is the interest and the first is called the principal. Now you have $100(1+r)$ in the bank. How much will you have at the end of the second month? By analogy to what was just done it would equal

$$100(1+r) + 100(1+r)r = 100(1+r)^2.$$

The amount you would have at the end of n months would be $100(1+r)^n$. (When a bank says they offer 6% compounded monthly, this means r, the rate per payment period equals $.06/12$.) In general, suppose you start with P and it sits in the bank for n payment periods. Then at the end of the n^{th} payment period, you would have $P(1+r)^n$ in the bank. In an ordinary annuity, you make payments, P at the end of each payment period, the first payment at the end of the first payment period. Thus there are n payments in all. Each accrue interest at the rate of r per payment period. Using Problem 11, find a formula for the amount you will have in the bank at the end of n payment periods? This is called the future value of an ordinary annuity. **Hint:** The first payment sits in the bank for $n-1$ payment periods and so this payment becomes $P(1+r)^{n-1}$. The second sits in the bank for $n-2$ payment periods so it grows to $P(1+r)^{n-2}$, etc.

13. Now suppose you want to buy a house by making n equal monthly payments. Typically, n is pretty large, 360 for a thirty year loan. Clearly a payment made 10 years from now can't be considered as valuable to the bank as one made today. This is because the one made today could be invested by the bank and having accrued interest for 10 years would be far larger. So what is a payment made at the end of k payment periods worth today assuming money is worth r per payment period? Shouldn't it be the amount, Q which when invested at a rate of r per payment period would yield P at the end of k payment periods? Thus from Problem 12 $Q(1+r)^k = P$ and so $Q = P(1+r)^{-k}$. Thus this payment of P at the end of n payment periods, is worth $P(1+r)^{-k}$ to the bank right now. It follows the amount of the loan should equal the sum of these "discounted payments". That is, letting A be the amount of the loan, $A = \sum_{k=1}^{n} P(1+r)^{-k}$. Using Problem 11, find a formula for the right side of the above formula. This is called the present value of an ordinary annuity.

14. Suppose the available interest rate is 7% per year and you want to take a loan for \$100,000 with the first monthly payment at the end of the first month. If you want to pay off the loan in 20 years, what should the monthly payments be? **Hint:** The rate per payment period is $.07/12$. See the formula you got in Problem 13 and solve for P.

15. Consider the first five rows of Pascal's[1] triangle

$$\begin{array}{c} 1 \\ 1\ 1 \\ 1\ 2\ 1 \\ 1\ 3\ 3\ 1 \\ 1\ 4\ 6\ 4\ 1 \end{array}$$

What is the sixth row? Now consider that $(x+y)^1 = 1x + 1y$, $(x+y)^2 = x^2 + 2xy + y^2$, and $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$. Give a conjecture about that $(x+y)^5$.

16. Based on Problem 15 conjecture a formula for $(x+y)^n$ and prove your conjecture by induction. **Hint:** Letting the numbers of the n^{th} row of Pascal's triangle be denoted by $\binom{n}{0}, \binom{n}{1}, \cdots, \binom{n}{n}$ in reading from left to right, there is a relation between the numbers on the $(n+1)^{st}$ row and those on the n^{th} row, the relation being $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$. This is used in the inductive step.

17. Let $\binom{n}{k} \equiv \frac{n!}{(n-k)!k!}$ where $0! \equiv 1$ and $(n+1)! \equiv (n+1)n!$ for all $n \geq 0$. Prove that whenever $k \geq 1$ and $k \leq n$, then $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$. Are these numbers, $\binom{n}{k}$ the same as those obtained in Pascal's triangle? Prove your assertion.

18. The binomial theorem states $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$. Prove the binomial theorem by induction. **Hint:** You might try using the preceding problem.

19. Show that for $p \in (0,1)$, $\sum_{k=0}^{n} \binom{n}{k} k p^k (1-p)^{n-k} = np$.

[1] Blaise Pascal lived in the 1600's and is responsible for the beginnings of the study of probability.

20. Show that for all $n \in \mathbb{N}$, $\left(1+\frac{1}{n}\right)^n \leq \left(1+\frac{1}{n+1}\right)^{n+1}$. **Hint:** Show first that $\binom{n}{k} = \frac{n\cdot(n-1)\cdots(n-k+1)}{k!}$. By the binomial theorem,
$$\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n}\binom{n}{k}\left(\frac{1}{n}\right)^k = \sum_{k=0}^{n}\frac{\overbrace{n\cdot(n-1)\cdots(n-k+1)}^{k\text{ factors}}}{k!n^k}.$$
Now consider the term $\frac{n\cdot(n-1)\cdots(n-k+1)}{k!n^k}$ and note that a similar term occurs in the binomial expansion for $\left(1+\frac{1}{n+1}\right)^{n+1}$ except that n is replaced with $n+1$ wherever this occurs. Argue the term got bigger and then note that in the binomial expansion for $\left(1+\frac{1}{n+1}\right)^{n+1}$, there are more terms.

21. Prove by induction that for all $k \geq 4$, $2^k \leq k!$

22. Use the Problems 21 and 20 to verify for all $n \in \mathbb{N}$, $\left(1+\frac{1}{n}\right)^n \leq 3$.

23. Prove by induction that $1+\sum_{i=1}^{n} i\,(i!) = (n+1)!$.

24. I can jump off the top of the Empire State Building without suffering any ill effects. Here is the proof by induction. If I jump from a height of one inch, I am unharmed. Furthermore, if I am unharmed from jumping from a height of n inches, then jumping from a height of $n+1$ inches will also not harm me. This is self evident and provides the induction step. Therefore, I can jump from a height of n inches for any n. What is the matter with this reasoning?

25. All horses are the same color. Here is the proof by induction. A single horse is the same color as himself. Now suppose the theorem that all horses are the same color is true for n horses and consider $n+1$ horses. Remove one of the horses and use the induction hypothesis to conclude the remaining n horses are all the same color. Put the horse which was removed back in and take out another horse. The remaining n horses are the same color by the induction hypothesis. Therefore, all $n+1$ horses are the same color as the $n-1$ horses which didn't get moved. This proves the theorem. Is there something wrong with this argument?

26. Let $\begin{pmatrix} n \\ k_1, k_2, k_3 \end{pmatrix}$ denote the number of ways of selecting a set of k_1 things, a set of k_2 things, and a set of k_3 things from a set of n things such that $\sum_{i=1}^{3} k_i = n$. Find a formula for $\begin{pmatrix} n \\ k_1, k_2, k_3 \end{pmatrix}$. Now give a formula for a trinomial theorem, one which expands $(x+y+z)^n$. Could you continue this way and get a multinomial formula?

2.10 Completeness of $\mathbb{R}$

By Theorem 2.7.9, between any two real numbers, points on the number line, there exists a rational number. This suggests there are a lot of rational numbers, but it is not clear from this Theorem whether the entire real line consists of only rational numbers. Some people might wish this were the case because then each real number could be described, not just as a point on a line but also algebraically, as the quotient of integers. Before 500 B.C., a group of mathematicians, led by Pythagoras believed in this, but they discovered their beliefs were false. It happened roughly like this. They knew they could construct the square root of two as the diagonal of a right triangle in which the two sides have unit length; thus they could regard $\sqrt{2}$ as a number. Unfortunately, they were also able to show $\sqrt{2}$ could not be written as the quotient of two integers. This discovery that the rational numbers could not even account for the results of geometric constructions was very upsetting to the Pythagoreans, especially when it became clear there were an endless supply of such "irrational" numbers.

This shows that if it is desired to consider all points on the number line, it is necessary to abandon the attempt to describe arbitrary real numbers in a purely algebraic manner using only the integers. Some might desire to throw out all the irrational numbers, and considering only the rational numbers, confine their attention to algebra, but this is not the approach to be followed here because it will effectively eliminate every major theorem of calculus and analysis. In this book real numbers will continue to be the points on the number line, a line which has no holes. This lack of holes is more precisely described in the following way.

Definition 2.10.1 *A non empty set, $S \subseteq \mathbb{R}$ is bounded above (below) if there exists $x \in \mathbb{R}$ such that $x \geq (\leq) s$ for all $s \in S$. If S is a nonempty set in $\mathbb{R}$ which is bounded above, then a number, l which has the property that l is an upper bound and that every other upper bound is no smaller than l is called a least upper bound, $l.u.b.(S)$ or often $\sup(S)$. If S is a nonempty set bounded below, define the greatest lower bound, $g.l.b.(S)$ or $\inf(S)$ similarly. Thus g is the $g.l.b.(S)$ means g is a lower bound for S and it is the largest of all lower bounds. If S is a nonempty subset of $\mathbb{R}$ which is not bounded above, this information is expressed by saying $\sup(S) = +\infty$ and if S is not bounded below, $\inf(S) = -\infty$.*

In an appendix, there is a proof that the real numbers can be obtained as equivalence classes of Cauchy sequences of rational numbers but in this book, we follow the historical development of the subject and accept it as an axiom. In other words, we will believe in the real numbers and this axiom.

The completeness axiom was identified by Bolzano as the reason for the truth of the intermediate value theorem for continuous functions around 1818. However, every existence theorem in calculus depends on some form of the completeness axiom.

Axiom 2.10.2 *(completeness) Every nonempty set of real numbers bounded above has a least upper bound and every nonempty set of real numbers which is bounded below has a greatest lower bound.*

It is this axiom which distinguishes Calculus from Algebra. A fundamental result about sup and inf is the following.

Proposition 2.10.3 *Let S be a nonempty set and suppose $\sup(S)$ exists. Then for every $\delta > 0$,*

$$S \cap (\sup(S) - \delta, \sup(S)] \neq \emptyset.$$

If $\inf(S)$ exists, then for every $\delta > 0$,

$$S \cap [\inf(S), \inf(S) + \delta) \neq \emptyset.$$

Proof: Consider the first claim. If the indicated set equals $\emptyset$, then $\sup(S) - \delta$ is an upper bound for S which is smaller than $\sup(S)$, contrary to the definition of $\sup(S)$ as the least upper bound. In the second claim, if the indicated set equals $\emptyset$, then $\inf(S)+\delta$ would be a lower bound which is larger than $\inf(S)$ contrary to the definition of $\inf(S)$.■

The wonderful thing about sup is that you can switch the order in which they occur. The same thing holds for inf. It is also convenient to generalize the notion of sup and inf so that we don't have to worry about whether it is a real number.

Definition 2.10.4 *Let $f(a,b) \in [-\infty,\infty]$ for $a \in A$ and $b \in B$ where A,B are nonempty sets which means that $f(a,b)$ is either a number, ∞, or $-\infty$. The symbol, $+\infty$ is interpreted as a point out at the end of the number line which is larger than every real number. Of course there is no such number. That is why it is called ∞. The symbol, $-\infty$ is interpreted similarly. Then $\sup_{a\in A} f(a,b)$ means $\sup(S_b)$ where $S_b \equiv \{f(a,b) : a \in A\}$. A similar convention holds for* inf.

Unlike limits, you can take the sup in different orders, same for inf.

Lemma 2.10.5 *Let $f(a,b) \in [-\infty,\infty]$ for $a \in A$ and $b \in B$ where A,B are sets. Then*

$$\sup_{a\in A}\sup_{b\in B} f(a,b) = \sup_{b\in B}\sup_{a\in A} f(a,b).$$

Also, you can replace sup *with* inf.

Proof: Note that for all a,b, $f(a,b) \leq \sup_{b\in B}\sup_{a\in A} f(a,b)$ and therefore, for all a, $\sup_{b\in B} f(a,b) \leq \sup_{b\in B}\sup_{a\in A} f(a,b)$. Therefore, it follows from the definition of sup that $\sup_{a\in A}\sup_{b\in B} f(a,b) \leq \sup_{b\in B}\sup_{a\in A} f(a,b)$. Repeat the same argument interchanging a and b, to get the conclusion of the lemma. The case of inf is similar. ■

2.11 Existence of Roots

What is $\sqrt[5]{7}$ and does it even exist? You can ask for it on your calculator and the calculator will give you a number which multiplied by itself 5 times will yield a number which is close to 7 but it isn't exactly right. Why should there exist a number which works exactly? Every one you find, appears to be some sort of approximation at best. If you can't produce one, why should you believe it is even there? The following is an argument that roots exist. You fill in the details of the argument. Basically, roots exist in analysis because of completeness of the real line. Here is a lemma.

Lemma 2.11.1 *Suppose $n \in \mathbb{N}$ and $a > 0$. Then if $x^n - a \neq 0$, there exists $\delta > 0$ such that whenever $y \in (x-\delta, x+\delta)$, it follows $y^n - a \neq 0$ and has the same sign as $x^n - a$.*

Proof: Let $y - x = \varepsilon$. Then we need to show that if $|\varepsilon|$ is small enough,

$$(x^n - a)((x+\varepsilon)^n - a) > 0.$$

From the binomial theorem

$$\begin{aligned}
(x^n - a)((x+\varepsilon)^n - a) &= (x^n - a)\left(\sum_{k=0}^{n}\binom{n}{k}\varepsilon^{n-k}x^k - a\right) \\
&= (x^n - a)\left((x^n - a) + \sum_{k=0}^{n-1}\binom{n}{k}\varepsilon^{n-k}x^k\right) \\
&\geq (x^n - a)^2 - |x^n - a|\,|\varepsilon| \sum_{k=0}^{n-1}\binom{n}{k}|\varepsilon|^{n-(k+1)}|x|^k
\end{aligned}$$

So let $|\varepsilon| < \min\left(1, \left(\frac{|x^n - a|}{2} \sum_{k=0}^{n-1} \binom{n}{k} |x|^k\right)^{-1}\right)$. Then for $|\varepsilon|$ this small,

$$(x^n - a)((x+\varepsilon)^n - a) = (x^n - a)(y^n - a) > \frac{(x^n - a)^2}{2} > 0 \blacksquare$$

Theorem 2.11.2 *Let $a > 0$ and let $n > 1$. Then there exists a unique $x > 0$ such that $x^n = a$.*

Proof: Let S denote those numbers $y \geq 0$ such that $t^n - a < 0$ for all $t \in [0, y]$. Now note that from the binomial theorem,

$$(1+a)^n - a = \sum_{k=0}^{n} \binom{n}{k} a^k 1^{n-k} - a \geq 1 + a - a = 1 > 0$$

Thus S is bounded above and $0 \in S$. Let $x \equiv \sup(S)$. Then by definition of sup, for every $\delta > 0$, there exists $t \in S$ with $|x - t| < \delta$.

If $x^n - a > 0$, then by the above lemma, for $t \in S$ sufficiently close to x,

$$(t^n - a)(x^n - a) > 0$$

which is a contradiction because the first factor is negative and the second is positive. Hence $x^n - a \leq 0$. If $x^n - a < 0$, then from the above lemma, there is a $\delta > 0$ such that if $t \in (x - \delta, x + \delta)$, $x^n - a$ and $t^n - a$ have the same sign. This is also a contradiction because then $x \neq \sup(S)$. It follows $x^n = a$. $\blacksquare$

From now on, we will use this fact that n^{th} roots exist whenever it is convenient to do so.

2.12 Exercises

1. Let $S = [2,5]$. Find $\sup S$. Now let $S = [2,5)$. Find $\sup S$. Is $\sup S$ always a number in S? Give conditions under which $\sup S \in S$ and then give conditions under which $\inf S \in S$.

2. Show that if $S \neq \emptyset$ and is bounded above (below) then $\sup S$ ($\inf S$) is unique. That is, there is only one least upper bound and only one greatest lower bound. If $S = \emptyset$ can you conclude that 7 is an upper bound? Can you conclude 7 is a lower bound? What about 13.5? What about any other number?

3. Let S be a set which is bounded above and let $-S$ denote the set $\{-x : x \in S\}$. How are $\inf(-S)$ and $\sup(S)$ related? **Hint:** Draw some pictures on a number line. What about $\sup(-S)$ and $\inf S$ where S is a set which is bounded below?

4. Which of the field axioms is being abused in the following argument that $0 = 2$? Let $x = y = 1$. Then
$$0 = x^2 - y^2 = (x - y)(x + y)$$
and so $0 = (x - y)(x + y)$. Now divide both sides by $x - y$ to obtain $0 = x + y = 1 + 1 = 2$.

5. Give conditions under which equality holds in the triangle inequality.

6. Let $k \leq n$ where k and n are natural numbers. $P(n,k)$, permutations of n things taken k at a time, is defined to be the number of different ways to form an ordered list of k of the numbers, $\{1,2,\cdots,n\}$. Show

$$P(n,k) = n \cdot (n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}.$$

7. Using the preceding problem, show the number of ways of selecting a set of k things from a set of n things is $\binom{n}{k}$.

8. Prove the binomial theorem from Problem 7. **Hint:** When you take $(x+y)^n$, note that the result will be a sum of terms of the form, $a_k x^{n-k} y^k$ and you need to determine what a_k should be. Imagine writing $(x+y)^n = (x+y)(x+y)\cdots(x+y)$ where there are n factors in the product. Now consider what happens when you multiply. Each factor contributes either an x or a y to a typical term.

9. Prove by induction that $n < 2^n$ for all natural numbers, $n \geq 1$.

10. Prove by the binomial theorem and Problem 7 that the number of subsets of a given finite set containing n elements is 2^n.

11. Let n be a natural number and let $k_1 + k_2 + \cdots k_r = n$ where k_i is a non negative integer. The symbol

$$\binom{n}{k_1 k_2 \cdots k_r}$$

denotes the number of ways of selecting r subsets of $\{1,\cdots,n\}$, which subsets contain $k_1, k_2 \cdots k_r$ elements in them. Find a formula for this number.

12. Is it ever the case that $(a+b)^n = a^n + b^n$ for a and b positive real numbers?

13. Is it ever the case that $\sqrt{a^2+b^2} = a+b$ for a and b positive real numbers?

14. Is it ever the case that $\frac{1}{x+y} = \frac{1}{x} + \frac{1}{y}$ for x and y positive real numbers?

15. Derive a formula for the multinomial expansion, $\left(\sum_{k=1}^{p} a_k\right)^n$ which is analogous to the binomial expansion. **Hint:** See Problem 8.

16. Suppose $a > 0$ and that x is a real number which satisfies the quadratic equation,

$$ax^2 + bx + c = 0.$$

Find a formula for x in terms of a and b and square roots of expressions involving these numbers. **Hint:** First divide by a to get $x^2 + \frac{b}{a}x + \frac{c}{a} = 0$. Then add and subtract the quantity $b^2/4a^2$. Verify that

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = \left(x + \frac{b}{2a}\right)^2.$$

Now solve the result for x. The process by which this was accomplished in adding in the term $b^2/4a^2$ is referred to as completing the square. You should obtain the quadratic formula,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

The expression b^2-4ac is called the discriminant. When it is positive there are two different real roots. When it is zero, there is exactly one real root and when it equals a negative number there are no real roots.

17. Find u such that $-\frac{b}{2}+u$ and $-\frac{b}{2}-u$ are roots of $x^2+bx+c=0$. Obtain the quadratic formula from this. [2]

18. Suppose $f(x)=3x^2+7x-17$. Find the value of x at which $f(x)$ is smallest by completing the square. Also determine $f(\mathbb{R})$ and sketch the graph of f. **Hint:**
$$\begin{aligned} f(x) &= 3\left(x^2+\frac{7}{3}x-\frac{17}{3}\right)=3\left(x^2+\frac{7}{3}x+\frac{49}{36}-\frac{49}{36}-\frac{17}{3}\right) \\ &= 3\left(\left(x+\frac{7}{6}\right)^2-\frac{49}{36}-\frac{17}{3}\right). \end{aligned}$$

19. Suppose $f(x)=-5x^2+8x-7$. Find $f(\mathbb{R})$. In particular, find the largest value of $f(x)$ and the value of x at which it occurs. Can you conjecture and prove a result about $y=ax^2+bx+c$ in terms of the sign of a based on these last two problems?

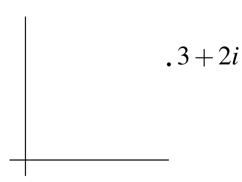
20. Show that if it is assumed $\mathbb{R}$ is complete, then the Archimedean property can be proved. **Hint:** Suppose completeness and let $a>0$. If there exists $x\in\mathbb{R}$ such that $na\leq x$ for all $n\in\mathbb{N}$, then x/a is an upper bound for $\mathbb{N}$. Let l be the least upper bound and argue there exists $n\in\mathbb{N}\cap[l-1/4,l]$. Now what about $n+1$?

21. Suppose you numbers a_k for each k a positive integer and that $a_1\leq a_2\leq\cdots$. Let A be the set of these numbers just described. Also suppose there exists an upper bound L such that each $a_k\leq L$. Then there exists N such that if $n\geq N$, then $(\sup A-\varepsilon< a_n\leq\sup A]$.

22. If $A\subseteq B$ for $A\neq\emptyset$ and A,B are sets of real numbers, show that $\inf(A)\geq\inf(B)$ and $\sup(A)\leq\sup(B)$.

2.13 The Complex Numbers

Just as a real number should be considered as a point on the line, a complex number is considered a point in the plane which can be identified 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, such a point is written as $a+ib$. For example, in the following picture, I have graphed the point $3+2i$. You see it corresponds to the point in the plane whose coordinates are $(3,2)$.



[2]The ancient Babylonians knew how to solve these quadratic equations sometime before 1700 B.C.

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

$$(a+ib)(c+id) = ac+iad+ibc+i^2bd = (ac-bd)+i(bc+ad).$$

Every non zero complex number, $a+ib$, with $a^2+b^2 \neq 0$, has a unique multiplicative inverse.

$$\frac{1}{a+ib} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}.$$

You should prove the following theorem.

Theorem 2.13.1 *The complex numbers with multiplication and addition defined as above form a field satisfying all the field axioms listed on Page 9.*

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.

$$\overline{a+ib} \equiv a-ib.$$

What it does is reflect a given complex number across the x axis. Algebraically, the following formula is easy to obtain.

$$\left(\overline{a+ib}\right)(a+ib) = a^2+b^2.$$

Definition 2.13.2 *Define the absolute value of a complex number as follows.*

$$|a+ib| \equiv \sqrt{a^2+b^2}.$$

Thus, denoting by z the complex number $z = a+ib$,

$$|z| = (z\bar{z})^{1/2}.$$

Be sure to verify the last claim in this definition. With this definition, it is important to note the following. Be sure to verify this. It is not too hard but you need to do it.

Remark 2.13.3 *: Let $z = a+ib$ and $w = c+id$. Then*

$$|z-w| = \sqrt{(a-c)^2+(b-d)^2}.$$

Thus the distance between the point in the plane determined by the ordered pair, (a,b) and the ordered pair (c,d) equals $|z-w|$ where z and w are as just described.

For example, consider the distance between $(2,5)$ and $(1,8)$. From the distance formula which you should have seen in either algebra or calculus, this distance is defined as

$$\sqrt{(2-1)^2+(5-8)^2} = \sqrt{10}.$$

On the other hand, letting $z = 2+i5$ and $w = 1+i8$, $z-w = 1-i3$ and so

$$(z-w)\left(\overline{z-w}\right) = (1-i3)(1+i3) = 10$$

so $|z-w| = \sqrt{10}$, the same thing obtained with the distance formula.

Notation 2.13.4 *From now on I will sometimes use the symbol $\mathbb{F}$ to denote either $\mathbb{C}$ or $\mathbb{R}$, rather than fussing over which one is meant because it often does not make any difference.*

The triangle inequality holds for the complex numbers just like it does for the real numbers.

Theorem 2.13.5 *Let $z, w \in \mathbb{C}$. Then*

$$|w+z| \leq |w| + |z|, \ ||z| - |w|| \leq |z-w|.$$

Proof: First note $|zw| = |z||w|$. Here is why: If $z = x+iy$ and $w = u+iv$, then

$$|zw|^2 = |(x+iy)(u+iv)|^2 = |xu - yv + i(xv+yu)|^2$$

$$= (xu-yv)^2 + (xv+yu)^2 = x^2u^2 + y^2v^2 + x^2v^2 + y^2u^2$$

Now look at the right side.

$$|z|^2 |w|^2 = (x+iy)(x-iy)(u+iv)(u-iv) = x^2u^2 + y^2v^2 + x^2v^2 + y^2u^2,$$

the same thing. Thus the rest of the proof goes just as before with real numbers. Using the results of Problem 6 on Page 38, the following holds.

$$\begin{aligned} |z+w|^2 &= (z+w)(\bar{z}+\overline{w}) = z\bar{z} + z\overline{w} + w\bar{z} + w\overline{w} \\ &= |z|^2 + |w|^2 + z\overline{w} + \overline{\overline{w}z} \\ &= |z|^2 + |w|^2 + 2\operatorname{Re} z\overline{w} \\ &\leq |z|^2 + |w|^2 + 2|z\overline{w}| = |z|^2 + |w|^2 + 2|z||w| \\ &= (|z| + |w|)^2 \end{aligned}$$

and so $|z+w| \leq |z| + |w|$ as claimed. The other inequality follows as before.

$$|z| \leq |z-w| + |w|$$

and so $|z| - |w| \leq |z-w| = |w-z|$. Now do the same argument switching the roles of z and w to conclude

$$|z| - |w| \leq |z-w|, \ |w| - |z| \leq |z-w|$$

which implies the desired inequality. ■

Since $\mathbb{R} \subseteq \mathbb{C}$ and the absolute value is consistently defined, the inequality holds also on $\mathbb{R}$.

2.14 Dividing Polynomials

It will be very important to be able to work with polynomials, especially in subjects like linear algebra and with the technique of partial fractions. It is surprising how useful this junior high material will be. In this section, a polynomial is an expression. Later, the expression will be used to define a function. These two ways of looking at a polynomial are very different.

Definition 2.14.1 *A polynomial is an expression of the form $p(\lambda) =$*

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0,$$

$a_n \neq 0$ where the a_i are (real or complex) numbers, more generally elements of a field of scalars. Two polynomials are equal means that ***the coefficients match for each power of*** *λ. The degree of a polynomial is the largest power of λ. Thus the degree of the above polynomial is n. Addition of polynomials is defined in the usual way as is multiplication of two polynomials. The leading term in the above polynomial is $a_n\lambda^n$. The coefficient of the leading term is called the leading coefficient. It is called a monic polynomial if $a_n = 1$. A* ***root*** *of a polynomial $p(\lambda)$ is μ such that $p(\mu) = 0$. This is also called a zero.*

Note that the degree of the zero polynomial is not defined in the above. The following is called the division algorithm. First is an important property of multiplication.

Lemma 2.14.2 *If $f(\lambda)g(\lambda) = 0$, then either $f(\lambda) = 0$ or $g(\lambda) = 0$. That is, there are no nonzero divisors of 0.*

Proof: Let $f(\lambda)$ have degree n and $g(\lambda)$ degree m. If $m+n = 0$, it is easy to see that the conclusion holds because both polynomials are constants. Suppose the conclusion holds for $m+n \leq M$ and suppose $m+n = M+1$. Then $f(\lambda)g(\lambda) =$

$$\begin{aligned}
&\left(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + a_n\lambda^n\right)\left(b_0 + b_1\lambda + \cdots + b_{m-1}\lambda^{m-1} + b_m\lambda^m\right) \\
=\;& (a(\lambda) + a_n\lambda^n)(b(\lambda) + b_m\lambda^m) \\
=\;& a(\lambda)b(\lambda) + b_m\lambda^m a(\lambda) + a_n\lambda^n b(\lambda) + a_n b_m\lambda^{n+m}
\end{aligned}$$

Either $a_n = 0$ or $b_m = 0$. Suppose $b_m = 0$. Then it must be the case that you have

$$(a(\lambda) + a_n\lambda^n)b(\lambda) = 0.$$

By induction, one of these polynomials in the product is 0. If $b(\lambda) \neq 0$, then this shows $a_n = 0$ and $a(\lambda) = 0$ so $f(\lambda) = 0$. If $b(\lambda) = 0$, then $g(\lambda) = 0$. The argument is the same if $a_n = 0$. ■

Say the degree of $r(\lambda)$ is $m \geq n$ where the degree of $g(\lambda)$ is n. Say $r(\lambda) = a\lambda^m + l(\lambda)$ with the degree of $l(\lambda) < m$ and $g(\lambda) = b\lambda^n + n(\lambda)$ where the degree of $n(\lambda)$ is less than n. Then $r(\lambda) - \frac{a}{b}\lambda^{m-n}g(\lambda)$ has degree smaller than m. This is used in the following fundamental lemma.

Lemma 2.14.3 *Let $f(\lambda)$ and $g(\lambda) \neq 0$ be polynomials. Then there exist polynomials, $q(\lambda)$ and $r(\lambda)$ such that*

$$f(\lambda) = q(\lambda)g(\lambda) + r(\lambda)$$

where the degree of $r(\lambda)$ is less than the degree of $g(\lambda)$ or $r(\lambda) = 0$. These polynomials $q(\lambda)$ and $r(\lambda)$ are unique.

Proof: Suppose that $f(\lambda) - q(\lambda)g(\lambda)$ is never equal to 0 for any $q(\lambda)$. If it is, then the conclusion follows. Now suppose

$$r(\lambda) = f(\lambda) - q(\lambda)g(\lambda) \tag{$*$}$$

where the degree of $r(\lambda)$ is as small as possible. Let it be m. Suppose $m \geq n$ where n is the degree of $g(\lambda)$. Say $r(\lambda) = b\lambda^m + a(\lambda)$ where $a(\lambda)$ is 0 or has degree less than m while $g(\lambda) = \hat{b}\lambda^n + \hat{a}(\lambda)$ where $\hat{a}(\lambda)$ is 0 or has degree less than n. Then

$$r(\lambda) - \frac{b}{\hat{b}}\lambda^{m-n}g(\lambda) = \overbrace{b\lambda^m + a(\lambda)}^{r(\lambda)} - \left(b\lambda^m + \frac{b}{\hat{b}}\lambda^{m-n}\hat{a}(\lambda)\right) = a(\lambda) - \tilde{a}(\lambda),$$

a polynomial having degree less than m. Therefore, from the above,

$$a(\lambda) - \tilde{a}(\lambda) = \overbrace{(f(\lambda) - q(\lambda)g(\lambda))}^{=r(\lambda)} - \frac{b}{\hat{b}}\lambda^{m-n}g(\lambda) = f(\lambda) - \hat{q}(\lambda)g(\lambda)$$

which is of the same form as $*$ having smaller degree. However, m was as small as possible. Hence $m < n$ after all.

As to uniqueness, if you have $r(\lambda), \hat{r}(\lambda), q(\lambda), \hat{q}(\lambda)$ which work, then you would have

$$(\hat{q}(\lambda) - q(\lambda))g(\lambda) = r(\lambda) - \hat{r}(\lambda)$$

Now if the polynomial on the right is not zero, then neither is the one on the left. Hence this would involve two polynomials which are equal although their degrees are different. This is impossible. Hence $r(\lambda) = \hat{r}(\lambda)$ and so, the above lemma shows $\hat{q}(\lambda) = q(\lambda)$. ■

Definition 2.14.4 *Let all coefficients of all polynomials come from a given field $\mathbb{F}$. For us, $\mathbb{F}$ will be the real numbers $\mathbb{R}$. Let $p(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$ be a polynomial. Recall it is called monic if $a_n = 1$. If you have polynomials*

$$\{p_1(\lambda), \cdots, p_m(\lambda)\},$$

the greatest common divisor $q(\lambda)$ is the monic polynomial which divides each, $p_k(\lambda) = q(\lambda)l_k(\lambda)$ for some $l_k(\lambda)$, written as $q(\lambda)/p_k(\lambda)$ and if $\hat{q}(\lambda)$ is any polynomial which divides each $p_k(\lambda)$, then $\hat{q}(\lambda)/q(\lambda)$. A set of polynomials

$$\{p_1(\lambda), \cdots, p_m(\lambda)\}$$

is relatively prime if the greatest common divisor is 1.

Lemma 2.14.5 *There is at most one greatest common divisor.*

Proof: If you had two, $\hat{q}(\lambda)$ and $q(\lambda)$, then $\hat{q}(\lambda)/q(\lambda)$ and $q(\lambda)/\hat{q}(\lambda)$ so $q(\lambda) = \hat{q}(\lambda)\hat{l}(\lambda) = q(\lambda)l(\lambda)\hat{l}(\lambda)$ and now it follows, since both $\hat{q}(\lambda)$ and $q(\lambda)$ are monic that $\hat{l}(\lambda)$ and $l(\lambda)$ are both equal to 1. ■

The next proposition is remarkable. It describes the greatest common divisor in a very useful way.

Proposition 2.14.6 *The greatest common divisor of $\{p_1(\lambda), \cdots, p_m(\lambda)\}$ exists and is characterized as the monic polynomial of smallest degree equal to an expression of the form*

$$\sum_{k=1}^{m} a_k(\lambda)p_k(\lambda), \text{ the } a_k(\lambda) \text{ being polynomials.} \tag{2.1}$$

Proof: First I need show that if $q(\lambda)$ is monic of the above form with smallest degree, then it is the greatest common divisor. If $q(\lambda)$ fails to divide $p_k(\lambda)$, then $p_k(\lambda) = q(\lambda)l(\lambda) + r(\lambda)$ where the degree of $r(\lambda)$ is smaller than the degree of $q(\lambda)$. Thus,

$$r(\lambda) = p_k(\lambda) - l(\lambda)\sum_{k=1}^{m} a_k(\lambda) p_k(\lambda)$$

which violates the condition that $q(\lambda)$ has smallest degree. Thus $q(\lambda)/p_k(\lambda)$ for each k. If $\hat{q}(\lambda)$ divides each $p_k(\lambda)$ then it must divide $q(\lambda)$ because $q(\lambda)$ is given by 2.1. Hence $q(\lambda)$ is the greatest common divisor.

Next, why does such greatest common divisor exist? Simply pick the monic polynomial which has smallest degree which is of the form $\sum_{k=1}^{m} a_k(\lambda) p_k(\lambda)$. Then from what was just shown, it is the greatest common divisor. ■

Proposition 2.14.7 *Let $p(\lambda)$ be a polynomial. Then there are polynomials $p_i(\lambda)$ such that*

$$p(\lambda) = a\prod_{i=1}^{m} p_i(\lambda)^{m_i} \tag{2.2}$$

where $m_i \in \mathbb{N}$ and $\{p_1(\lambda), \cdots, p_m(\lambda)\}$ are monic and relatively prime. Every subset of $\{p_1(\lambda), \cdots, p_m(\lambda)\}$ having at least 2 elements is also relatively prime.

Proof: If there is no polynomial of degree larger than 0 dividing $p(\lambda)$, then we are done. Simply pick a such that $p(\lambda)$ is monic. Otherwise $p(\lambda) = ap_1(\lambda)p_2(\lambda)$ where $p_i(\lambda)$ is monic and each has degree at least 1. These could be the same polynomial. If some nonconstant polynomial divides each $p_i(\lambda)$, factor further. Continue doing this. Eventually the process must end with a factorization as described in 2.2 because the degrees of the factors are decreasing. The claim about the subsets is clear because each polynomial is irreducible so the only monic polynomial dividing any of them is itself and 1. ■

2.15 The Cauchy Schwarz Inequality

This fundamental inequality takes several forms. I will present the version first given by Cauchy although I am not sure if the proof is the same.

Proposition 2.15.1 *Let z_j, w_j be complex numbers. Then*

$$\left|\sum_{j=1}^{p} z_j \overline{w_j}\right| \leq \left(\sum_{j=1}^{p} |z_j|^2\right)^{1/2} \left(\sum_{j=1}^{p} |w_j|^2\right)^{1/2}$$

Proof: First note that $\sum_{j=1}^{p} z_j \overline{z_j} = \sum_{j=1}^{p} |z_j|^2 \geq 0$. Next, if $a+ib$ is a complex number, consider $\theta = 1$ if both a, b are zero and $\theta = \frac{a-ib}{\sqrt{a^2+b^2}}$ if the complex number is not zero. Thus, in either case, there exists a complex number θ such that $|\theta| = 1$ and $\theta(a+ib) = |a+ib| \equiv \sqrt{a^2+b^2}$. Now let $|\theta| = 1$ and

$$\theta\sum_{j=1}^{p} z_j \overline{w_j} = \left|\sum_{j=1}^{p} z_j \overline{w_j}\right|$$

Then for t a real number,

$$
\begin{aligned}
0 \quad \leq \quad & p(t) \equiv \sum_{j=1}^{p} (z_j + t\theta w_j)\left(\overline{z_j} + t\overline{\theta w_j}\right) \\
= \quad & \overbrace{\sum_{j=1}^{p} z_j\overline{z_j}}^{a^2} + \sum_{j=1}^{p} z_j t\overline{\theta w_j} + \sum_{j=1}^{p} t\theta w_j\overline{z_j} + t^2\overbrace{\sum_{j=1}^{p} w_j\overline{w_j}}^{b^2}
\end{aligned}
$$

$$
\equiv a^2 + 2t \operatorname{Re}\theta \sum_{j=1}^{p} w_j\overline{z_j} + t^2 b^2 = a^2 + 2t\left|\sum_{j=1}^{p} w_j\overline{z_j}\right| + t^2 b^2
$$

Since this is always nonnegative for all real t, it follows from the quadratic formula that

$$
4\left|\sum_{j=1}^{p} w_j\overline{z_j}\right|^2 - 4a^2b^2 = 4\left|\sum_{j=1}^{p} w_j\overline{z_j}\right|^2 - 4\left(\sum_{j=1}^{p} z_j\overline{z_j}\right)\left(\sum_{j=1}^{p} w_j\overline{w_j}\right) \leq 0
$$

Indeed, $p(t) = 0$ either has exactly one real root or no real roots. Thus the desired inequality follows. ■

2.16 Integer Multiples of Irrational Numbers

This section will give a proof of a remarkable result. I think its proof, based on the pigeon hole principle, is even more interesting than the result obtained. Dirichlet proved it in the 1830's. Jacobi used similar ideas around the same time in studying elliptic functions. The theorem involves the sum of integer multiples of numbers whose ratio is irrational. If a/b is irrational, then it is not possible that $ma + nb = 0, m, m \in \mathbb{Z}$ because if this were so, you would have $\frac{-m}{n} = \frac{b}{a}$ and so the ratio of a, b is rational after all. Even though you cannot get 0 (which you can get if the ratio of a and b is rational) you can get such an integer combination arbitrarily small. Dirichlet did this in the 1830's long before Dedekind constructed the real numbers in 1858, published in 1872.

Theorem 2.16.1 *If a, b are real numbers and a/b is not rational, then for every $\varepsilon > 0$ there exist integers m, n such that $|ma + nb| < \varepsilon$.*

Proof: Let $P_N, N \geq 1$ denote all combinations of the form $ma + nb$ where m, n are integers and $|m|, |n| \leq N$. Thus there are $(2N+1)^2$ of these integer combinations and all of them are contained in the interval $I \equiv [-N(|a| + |b|), N(|a| + |b|)]$. Now pick an integer M such that

$$
(2N)^2 < M < (2N+1)^2
$$

I know such an integer exists because $(2N+1)^2 - (2N)^2 = 4N + 1 > 2$. Now partition the interval I into M equal intervals. If l is the length of one of these intervals, then

$$
lM = 2N(|a| + |b|), \text{ so } (2N)^2 l < 2N(|a| + |b|) \text{ and } l < \frac{(|a| + |b|)}{2N} \equiv \frac{C}{N}
$$

Now as mentioned, all of the points of P_N are contained in I and there are more of these points, $(2N+1)^2$ than there are intervals, M. Therefore, some interval contains two points of P_N.[3] But each interval has length no more than C/N and so there exist $k,\hat{k},l,\hat{l}$ integers such that

$$\left|ka+lb-\left(\hat{k}a+\hat{l}b\right)\right| \equiv |ma+nb| < \frac{C}{N}$$

Now let $\varepsilon > 0$ be given. Choose N large enough that $C/N < \varepsilon$. Then the above inequality holds for some integers m,n. ■

2.17 Exercises

1. Let $z = 5+i9$. Find z^{-1}.

2. Let $z = 2+i7$ and let $w = 3-i8$. Find $zw, z+w, z^2$, and w/z.

3. If z is a complex number, show there exists ω a complex number with $|\omega| = 1$ and $\omega z = |z|$.

4. For those who know about the trigonometric functions [4], De Moivre's theorem says $[r(\cos t+i\sin t)]^n = r^n(\cos nt+i\sin nt)$ for n a positive integer. Prove this formula by induction. Does this formula continue to hold for all integers n, even negative integers? Explain.

5. Using De Moivre's theorem from Problem 4, derive a formula for $\sin(5x)$ and one for $\cos(5x)$. **Hint:** Use Problem 18 on Page 25 and if you like, you might use Pascal's triangle to construct the binomial coefficients.

6. If z,w are complex numbers prove $\overline{zw} = \overline{z}\,\overline{w}$. Show that $\overline{z_1\cdots z_m} = \overline{z_1}\cdots\overline{z_m}$. Also verify that $\overline{\sum_{k=1}^m z_k} = \sum_{k=1}^m \overline{z_k}$. In words this says the conjugate of a product equals the product of the conjugates and the conjugate of a sum equals the sum of the conjugates.

7. Suppose $p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ where all the a_k are real numbers. Suppose also that $p(z) = 0$ for some $z \in \mathbb{C}$. Show it follows that $p(\overline{z}) = 0$ also.

8. I claim that $1 = -1$. Here is why. $-1 = i^2 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)^2} = \sqrt{1} = 1$. This is clearly a remarkable result but is there something wrong with it? If so, what is wrong? **Hint:** When we push symbols without consideration of their meaning, we can accomplish many strange and wonderful but false things.

9. De Moivre's theorem of Problem 4 is really a grand thing. I plan to use it now for rational exponents, not just integers. $1 = 1^{(1/4)} = (\cos 2\pi + i\sin 2\pi)^{1/4} = \cos(\pi/2) + i\sin(\pi/2) = i$. Therefore, squaring both sides it follows $1 = -1$ as in the previous problem. What does this tell you about De Moivre's theorem? Is there a profound difference between raising numbers to integer powers and raising numbers to non integer powers?

[3] This is called the pigeon hole principle. It was used by Jacobi and Dirichlet. Later, Besicovitch used it in his amazing covering theorem. In terms of pigeons, it says that if you have more pigeons than holes and they each need to go in a hole, then some hole must have more than one pigeon. In contrast to Dirichlet, Jacobi and others, (those who had common sense) this simple observation seems to have not been understood by people like Brigham Young the Utah Mormon leader who made polygyny (multiple wives for a single man) a religious expectation in the 1850's. In Utah there were more males than females.

[4] I will present a treatment of the trig functions which is independent of plane geometry a little later.

10. Review Problem 4 at this point. Now here is another question: If n is an integer, is it always true that $(\cos\theta - i\sin\theta)^n = \cos(n\theta) - i\sin(n\theta)$? Explain.

11. Suppose you have any polynomial in $\cos\theta$ and $\sin\theta$. By this I mean an expression of the form $\sum_{\alpha=0}^{m}\sum_{\beta=0}^{n} a_{\alpha\beta}\cos^{\alpha}\theta\sin^{\beta}\theta$ where $a_{\alpha\beta}\in\mathbb{C}$. Can this always be written in the form $\sum_{\gamma=-(n+m)}^{m+n} b_{\gamma}\cos\gamma\theta + \sum_{\tau=-(n+m)}^{n+m} c_{\tau}\sin\tau\theta$? Explain.

12. Does there exist a subset of $\mathbb{C}$, $\mathbb{C}^{+}$ which satisfies 2.4.1 - 2.4.3? **Hint:** You might review the theorem about order. Show -1 cannot be in $\mathbb{C}^{+}$. Now ask questions about $-i$ and i. In mathematics, you can sometimes show certain things do not exist. It is very seldom you can do this outside of mathematics. For example, does the Loch Ness monster exist? Can you prove it does not?

13. Show that if a/b is irrational, then $\{ma+nb\}_{m,n\in\mathbb{Z}}$ is dense in $\mathbb{R}$, each an irrational number. If a/b is rational, show that $\{ma+nb\}_{m,n\in\mathbb{Z}}$ is not dense. **Hint:** From Theorem 2.16.1 there exist integers, m_l, n_l such that $|m_l a + n_l b| < 2^{-l}$. Let $P_l \equiv \cup_{k\in\mathbb{Z}}\{k(m_l a + n_l b)\}$. Thus this is a collection of numbers which has successive numbers 2^{-l} apart. Then consider $\cup_{l\in\mathbb{N}} P_l$. In case the ratio is rational and $\{ma+nb\}_{m,n\in\mathbb{Z}}$ is dense, explain why there are relatively prime integers p, q such that $p/q = a/b$ is rational and $\{mp+nq\}_{m,n\in\mathbb{Z}}$ would be dense. Isn't this last a collection of integers?

14. This problem will show, as a special case, that the rational numbers are dense in $\mathbb{R}$. Referring to the proof of Theorem 2.16.1.

    (a) Suppose $\alpha\in(0,1)$ and is irrational. Show that if N is a positive integer, then there are integers m, n such that $0 < n \le N$ and $|n\alpha - m| < \frac{1}{N}\frac{1}{2}(1+\alpha) < \frac{1}{N}$. Thus $\left|\alpha - \frac{m}{n}\right| < \frac{1}{nN} \le \frac{1}{n^2}$.

    (b) Show that if β is any nonzero irrational number, and N is a positive integer, there exists $0 < n \le N$ and an integer m such that $\left|\beta - \frac{m}{n}\right| < \frac{1}{nN} \le \frac{1}{n^2}$. **Hint:** You might consider $\beta - [\beta] \equiv \alpha$ where $[\beta]$ is the integer no larger than β which is as large as possible.

    (c) Next notice that from the proof, the same will hold for any β a positive number. **Hint:** In the proof, if there is a repeat in the list of numbers, then you would have an exact approximation. Otherwise, the pigeon hole principle applies as before. Now explain why nothing changes if you only assume β is a nonzero real number.

15. This problem outlines another way to see that rational numbers are dense in $\mathbb{R}$. Pick $x\in\mathbb{R}$. Explain why there exists m_l, the smallest integer such that $2^{-l}m_l \ge x$ so $x \in (2^{-l}(m_l - 1), 2^{-l}m_l]$. Now note that $2^{-l}m_l$ is rational and closer to x than 2^{-l}.

16. You have a rectangle R having length 4 and height 3. There are six points in R. One is at the center. Show that two of them are as close as $\sqrt{5}$. You might use pigeon hole principle.

17. Do the same problem without assuming one point is at the center. **Hint:** Consider the pictures. If not, then by pigeon hole principle, there is exactly one point in each of the six rectangles in first two pictures.

18. Suppose $r(\lambda) = \frac{a(\lambda)}{p(\lambda)^m}$ where $a(\lambda)$ is a polynomial and $p(\lambda)$ is an irreducible polynomial meaning that the only polynomials dividing $p(\lambda)$ are numbers and scalar multiples of $p(\lambda)$. That is, you can't factor it any further. Show that $r(\lambda)$ is of the form

$$r(\lambda) = q(\lambda) + \sum_{k=1}^{m} \frac{b_k(\lambda)}{p(\lambda)^k}, \text{ where degree of } b_k(\lambda) < \text{ degree of } p(\lambda)$$

19. $\uparrow$Suppose you have a rational function $\frac{a(\lambda)}{b(\lambda)}$.

    (a) Show it is of the form $p(\lambda) + \frac{n(\lambda)}{\prod_{i=1}^{m} p_i(\lambda)^{m_i}}$ where $\{p_1(\lambda), \cdots, p_m(\lambda)\}$ are relatively prime and the degree of $n(\lambda)$ is less than the degree of $\prod_{i=1}^{m} p_i(\lambda)^{m_i}$.

    (b) Using Proposition 2.14.6 and the division algorithm for polynomials, show that the original rational function is of the form

$$\hat{p}(\lambda) + \sum_{i=1}^{m} \sum_{k=1}^{m_i} \frac{n_{ki}(\lambda)}{p_i(\lambda)^k}$$

    where the degree of $n_{ki}(\lambda)$ is less than the degree of $p_i(\lambda)$ and $\hat{p}(\lambda)$ is some polynomial.

    This is the partial fractions expansion of the rational function. Actually carrying out this computation may be impossible, but this shows the existence of such a partial fractions expansion. **Hint:** You might induct on the sum of the m_i and use Proposition 2.14.6.

20. One can give a fairly simple algorithm to find the $g.c.d.$, greatest common divisor of two polynomials. The coefficients are in some field. For us, this will be either the real, rational, or complex numbers. However, in general, the algorithm for long division would be carried out in whatever field includes the coefficients. Explain the following steps. Let $r_0(\lambda), r_1(\lambda)$ be polynomials with the degree of $r_0(\lambda)$ at least as large as the degree of $r_1(\lambda)$. Then do division.

$$r_0(\lambda) = r_1(\lambda) f_1(\lambda) + r_2(\lambda)$$

    where $r_2(\lambda)$ has smaller degree than $r_1(\lambda)$ or else is 0. If $r_2(\lambda)$ is 0, then the $g.c.d.$ of $r_1(\lambda), r_0(\lambda)$ is $r_1(\lambda)$. Otherwise, $l(\lambda) / r_0(\lambda), r_1(\lambda)$ if and only if

$$l(\lambda) / r_1(\lambda), r_2(\lambda).$$

    Do division again

$$r_1(\lambda) = r_2(\lambda) f_2(\lambda) + r_3(\lambda)$$

where $\deg(r_3(\lambda)) < \deg(r_2(\lambda))$ or $r_3(\lambda)$ is 0. Then $l(\lambda)/r_2(\lambda), r_3(\lambda)$ if and only if $l(\lambda)$ divides both $r_1(\lambda)$ and $r_2(\lambda)$ if and only if $l(\lambda)/r_0(\lambda), r_1(\lambda)$. If $r_3(\lambda) = 0$, then $r_2(\lambda)/r_2(\lambda), r_1(\lambda)$ so also $r_2(\lambda)/r_0(\lambda), r_1(\lambda)$ and also, if

$$l(\lambda)/r_0(\lambda), r_1(\lambda),$$

then $l(\lambda)/r_1(\lambda), r_2(\lambda)$ and in particular, $l(\lambda)/r_1(\lambda)$ so if this happens, then $r_2(\lambda)$ is the $g.c.d.$ of $r_0(\lambda)$ and $r_1(\lambda)$. Continue doing this. Eventually either $r_{m+1}(\lambda) = 0$ or has degree 0. If $r_{m+1}(\lambda) = 0$, then $r_m(\lambda)$ multiplied by a suitable scalar to make the result a monic polynomial is the $g.c.d.$ of $r_0(\lambda)$ and $r_1(\lambda)$. If the degree is 0, then the two polynomials $r_0(\lambda), r_1(\lambda)$ must be relatively prime. It is really significant that this can be done because fundamental theorems in linear algebra depend on whether two polynomials are relatively prime having $g.c.d.$ equal to 1. In this application, it is typically a question about a polynomial and its derivative.

21. Find the $g.c.d.$ for $\left(x^4 + 3x^2 + 2\right), \left(x^2 + 3\right)$.

22. Find $g.c.d.$ of $\left(x^5 + 3x^3 + x^2 + 3\right), \left(x^2 + 3\right)$.

23. Find the $g.c.d.$ of $\left(x^6 + 2x^5 + x^4 + 3x^3 + 2x^2 + x + 2\right), \left(x^4 + 2x^3 + x + 2\right)$.

24. Find the $g.c.d.$ of $\left(x^4 + 3x^3 + 2x + 1\right)$, $\left(4x^3 + 9x^2 + 2\right)$. If you do this one by hand, it might be made easier to note that the question of interest is resolved if you multiply everything with a nonzero scalar before you do long division.

25. Prove the pigeon hole principle by induction.

Chapter 3

Set Theory

3.1 Basic Definitions

This chapter has more on set theory. Recall a set is a collection of things called elements of the set. For example, the set of integers, the collection of signed whole numbers such as 1,2,-4, etc. This set whose existence will be assumed 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 parentheses, $\{\ \}$ specify a set by listing 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 $\{-1,0,1,2\}$. The notation signifying x is an element of a set S, is written as $x \in S$. Thus, $1 \in \{-1,0,1,2,3\}$. Here are some axioms about sets.

Axiom 3.1.1 *Two sets are equal if and only if they have the same elements.*

Axiom 3.1.2 *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.*

Axiom 3.1.3 *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.*

Axiom 3.1.4 *The Cartesian product of a nonempty family of nonempty sets is nonempty.*

Axiom 3.1.5 *If A is a set there exists a set $\mathscr{P}(A)$, such that $\mathscr{P}(A)$ is the set of all subsets of A. This is called the power set.*

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 you 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 you should be able 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, 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.

Another example of political interest, could be the set of all judges who are not judicial activists. I think you can see this last is not a very precise condition since there is no way to determine to everyone's satisfaction whether a given judge is an activist. Also, **just because something is grammatically correct does not mean it makes any sense.** For example consider the following nonsense.

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

So what is a condition?

We will leave these sorts of considerations and assume our conditions "make sense". The axiom of unions states that for 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 $\mathscr{S}$ is such a set whose elements are sets,

$$\cup\{A : A \in \mathscr{S}\} \text{ or } \cup\mathscr{S}$$

signifies this union.

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 $\mathscr{S}$ is a set whose elements are sets, $\prod_{A\in\mathscr{S}} A$ signifies the Cartesian product.

The Cartesian product is the set of choice functions, a choice function being a function which selects exactly one element of each set of $\mathscr{S}$. Functions will be described precisely soon. The idea is that there is something, which will produce a set consisting of exactly one element of each set of $\mathscr{S}$. 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, things which never happen without the axiom of choice.

A is a subset of B, written $A \subseteq B$, if every element of A is also an element of B. This can also be written as $B \supseteq A$. A is a proper subset of B, written $A \subset B$ or $B \supset A$ if A is a subset of B but A is not equal to $B, A \neq B$. However, this is not entirely consistent. Sometimes people write $A \subset B$ when they mean $A \subseteq B$. $A \cap B$ denotes the intersection of the two sets, A and 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 $\emptyset$. $A \cup B$ denotes the union of the two sets, A and B and it means the set of all elements which are in either of the sets. It is a set because of 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, a set which contains the one whose complement is being taken. Thus, the complement of A, denoted as A^C (or more precisely as $X \setminus A$) is a set obtained from using the axiom of specification to write

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

The symbol $\notin$ means: "is not an element of". Note the axiom of specification takes place relative to a given set. Without this universal set, it makes no sense to use the axiom of specification to obtain the complement.

Words such as "all" or "there exists" are called quantifiers and they must be understood relative to some given set. For example, the set of all integers larger than 3. Or 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.

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

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

and

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

These laws follow directly from the definitions. Also following directly from the definitions are:

Let $\mathscr{S}$ be a set of sets then

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

and: Let $\mathscr{S}$ be a set of sets show

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

Unfortunately, there is no single universal set which can be used for all sets. Here is why: Suppose there were. Call it S. Then you could consider A the set of all elements of S which are not elements of themselves, this from the axiom of specification. If A is an element of itself, then it fails to qualify for inclusion in A. Therefore, it must not be an element of itself. However, if A is not an element of itself, it qualifies for inclusion in A so it is an element of itself and so this can't be true either. Thus the most basic of conditions you could imagine, that of being an element of, is meaningless and so allowing such a set causes the whole theory to be meaningless. The solution is to not allow a universal set. As mentioned by Halmos in Naive set theory, "Nothing contains everything". Always beware of statements involving quantifiers wherever they occur, even this one. This little observation described above is due to Bertrand Russell and is called Russell's paradox.

Example 3.1.6 *Various religions, including my own, use the word "omnipotent" as an attribute of god. It "means" god can do all things. Isn't there a universal quantifier with no universal set specified? Incidentally, when speaking to religious people, it is often best not to call attention to this fact so they won't think you are an atheist like Russell. Many of the same people who believe in an "omnipotent" god are concerned with the problem of evil (theodicy). Why does god allow evil, suffering, and sorrow? This leads to: Why does an omnipotent god allow these things? Is god even "good"? I have heard much agonizing over the latter question in my life, but never any consideration whether it makes sense.*

Theodicy has concerned intelligent people since the time of Jeremiah. See Chapter 12 of Jeremiah for example, and the profound discussion in the book of Job. However, linking theodicy to illogical words only makes it even more difficult and challenges the existence of God for those who don't realize that the omni words don't make good sense. It only takes one of these to make their god's existence meaningless, but religious people usually insist on saddling god with several of them. If they knew about Russell's paradox it would help.

3.2 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. The Cartesian product is discussed above. The next definition reviews this and defines the abstract notion of a function.

Definition 3.2.1 *Let X, 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, also called a mapping, 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*

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

written as $f:D(f)\to Y$ *and we write* $y=f(x)$*. Another notation which is used is the following*

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

This is called the inverse image.

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 like 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 or mapping has an input, x and a unique output y, denoted as $f(x)$ while the name of the function is f. "mapping" is often a noun meaning function. However, it also is a verb as in "f is mapping A to B". That which a function is thought of as doing is also referred to using the word "maps" as in: f maps X to Y. However, a set of functions may be called a set of maps so this word might also be used as the plural of a noun. There is no help for it. You just have to suffer with this nonsense.

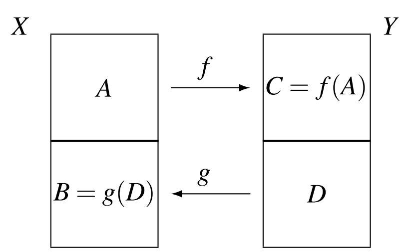
The following theorem which is interesting for its own sake will be used to prove the Schroder Bernstein theorem, proved by Dedekind in 1887. The shortest proof I have seen is in Hewitt and Stromberg [17] and this is the version given here. There is another version in Halmos [15].

Theorem 3.2.2 *Let* $f:X\to Y$ *and* $g:Y\to X$ *be two functions. 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: Consider the empty set, $\emptyset\subseteq X$. If $y\in Y\setminus f(\emptyset)$, then $g(y)\notin\emptyset$ because $\emptyset$ has no elements. Also, if A,B,C, and D are as described above, A also would have this same property that the empty set has. However, A is probably larger. Therefore, say $A_0\subseteq X$ satisfies $\mathscr{P}$ if whenever $y\in Y\setminus f(A_0)$, $g(y)\notin A_0$.

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

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

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

It only remains to verify that $g(D) = B$. It was just shown that $g(D) \subseteq B$.

Suppose $x \in B = X \setminus A$. Then $A \cup \{x\}$ does not satisfy $\mathscr{P}$ because it is too large, and so there exists $y \in Y \setminus f(A \cup \{x\}) \subseteq D$ such that $g(y) \in A \cup \{x\}$. But $y \notin f(A)$ and so since A satisfies $\mathscr{P}$, it follows $g(y) \notin A$. Hence $g(y) = x$ and so $x \in g(D)$. Hence $g(D) = B$. ■

Theorem 3.2.3 *(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 Theorem 3.2.2 and define

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

Then h is the desired one to one and onto mapping. ■

Recall that the Cartesian product may be considered as the collection of choice functions.

Definition 3.2.4 *Let I be a set and let X_i be a nonempty set for each $i \in I$. f is a choice function written as*

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

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

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

Sometimes the two functions, f and g are onto but not one to one. It turns out that with the axiom of choice, a similar conclusion to the above may be obtained.

Corollary 3.2.5 *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 $y \in Y$, $f^{-1}(y) \equiv \{x \in X : f(x) = y\} \neq \emptyset$. Therefore, by the axiom of choice, there exists $f_0^{-1} \in \prod_{y \in Y} f^{-1}(y)$ which is the same as saying that for each $y \in Y$, $f_0^{-1}(y) \in f^{-1}(y)$. Similarly, there exists $g_0^{-1}(x) \in g^{-1}(x)$ for all $x \in X$. Then f_0^{-1} is one to one because if $f_0^{-1}(y_1) = f_0^{-1}(y_2)$, then

$$y_1 = f\left(f_0^{-1}(y_1)\right) = f\left(f_0^{-1}(y_2)\right) = y_2.$$

Similarly g_0^{-1} is one to one. Therefore, by the Schroder Bernstein theorem, there exists $h : X \to Y$ which is one to one and onto. ■

We have already made reference to finite sets in the pigeon hole principle. The following is just a more formal definition of what is meant by a finite set and this is generalized to what is meant by a countable set.

Definition 3.2.6 *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. 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, is written as $\theta : A \to B$ in the future.) Here $\mathbb{N} \equiv \{1,2,\cdots\}$, the natural numbers. S is at most countable if there exists a map $\theta : \mathbb{N} \to S$ which is onto.*

The property of being at most countable is often referred to as being countable 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 give each element of the set a natural number. The possibility that a single element of the set may be counted more than once is often not important.

Theorem 3.2.7 *If X and Y are both at most countable, then $X \times Y$ is also at most countable. If either X or Y is countable, then $X \times Y$ is also countable.*

Proof: It is given that there exists a mapping $\eta : \mathbb{N} \to X$ which is onto. Define $\eta(i) \equiv x_i$ and consider X as the set $\{x_1,x_2,x_3,\cdots\}$. Similarly, consider Y as the set $\{y_1,y_2,y_3,\cdots\}$. It follows the elements of $X \times Y$ are included in the following rectangular array.

$$\begin{array}{ccccl}
(x_1,y_1) & (x_1,y_2) & (x_1,y_3) & \cdots & \leftarrow \text{Those which have } x_1 \text{ in first slot.} \\
(x_2,y_1) & (x_2,y_2) & (x_2,y_3) & \cdots & \leftarrow \text{Those which have } x_2 \text{ in first slot.} \\
(x_3,y_1) & (x_3,y_2) & (x_3,y_3) & \cdots & \leftarrow \text{Those which have } x_3 \text{ in first slot.} \\
\vdots & \vdots & \vdots & & \qquad\qquad \vdots
\end{array} .$$

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. This assigns a number from $\mathbb{N}$ to each element of $X \times Y$. Thus $X \times Y$ is at most countable.

It remains to show the last claim. 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 $\beta((x,y)) \equiv \alpha^{-1}(x)$. Thus β is onto $\mathbb{N}$. By the first part there exists a function from $\mathbb{N}$ onto $X \times Y$. Therefore, by Corollary 3.2.5, there exists a one to one and onto mapping from $X \times Y$ to $\mathbb{N}$. ■

Theorem 3.2.8 *If X and Y are at most countable, then $X \cup Y$ is at most countable. If either X or Y is infinite, then $X \cup Y$ is countable.*

Proof: As in the preceding theorem,

$$X = \{x_1,x_2,x_3,\cdots\}$$

and

$$Y = \{y_1,y_2,y_3,\cdots\}.$$

Consider the following array consisting of $X \cup Y$ and path through it.

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

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.

Consider the second claim. By the first part, there is a map from $\mathbb{N}$ onto $X \times Y$. Suppose without loss of generality that X is countable and $\alpha : \mathbb{N} \to X$ 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 this shows there exist two onto maps, one mapping $X \cup Y$ onto $\mathbb{N}$ and the other mapping $\mathbb{N}$ onto $X \cup Y$. Then Corollary 3.2.5 yields the conclusion. ■

Note that by induction this shows that if you have any finite set whose elements are countable sets, then the union of these is countable. In fact, you can say that a countable union of countable sets is countable.

Theorem 3.2.9 *Let A_i be a countable set. Thus $A_i = \left\{r_j^i\right\}_{j=1}^{\infty}$. Then $\cup_{i=1}^{\infty} A_i$ is also at most a countable set. If it is an infinite set, then it is countable.*

Proof: This is proved like Theorem 3.2.7 arrange $\cup_{i=1}^{\infty} A_i$ as follows.

$$\begin{array}{cccc} r_1^1 & r_2^1 & r_3^1 & \cdots \\ r_1^2 & r_2^2 & r_3^2 & \cdots \\ r_1^3 & r_2^3 & r_3^3 & \cdots \\ \vdots & \vdots & \vdots & \end{array}$$

Now take a route through this rectangular array as in Theorem 3.2.7, identifying an enumeration in the order in which the displayed elements are encountered as done in that theorem. Thus there is an onto mapping from $\mathbb{N}$ to $\cup_{i=1}^{\infty} A_i$ and so $\cup_{i=1}^{\infty} A_i$ is at most countable, meaning its elements can be enumerated. However, if any of the A_i is infinite or if the union is, then there is an onto map from $\cup_{i=1}^{\infty} A_i$ onto $\mathbb{N}$ and so from Corollary 3.2.5, there would be a one to one and onto map between $\mathbb{N}$ and $\cup_{i=1}^{\infty} A_i$. ■

As mentioned, in virtually all applications to analysis, the topic of main interest is "at most countable" meaning that the elements of a set S can be listed with subscripts from $\mathbb{N}$ to obtain them all. More precisely, there is a map from $\mathbb{N}$ onto the set S. Often people simply refer to such a set as countable.

3.3 Equivalence Relations

There are many ways to compare elements of a set other than to say two elements are equal or the same. For example, in the set of people let two people be equivalent if they have the same weight. This would not be saying they were the same person, just that they weighed the same. Often such relations involve considering one characteristic of the elements of a set and then saying the two elements are equivalent if they are the same as far as the given characteristic is concerned.

Definition 3.3.1 *Let S be a set. $\sim$ is an equivalence relation on S if it satisfies the following axioms.*

1. *$x \sim x$ for all $x \in S$. (Reflexive)*

2. *If $x \sim y$ then $y \sim x$. (Symmetric)*

3. *If $x \sim y$ and $y \sim z$, then $x \sim z$. (Transitive)*

Definition 3.3.2 *$[x]$ denotes the set of all elements of S which are equivalent to x and $[x]$ is called the equivalence class determined by x or just the equivalence class of x.*

With the above definition one can prove the following simple theorem.

Theorem 3.3.3 *Let $\sim$ be an equivalence relation defined on a set, S and let $\mathscr{H}$ denote the set of equivalence classes. Then if $[x]$ and $[y]$ are two of these equivalence classes, either $x \sim y$ and $[x] = [y]$ or it is not true that $x \sim y$ and $[x] \cap [y] = \emptyset$.*

3.4 Hausdorff Maximal Theorem*

The Hausdorff maximal theorem is equivalent to the axiom of choice. Hausdorff proved it in 1914. The useful direction is what I will prove below. It will not be used much in the rest of the book which is mostly nineteenth century material. I am including it because it or something like it is either absolutely essential, as in the Hahn Banach theorem, or extremely useful to have.

Definition 3.4.1 *For any set S, $\mathscr{P}(S)$ denotes the set of all subsets of S. It is sometimes called the power set of S and is also sometimes denoted as 2^S. A nonempty set $\mathscr{F}$ is called a partially ordered set if it has a partial order denoted by $\prec$. This means it satisfies the following. If $x \prec y$ and $y \prec z$, then $x \prec z$. Also $x \prec x$. It is like $\subseteq$ on the set of all subsets of a given set. It is not the case that given two elements of $\mathscr{F}$ that they are related. In other words, you cannot conclude that either $x \prec y$ or $y \prec x$. A chain, denoted by $\mathscr{C} \subseteq \mathscr{F}$ has the property that it is totally ordered meaning that if $x, y \in \mathscr{C}$, either $x \prec y$ or $y \prec x$. A maximal chain is a chain $\mathscr{C}$ which has the property that there is no strictly larger chain. In other words, if $x \in \mathscr{F} \setminus \cup \mathscr{C}$, then $\mathscr{C} \cup \{x\}$ is no longer a chain.*

Here is the Hausdorff maximal theorem. The proof is a proof by contradiction. We assume there is no maximal chain and then show that this cannot happen. The axiom of choice is used in choosing the $x_{\mathscr{C}}$ right at the beginning of the argument. See Hewitt and Stromberg [17] for more of this kind of thing.

Theorem 3.4.2 *Let $\mathscr{F}$ be a nonempty partially ordered set with order $\prec$. Then there exists a maximal chain.*

Proof: Suppose not. Then for $\mathscr{C}$ a chain, let $\theta\mathscr{C}$ denote $\mathscr{C} \cup \{x_{\mathscr{C}}\}$. Thus for $\mathscr{C}$ a chain, $\theta\mathscr{C}$ is a larger chain which has exactly one more element of $\mathscr{F}$. Since $\mathscr{F} \neq \emptyset$, pick $x_0 \in \mathscr{F}$. Note that $\{x_0\}$ is a chain. Let $\mathscr{X}$ be the set of all chains $\mathscr{C}$ such that $x_0 \in \cup\mathscr{C}$. Thus $\mathscr{X}$ contains $\{x_0\}$. Call two chains comparable if one is a subset of the other. Also, if $\mathscr{S}$ is a nonempty subset of $\mathscr{F}$ in which all chains are comparable, then $\cup\mathscr{S}$ is also a chain. From now on $\mathscr{S}$ **will always refer to a nonempty set of chains in which any pair are comparable**. Then summarizing,

1. $x_0 \in \cup\mathscr{C}$ for all $\mathscr{C} \in \mathscr{X}$.

2. $\{x_0\} \in \mathscr{X}$

3. If $\mathscr{C} \in \mathscr{X}$ then $\theta\mathscr{C} \in \mathscr{X}$.

4. If $\mathscr{S} \subseteq \mathscr{X}$ then $\cup\mathscr{S} \in \mathscr{X}$.

A subset $\mathscr{Y}$ of $\mathscr{X}$ will be called a "tower" if $\mathscr{Y}$ satisfies 1.) - 4.). Let $\mathscr{Y}_0$ be the intersection of all towers. Then $\mathscr{Y}_0$ is also a tower, the smallest one. Then the next claim might seem to be so because if not, $\mathscr{Y}_0$ would not be the smallest tower.

Claim 1: If $\mathscr{C}_0 \in \mathscr{Y}_0$ is comparable to every chain $\mathscr{C} \in \mathscr{Y}_0$, then if $\mathscr{C}_0 \subsetneq \mathscr{C}$, it must be the case that $\theta\mathscr{C}_0 \subseteq \mathscr{C}$. In other words, $x_{\mathscr{C}_0} \in \cup\mathscr{C}$. The symbol $\subsetneq$ indicates proper subset.

This is done by considering a set $\mathscr{B} \subseteq \mathscr{Y}_0$ consisting of $\mathscr{D}$ which acts like $\mathscr{C}$ in the above and showing that it actually equals $\mathscr{Y}_0$ because it is a tower.

Proof of Claim 1: Consider $\mathscr{B} \equiv \left\{\mathscr{D} \in \mathscr{Y}_0 : \mathscr{D} \subseteq \mathscr{C}_0 \text{ or } x_{\mathscr{C}_0} \in \cup\mathscr{D}\right\}$. Let $\mathscr{Y}_1 \equiv \mathscr{Y}_0 \cap \mathscr{B}$. I want to argue that $\mathscr{Y}_1$ is a tower. By definition all chains of $\mathscr{Y}_1$ contain x_0 in their unions. If $\mathscr{D} \in \mathscr{Y}_1$, is $\theta\mathscr{D} \in \mathscr{Y}_1$? If $\mathscr{S} \subseteq \mathscr{Y}$, is $\cup\mathscr{S} \in \mathscr{Y}_1$? Is $\{x_0\} \in \mathscr{B}$?

$\{x_0\}$ cannot properly contain $\mathscr{C}_0$ since $x_0 \in \cup\mathscr{C}_0$. Therefore, $\mathscr{C}_0 \supseteq \{x_0\}$ so $\{x_0\} \in \mathscr{B}$.

If $\mathscr{S} \subseteq \mathscr{Y}_1$, and $\mathscr{D} \equiv \cup\mathscr{S}$, is $\mathscr{D} \in \mathscr{Y}_1$? Since $\mathscr{Y}_0$ is a tower, $\mathscr{D}$ is comparable to $\mathscr{C}_0$. If $\mathscr{D} \subseteq \mathscr{C}_0$, then $\mathscr{D}$ is in $\mathscr{B}$. Otherwise $\mathscr{D} \supsetneq \mathscr{C}_0$ and in this case, why is $\mathscr{D}$ in $\mathscr{B}$? Why is $x_{\mathscr{C}_0} \in \cup\mathscr{D}$? The chains of $\mathscr{S}$ are in $\mathscr{B}$ so one of them, called $\tilde{\mathscr{C}}$ must properly contain $\mathscr{C}_0$ and so $x_{\mathscr{C}_0} \in \cup\tilde{\mathscr{C}} \subseteq \cup\mathscr{D}$. Therefore, $\mathscr{D} \in \mathscr{B} \cap \mathscr{Y}_0 = \mathscr{Y}_1$. 4.) holds. Two cases remain, to show that $\mathscr{Y}_1$ satisfies 3.).

case 1: $\mathscr{D} \supsetneq \mathscr{C}_0$. Then by definition of $\mathscr{B}$, $x_{\mathscr{C}_0} \in \cup\mathscr{D}$ and so $x_{\mathscr{C}_0} \in \cup\theta\mathscr{D}$ so $\theta\mathscr{D} \in \mathscr{Y}_1$.

case 2: $\mathscr{D} \subseteq \mathscr{C}_0$. $\theta\mathscr{D} \in \mathscr{Y}_0$ so $\theta\mathscr{D}$ is comparable to $\mathscr{C}_0$. First suppose $\theta\mathscr{D} \supsetneq \mathscr{C}_0$. Thus $\mathscr{D} \subseteq \mathscr{C}_0 \subsetneq \mathscr{D} \cup \{x_{\mathscr{D}}\}$. If $x \in \mathscr{C}_0$ and x is not in $\mathscr{D}$ then $\mathscr{D} \cup \{x\} \subseteq \mathscr{C}_0 \subsetneq \mathscr{D} \cup \{x_{\mathscr{D}}\}$. This is impossible. Consider x. Thus in this case that $\theta\mathscr{D} \supsetneq \mathscr{C}_0$, $\mathscr{D} = \mathscr{C}_0$. It follows that $x_{\mathscr{D}} = x_{\mathscr{C}_0} \in \cup\theta\mathscr{C}_0 = \cup\theta\mathscr{D}$ and so $\theta\mathscr{D} \in \mathscr{Y}_1$. The other case is that $\theta\mathscr{D} \subseteq \mathscr{C}_0$ so $\theta\mathscr{D} \in \mathscr{B}$ by definition. This shows 3.) so $\mathscr{Y}_1$ is a tower and must equal $\mathscr{Y}_0$.

Claim 2: Any two chains in $\mathscr{Y}_0$ are comparable.

Proof of Claim 2: Let $\mathscr{Y}_1$ consist of all chains of $\mathscr{Y}_0$ which are comparable to every chain of $\mathscr{Y}_0$. $\{x_0\}$ is in $\mathscr{Y}_1$ by definition. All chains of $\mathscr{Y}_0$ have x_0 in their union. If $\mathscr{S} \subseteq \mathscr{Y}_1$, is $\cup\mathscr{S} \in \mathscr{Y}_1$? Given $\mathscr{D} \in \mathscr{Y}_0$ either every chain of $\mathscr{S}$ is contained in $\mathscr{D}$ or at least one contains $\mathscr{D}$. Either way $\mathscr{D}$ is comparable to $\cup\mathscr{S}$ so $\cup\mathscr{S} \in \mathscr{Y}_1$. It remains to show 3.). Let $\mathscr{C} \in \mathscr{Y}_1$ and $\mathscr{D} \in \mathscr{Y}_0$. Since $\mathscr{C}$ is comparable to all chains in $\mathscr{Y}_0$, it follows from Claim 1 either $\mathscr{C} \subsetneq \mathscr{D}$ when $x_{\mathscr{C}} \in \cup\mathscr{D}$ and $\theta\mathscr{C} \subseteq \mathscr{D}$ or $\mathscr{C} \supseteq \mathscr{D}$ when $\theta\mathscr{C} \supseteq \mathscr{D}$. Hence $\mathscr{Y}_1 = \mathscr{Y}_0$ because $\mathscr{Y}_0$ is as small as possible.

Since every pair of chains in $\mathscr{Y}_0$ are comparable and $\mathscr{Y}_0$ is a tower, it follows that $\cup\mathscr{Y}_0 \in \mathscr{Y}_0$ so $\cup\mathscr{Y}_0$ is a chain. However, $\theta \cup \mathscr{Y}_0$ is a chain which properly contains $\cup\mathscr{Y}_0$ and since $\mathscr{Y}_0$ is a tower, $\theta \cup \mathscr{Y}_0 \in \mathscr{Y}_0$. Thus $\cup(\theta \cup \mathscr{Y}_0) \supsetneq \cup(\cup\mathscr{Y}_0) \supseteq \cup(\theta \cup \mathscr{Y}_0)$ which is a contradiction. Therefore, for some chain $\mathscr{C}$ it is impossible to obtain the x_C described above and so, this $\mathscr{C}$ is a maximal chain. ■

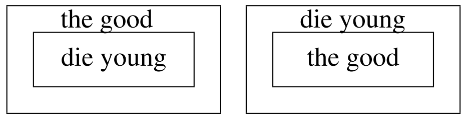
3.5 Exercises

1. The Barber of Seville is a man and he shaves exactly those men who do not shave themselves. Who shaves the Barber?

2. 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? (This is

not hypothetical. Tyrants usually seek to deprive others of their agency to do what they want. Do they have a right to do this if they want to?)

3. Only the good die young. It says so in a song. Which is the correct diagram to correspond to this statement? Sometimes such pictures are helpful.



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

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

and

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

5. Let $\mathscr{S}$ be a set of sets show $B \cup \cup\{A : A \in \mathscr{S}\} = \cup\{B \cup A : A \in \mathscr{S}\}$.

6. Let $\mathscr{S}$ be a set of sets show $B \cap \cup\{A : A \in \mathscr{S}\} = \cup\{B \cap A : A \in \mathscr{S}\}$.

7. Show the rational numbers are countable, this is in spite of the fact that between any two integers there are infinitely many rational numbers. What does this show about the usefulness of common sense and instinct in mathematics?

8. From Problem 7 the rational numbers can be listed as $\{r_i\}_{i=1}^{\infty}$. Let $j \in \mathbb{N}$. Show that

$$\mathbb{Q} = \cup_{i=1}^{\infty} \cap_{j=1}^{\infty} \left(r_i - \frac{1}{j}, r_i + \frac{1}{j}\right), \quad \mathbb{R} = \cap_{j=1}^{\infty} \cup_{i=1}^{\infty} \left(r_i - \frac{1}{j}, r_i + \frac{1}{j}\right)$$

Thus you can't switch intersections and unions in general.

9. 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 $\mathbb{N}$?

10. 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 observation that any such equation has at most n solutions, show the set of all algebraic numbers is countable.

11. Let A be a nonempty set and let $\mathscr{P}(A)$ be its power set, the set of all subsets of A. Show there does not exist any function f, which maps A onto $\mathscr{P}(A)$. 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.

12. 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?

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

14. Let $\mathscr{S}$ denote the set of all sequences which have either 0 or 1 in every entry. You have seen sequences in calculus. They will be discussed more formally later. Show that the set of all such sequences cannot be countable. **Hint:** Such a sequence can be thought of as an ordered list $a_1 a_2 a_3 \cdots$ where each a_i is either 0 or 1. Suppose you could list them all as follows.
$$\begin{aligned} \mathbf{a}_1 &= a_{11} a_{12} a_{13} \cdots \\ \mathbf{a}_2 &= a_{21} a_{22} a_{23} \cdots \\ \mathbf{a}_3 &= a_{31} a_{32} a_{33} \cdots \\ &\vdots \end{aligned}$$
Then consider the sequence $a_{11} a_{22} a_{33} \cdots$. Obtain a sequence which can't be in the list by considering the sequence $b_1 b_2 b_3 \cdots$ where b_k is obtained by changing a_{kk}. Explain why this sequence can't be any of the ones which are listed.

15. Show that the collection of sequences $a_1 a_2 \cdots a_n$ such that each a_k is either 0 or 1 such that $a_k = 0$ for all k larger than n is countable. Now show that the collection of sequences consisting of either 0 or 1 such that a_k is 0 for all k larger than **some** n is also countable. However, the set of all sequences of 0 and 1 is not countable.

16. Let $\mathscr{S}$ be the set of sequences of 0 or 1. Show there exists a mapping $\theta : [0,1] \to \mathscr{S}$ which is onto. Explain why this requires $[0,1]$ to be uncountable.

17. Prove Theorem 3.3.3, the theorem about partitioning using an equivalence relation into equivalence classes.

18. Let S be a set and consider a function f which maps $\mathscr{P}(S)$ to $\mathscr{P}(S)$ which satisfies the following. If $A \subseteq B$, then $f(A) \subseteq f(B)$. Then there exists A such that $f(A) = A$. **Hint:** You might consider the following subset of $\mathscr{P}(S)$.
$$\mathscr{C} \equiv \{B \in \mathscr{P}(S) : B \subseteq f(B)\}$$
Then consider $A \equiv \cup \mathscr{C}$. Argue A is the "largest" set in $\mathscr{C}$ which implies A cannot be a proper subset of $f(A)$. This is a case of the Tarski fixed point theorem. If X is a subset of $\mathscr{P}(S)$ such that if $\mathscr{F} \subseteq X$, then $\cup \mathscr{F} \in X$ and if f is increasing as above and $f(x) \in X$ for all $x \in X$, then the same result follows.

19. Another formulation of the Hausdorff maximal theorem is Zorn's lemma. This says that if you have a nonempty partially ordered set and every chain has an upper bound, then there exists a maximal element, one which has no element of the partial order which is larger. Show these two formulations are equivalent and each is equivalent to the axiom of choice.

20. A nonempty set X is well ordered if there exists an order $\leq$ which is a total order of the elements of X and in addition has the property that every nonempty subset of X has a smallest element. Zermelo showed that for every nonempty set X, there exists $\leq$ which makes X a well ordered set. To prove Zermelo's theorem, let $\mathscr{F} = \{S \subseteq X :$ there exists a well order for $S\}$. Let $S_1 \prec S_2$ if $S_1 \subseteq S_2$ and there exists a well order

for S_2, $\leq_2$which agrees with $\leq_1$ on S_1. Now use the Hausdorff maximal theorem. You need to show its union is all of X. If $X = \mathbb{R}$ this well order has NOTHING to do with the usual order on $\mathbb{R}$. Explain why.

Chapter 4

Functions and Sequences

4.1 General Considerations

As discussed earlier, the concept of a function is that of something which gives a unique output for a given input.

Definition 4.1.1 *Consider two sets, D and R along with a rule which assigns a unique element of R to every element of D. This rule is called a function and it is denoted by a letter such as f. The symbol, $D(f) = D$ is called the domain of f. The set R, also written $R(f)$, is called the range of f. The set of all elements of R which are of the form $f(x)$ for some $x \in D$ is often denoted by $f(D)$. When $R = f(D)$, the function f is said to be onto. It is common notation to write $f : D(f) \to R$ to denote the situation just described in this definition where f is a function defined on D having values in R.*

Example 4.1.2 *Consider the list of numbers, $\{1,2,3,4,5,6,7\} \equiv D$. Define a function which assigns an element of D to $R \equiv \{2,3,4,5,6,7,8\}$ by $f(x) \equiv x+1$ for each $x \in D$.*

In this example there was a clearly defined procedure which determined the function. However, sometimes there is no discernible procedure which yields a particular function.

Example 4.1.3 *Consider the ordered pairs, $(1,2),(2,-2),(8,3),(7,6)$ and let the domain be $D \equiv \{1,2,8,7\}$, the set of first entries in the given set of ordered pairs, $R \equiv \{2,-2,3,6\}$, the set of second entries, and let $f(1) = 2, f(2) = -2, f(8) = 3$, and $f(7) = 6$.*

Sometimes functions are not given in terms of a formula. For example, consider the following function defined on the positive real numbers having the following definition.

Example 4.1.4 *For $x \in \mathbb{R}$ define*

$$f(x) = \begin{cases} \frac{1}{n} \text{ if } x = \frac{m}{n} \text{ in lowest terms for } m,n \in \mathbb{Z} \\ 0 \text{ if } x \text{ is not rational} \end{cases} \tag{4.1}$$

This is a very interesting function called the Dirichlet function. Note that it is not defined in a simple way from a formula.

Example 4.1.5 *Let D consist of the set of people who have lived on the earth except for Adam and for $d \in D$, let $f(d) \equiv$ the biological father of d. Then f is a function.*

This function is not the sort of thing studied in calculus but it is a function just the same. When $D(f)$ is not specified and f is given by a formula, it is understood to consist of everything for which f makes sense. The following definition gives several ways to make new functions from old ones.

Definition 4.1.6 *Let f,g be functions with values in $\mathbb{F}$. Let a,b be points of $\mathbb{F}$. Then $af+bg$ is the name of a function whose domain is $D(f) \cap D(g)$ which is defined as $(af+bg)(x) = af(x)+bg(x)$. The function fg is the name of a function which is defined on $D(f) \cap D(g)$ given by $(fg)(x) = f(x)g(x)$. Similarly for k an integer, f^k is the name of a function defined as $f^k(x) = (f(x))^k$. The function f/g is the name of a function whose domain is $D(f) \cap \{x \in D(g) : g(x) \neq 0\}$ defined as $(f/g)(x) = f(x)/g(x)$. If $f : D(f) \to X$ and $g : D(g) \to Y$, then $g \circ f$ is the name of a function whose domain is $\{x \in D(f) : f(x) \in D(g)\}$ which is defined as $g \circ f(x) \equiv g(f(x))$. This is called the composition of the two functions.*

You should note that $f(x)$ is not a function. It is the value of the function at the point x. The name of the function is f. Nevertheless, people often write $f(x)$ to denote a function and it doesn't cause too many problems in beginning courses. When this is done, the variable x should be considered as a generic variable free to be anything in $D(f)$.

Sometimes people get hung up on formulas and think that the only functions of importance are those which are given by some simple formula. It is a mistake to think this way. Functions involve a domain and a range and a function is determined by what it does. This is an old idea. See Luke 6:44 where Jesus says that you know a tree by its fruit. See also Matt. 7 about how to recognize false prophets. You look at what it does to determine what it is. As it is with false prophets and trees, so it is with functions. [1] Although the idea is very old, its application to mathematics started with Dirichlet[2] in the early 1800's because he was concerned with piecewise continuous functions which would be given by different descriptions on different intervals. Before his time, they did tend to think of functions in terms of formulas.

Example 4.1.7 *Let $f(t) = t$ and $g(t) = 1+t$. Then $fg : \mathbb{R} \to \mathbb{R}$ is given by $fg(t) = t(1+t) = t+t^2$.*

Example 4.1.8 *Let $f(t) = 2t+1$ and $g(t) = \sqrt{1+t}$. Then*

$$g \circ f(t) = \sqrt{1+(2t+1)} = \sqrt{2t+2}$$

for $t \geq -1$. If $t < -1$ the inside of the square root sign is negative so makes no sense. Therefore, $g \circ f : \{t \in \mathbb{R} : t \geq -1\} \to \mathbb{R}$.

Note that in this last example, it was necessary to fuss about the domain of $g \circ f$ because g is only defined for certain values of t.

The concept of a one to one function is very important. This is discussed in the following definition.

Definition 4.1.9 *For any function $f : D(f) \subseteq X \to Y$, define the following set known as the inverse image of y.*

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

There may be many elements in this set, but when there is always only one element in this set for all $y \in f(D(f))$, the function f is one to one sometimes written, $1-1$. Thus f is one to one, $1-1$, if whenever $f(x) = f(x_1)$, then $x = x_1$. If f is one to one, the inverse function f^{-1} is defined on $f(D(f))$ and $f^{-1}(y) = x$ where $f(x) = y$. Thus from the definition, $f^{-1}(f(x)) = x$ for all $x \in D(f)$ and $f\left(f^{-1}(y)\right) = y$ for all $y \in f(D(f))$. Defining id *by* $\mathrm{id}(z) \equiv z$ *this says $f \circ f^{-1} = \mathrm{id}$ and $f^{-1} \circ f = \mathrm{id}$. Note that this is sloppy notation because the two* id *are totally different functions.*

[1] In many religions, including mine, epistemology based on knowledge of good and evil along with known facts, as urged by Jesus, is disparaged and replaced with other criteria like feelings, proof texts of scripture, traditions, sacrifice of believers, allegations of miracles, emotions, claimed authority, and social pressure.

[2] Peter Gustav Lejeune Dirichlet, 1805-1859 was a German mathematician who did fundamental work in analytic number theory. He also gave the first proof that Fourier series tend to converge to the mid-point of the jump of the function. He is a very important figure in the development of analysis in the nineteenth century. An interesting personal fact is that the great composer Felix Mendelsson was his brother in law.

Polynomials and rational functions are particularly easy functions to understand because they do come from a simple formula.

Definition 4.1.10 *A function f given by $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ is called a polynomial. Here the a_i are real or complex numbers and n is a nonnegative integer. In this case the degree of the polynomial $f(x)$ is n. Thus the degree of a polynomial is the largest exponent appearing on the variable.*

f is a rational function if $f(x) = \frac{h(x)}{g(x)}$ where h and g are polynomials.

For example, $f(x) = 3x^5 + 9x^2 + 7x + 5$ is a polynomial of degree 5 and $\frac{3x^5+9x^2+7x+5}{x^4+3x+x+1}$ is a rational function.

Note that in the case of a rational function, the domain of the function might not be all of $\mathbb{F}$. For example, if $f(x) = \frac{x^2+8}{x+1}$, the domain of f would be all complex numbers not equal to -1.

Closely related to the definition of a function is the concept of the graph of a function.

Definition 4.1.11 *Given two sets, X and Y, the Cartesian product of the two sets, written as $X \times Y$, is assumed to be a set described as follows.*

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

$\mathbb{F}^2$ denotes the Cartesian product of $\mathbb{F}$ with $\mathbb{F}$. Recall $\mathbb{F}$ could be either $\mathbb{R}$ or $\mathbb{C}$.

The notion of Cartesian product is just an abstraction of the concept of identifying a point in the plane with an ordered pair of numbers.

Definition 4.1.12 *Let $f : D(f) \to R(f)$ be a function. The graph of f consists of the set,*

$$\{(x, y) : y = f(x) \text{ for } x \in D(f)\}.$$

Note that knowledge of the graph of a function is equivalent to knowledge of the function. To find $f(x)$, simply observe the ordered pair which has x as its first element and the value of y equals $f(x)$.

4.2 Sequences

Functions defined on the set of integers larger than a given integer are called sequences.

Definition 4.2.1 *A function whose domain is defined as a set of the form*

$$\{k, k+1, k+2, \cdots\}$$

for k an integer is known as a sequence. Thus you can consider

$$f(k), f(k+1), f(k+2),$$

etc. Usually the domain of the sequence is either $\mathbb{N}$, the natural numbers consisting of $\{1, 2, 3, \cdots\}$ or the nonnegative integers, $\{0, 1, 2, 3, \cdots\}$. Also, it is traditional to write f_1, f_2, etc. instead of $f(1), f(2), f(3)$ etc. when referring to sequences. In the above context, f_k is called the first term, f_{k+1} the second and so forth. It is also common to write the sequence, not as f but as $\{f_i\}_{i=k}^{\infty}$ or just $\{f_i\}$ for short.

Example 4.2.2 *Let* $\{a_k\}_{k=1}^{\infty}$ *be defined by* $a_k \equiv k^2+1$.

This gives a sequence. In fact, $a_7 = a(7) = 7^2+1 = 50$ just from using the formula for the k^{th} term of the sequence.

It is nice when sequences come in this way from a formula for the k^{th} term. However, this is often not the case. Sometimes sequences are defined recursively. This happens, when the first several terms of the sequence are given and then a rule is specified which determines a_{n+1} from knowledge of $a_1, \cdots, a_n$. This rule which specifies a_{n+1} from knowledge of a_k for $k \leq n$ is known as a recurrence relation.

Example 4.2.3 *Let* $a_1 = 1$, $a_2 = 1$. *Assuming* $a_1, \cdots, a_{n+1}$ *are known,* $a_{n+2} \equiv a_n + a_{n+1}$.

Thus the first several terms of this sequence, listed in order, are 1, 1, 2, 3, 5, 8,$\cdots$. This particular sequence is called the Fibonacci sequence and is important in the study of reproducing rabbits. Note this defines a function without giving a formula for it. Such sequences occur naturally in the solution of differential equations using power series methods and in many other situations of great importance.

For sequences, it is very important to consider something called a subsequence.

Definition 4.2.4 *Let* $\{a_n\}$ *be a sequence and let* $n_1 < n_2 < n_3, \cdots$ *be any strictly increasing list of integers such that* n_1 *is at least as large as the first number in the domain of the function. Then if* $b_k \equiv a_{n_k}$, $\{b_k\}$ *is called a subsequence of* $\{a_n\}$. *Here* a_n *is in some given set.*

For example, suppose $a_n = (n^2+1)$. Thus $a_1 = 2$, $a_3 = 10$, etc. If $n_1 = 1, n_2 = 3, n_3 = 5, \cdots, n_k = 2k-1$, then letting $b_k = a_{n_k}$, it follows

$$b_k = \left((2k-1)^2+1\right) = 4k^2-4k+2.$$

However, you might not be able to describe a subsequence by a formula as I just did.

4.3 Exercises

1. Let $g(t) \equiv \sqrt{2-t}$ and let $f(t) = \frac{1}{t}$. Find $g \circ f$. Include the domain of $g \circ f$.

2. Give the domains of the following functions.

   (a) $f(x) = \frac{x+3}{3x-2}$

   (b) $f(x) = \sqrt{x^2-4}$

   (c) $f(x) = \sqrt{4-x^2}$

   (d) $f(x) = \sqrt{\frac{x-4}{3x+5}}$

   (e) $f(x) = \sqrt{\frac{x^2-4}{x+1}}$

3. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(t) \equiv t^3+1$. Is f one to one? Can you find a formula for f^{-1}?

4. Suppose $a_1 = 1, a_2 = 3$, and $a_3 = -1$. Suppose also that for $n \geq 4$ it is known that $a_n = a_{n-1} + 2a_{n-2} + 3a_{n-3}$. Find a_7. Are you able to guess a formula for the k^{th} term of this sequence?

5. Let $f : \{t \in \mathbb{R} : t \neq -1\} \to \mathbb{R}$ be defined by $f(t) \equiv \frac{t}{t+1}$. Find f^{-1} if possible.

6. A function $f : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function if whenever $x < y$, it follows that $f(x) < f(y)$. If f is a strictly increasing function, does f^{-1} always exist? Explain your answer.

7. Let $f(t)$ be defined by $f(t) = \begin{cases} 2t+1 \text{ if } t \leq 1 \\ t \text{ if } t > 1 \end{cases}$. Find f^{-1} if possible.

8. Suppose $f : D(f) \to R(f)$ is one to one, $R(f) \subseteq D(g)$, and $g : D(g) \to R(g)$ is one to one. Does it follow that $g \circ f$ is one to one?

9. If $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are two one to one functions, which of the following are necessarily one to one on their domains? Explain why or why not by giving a proof or an example.

   (a) $f + g$

   (b) fg

   (c) f^3

   (d) f/g

10. Draw the graph of the function $f(x) = x^3 + 1$.

11. Draw the graph of the function $f(x) = x^2 + 2x + 2$.

12. Draw the graph of the function $f(x) = \frac{x}{1+x}$.

13. Suppose $a_n = \frac{1}{n}$ and let $n_k = 2^k$. Find b_k where $b_k = a_{n_k}$.

14. If X_i are sets and for some j, $X_j = \emptyset$, the empty set. Verify carefully that $\prod_{i=1}^{n} X_i = \emptyset$.

15. Suppose $f(x) + f\left(\frac{1}{x}\right) = 7x$ and f is a function defined on $\mathbb{R} \setminus \{0\}$, the nonzero real numbers. Find all values of x where $f(x) = 1$ if there are any. Does there exist any such function?

16. Does there exist a function f, satisfying $f(x) - f\left(\frac{1}{x}\right) = 3x$ which has both x and $\frac{1}{x}$ in the domain of f?

17. In the situation of the Fibonacci sequence show that the formula for the n^{th} term can be found and is given by $a_n = \frac{\sqrt{5}}{5}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{\sqrt{5}}{5}\left(\frac{1-\sqrt{5}}{2}\right)^n$. **Hint:** You might be able to do this by induction but a better way would be to look for a solution to the recurrence relation, $a_{n+2} \equiv a_n + a_{n+1}$ of the form r^n. You will be able to show that there are two values of r which work, one of which is $r = \frac{1+\sqrt{5}}{2}$. Next you can observe that if r_1^n and r_2^n both satisfy the recurrence relation then so does $cr_1^n + dr_2^n$ for any choice of constants c, d. Then you try to pick c and d such that the conditions, $a_1 = 1$ and $a_2 = 1$ both hold.

18. In an ordinary annuity, you make constant payments, P at the beginning of each payment period. These accrue interest at the rate of r per payment period. This means at the start of the first payment period, there is the payment $P \equiv A_1$. Then this produces an amount rP in interest so at the beginning of the second payment period, you would have $rP + P + P \equiv A_2$. Thus $A_2 = A_1(1+r) + P$. Then at the beginning of the third payment period you would have $A_2(1+r) + P \equiv A_3$. Continuing in this way, you see that the amount in at the beginning of the n^{th} payment period would be A_n given by $A_n = A_{n-1}(1+r) + P$ and $A_1 = P$. Thus A is a function defined on

the positive integers given recursively as just described and A_n is the amount at the beginning of the n^{th} payment period. Now if you wanted to find out A_n for large n, how would you do it? One way would be to use the recurrance relation n times. A better way would be to find a formula for A_n. Look for one in the form $A_n = Cz^n + s$ where C, z and s are to be determined. Show that $C = \frac{P}{r}, z = (1+r)$, and $s = -\frac{P}{r}$.

19. A well known puzzle consists of three pegs and several disks each of a different diameter, each having a hole in the center which allows it to be slid down each of the pegs. These disks are piled one on top of the other on one of the pegs, in order of decreasing diameter, the larger disks always being below the smaller disks. The problem is to move the whole pile of disks to another peg such that you never place a disk on a smaller disk. If you have n disks, how many moves will it take? Of course this depends on n. If $n = 1$, you can do it in one move. If $n = 2$, you would need 3. Let A_n be the number required for n disks. Then in solving the puzzle, you must first obtain the top $n-1$ disks arranged in order on another peg before you can move the bottom disk of the original pile. This takes A_{n-1} moves. Explain why $A_n = 2A_{n-1} + 1, A_1 = 1$ and give a formula for A_n. Look for one in the form $A_n = Cr^n + s$. This puzzle is called the Tower of Hanoi. When you have found a formula for A_n, explain why it is not possible to do this puzzle if n is very large.

4.4 The Limit of a Sequence

The concept of the limit of a sequence was defined precisely by Bolzano.[3] The following is the precise definition of what is meant by the limit of a sequence. Our sequences will have values in $\mathbb{F}^p \equiv \left\{(x = (x_1, \cdots, x_p)) : x_i \in \mathbb{F}\right\}$.

Definition 4.4.1 *A sequence $\{a_n\}_{n=1}^{\infty}$ converges to a, written as*

$$\lim_{n\to\infty} a_n = a \text{ or } a_n \to a$$

if and only if for every $\varepsilon > 0$ there exists n_ε such that whenever $n \geq n_\varepsilon$, $\|a_n - a\| < \varepsilon$. Here a and a_n are assumed to be in $\mathbb{F}^p$ for some integer $p \geq 1$. Thus a_n is in the Cartesian product

[3] Bernhard Bolzano lived from 1781 to 1848. He had an Italian father but was born in Bohemia, and he wrote in German. He was a Catholic priest and held a position in philosophy at the University of Prague. It appears that Bolzano believed in the words of Jesus and did not hesitate to enthusiastically promote them. This got him in trouble with the political establishment of Austria who forced him out of the university and did not allow him to publish. He also displeased the Catholic hierarchy for being too rational.

Bolzano believed in absolute rigor in mathematics. He also was interested in physics, theology, and especially philosophy. His contributions in philosophy are very influential. He originated anti-psychologism also called logical objectivism which holds that logical truth exists independent of our opinions about it, contrary to the notion that truth can in any way depend on our feelings. This is the right way to regard truth in mathematics.

The intermediate value theorem from calculus is due to him. These days, this theorem is considered obvious and is not discussed well in calculus texts, but Bolzano knew better and gave a proof which identified exactly what was needed instead of relying on vague intuition.

Like many of the other mathematicians, he was concerned with the notion of infinitesimals which had been popularized by Leibniz. Some tried to strengthen this idea and others sought to get rid of it. They realized that something needed to be done about this fuzzy idea. Bolzano was one who contributed to removing it from calculus. He also proved the extreme value theorem in 1830's and gave the first formal $\varepsilon\delta$ description of continuity and limits.

This notion of infinitesimals did not completely vanish. These days, it is called non standard analysis. It can be made mathematically respectable but not in this book.

$\mathbb{F}\times\cdots\times\mathbb{F}$ *where* $\mathbb{F}$ *consists of real or complex numbers and, although other definitions are used,*

$$\|a\|\equiv\max\{|a_i|:i\leq p\}$$

for $a=(a_1,\cdots,a_p)\in\mathbb{F}^p$. *In this book, it is usually the case that* $p=1$, *but there is no difficulty in considering a more general case.*

Proposition 4.4.2 *The usual properties of absolute value hold for* $\|\cdot\|$ *with addition of the vectors, and multiplication by a scalar* α, *as presented in elementary calculus*

$$a+b\equiv(a_1+b_1,\cdots,a_p+b_p),\ \alpha a\equiv(\alpha a_1,\cdots,\alpha a_p)$$

Proof: From the triangle inequality for complex numbers,

$$\|a+b\|\equiv\max\{|a_i+b_i|,i\leq p\}\leq\max\{|a_i|,i\leq p\}+\max\{|b_i|,i\leq p\}=\|a\|+\|b\|. \tag{4.2}$$

Also, for $\alpha\in\mathbb{F}$, $\|\alpha a\|\equiv\max\{|\alpha a_i|:i\leq p\}=|\alpha|\max\{|a_i|:i\leq p\}=|\alpha|\,\|a\|$. By definition, $\|a\|\geq 0$ and is 0 if and only if $a_i=0$ for each i if and only if $a=0\equiv(0,\cdots,0)$. Also $\|a\|=\|a-b+b\|\leq\|a-b\|+\|b\|$ so $\|a\|-\|b\|\leq\|a-b\|$. Similarly $\|b\|-\|a\|\leq\|b-a\|=\|a-b\|$ and so $|\|a\|-\|b\||\leq\|a-b\|$. ■

In words the definition says that given any measure of closeness ε, the terms of the sequence are eventually this close to a. Here, the word "eventually" refers to n being sufficiently large. The above definition is always the definition of what is meant by the limit of a sequence.

Proposition 4.4.3 *Let* $a_n=\left(a_1^n,\cdots,a_p^n\right)$. *Then* $a_n\to a$ *if and only if for each* $i\leq p$, $a_i^n\to a_i$.

Proof: $\Rightarrow$ is obvious because $|a_i^n-a_i|\leq\|a_n-a\|$.
$\Leftarrow$ There exists n_i such that $\left|a_i^k-a_i\right|<\varepsilon$ whenever $k>n_i$. Let

$$N\geq\max\{n_i:i\leq p\}$$

Then for $n\geq N$, $\|a_n-a\|\equiv\max\{|a_i^n-a_i|,i\leq p\}<\varepsilon$. ■

Theorem 4.4.4 *If* $\lim_{n\to\infty}a_n=a$ *and* $\lim_{n\to\infty}a_n=\hat{a}$ *then* $\hat{a}=a$.

Proof: Suppose $\hat{a}\neq a$. Then let $0<\varepsilon<\|\hat{a}-a\|/2$ in the definition of the limit. It follows there exists n_ε such that if $n\geq n_\varepsilon$, then $\|a_n-a\|<\varepsilon$ and $\|a_n-\hat{a}\|<\varepsilon$. Therefore, for such n, $\|\hat{a}-a\|\leq\|\hat{a}-a_n\|+\|a_n-a\|<\varepsilon+\varepsilon<\|\hat{a}-a\|/2+\|\hat{a}-a\|/2=\|\hat{a}-a\|$, a contradiction. ■

Example 4.4.5 *Let* $a_n=\left(\frac{1}{n^2+1},\frac{i}{n}\right)\in\mathbb{F}^2$.

Then it seems clear that $\lim_{n\to\infty}a_n=(0,0)$. In fact, this is true from the definition. Let $\varepsilon>0$ be given. Let $n_\varepsilon\geq\max\left(\sqrt{\varepsilon^{-1}},\frac{1}{\varepsilon}\right)$. Then if $n>n_\varepsilon\geq\sqrt{\varepsilon^{-1}}$, it follows that $n^2+1>\varepsilon^{-1}$ and so $0<\frac{1}{n^2+1}=a_n<\varepsilon$ and also $n\geq 1/\varepsilon$ so $1/n<\varepsilon$. Thus $\|a_n-(0,0)\|\equiv\max\left(\left|\frac{1}{n^2+1}-0\right|,\left|\frac{i}{n}-0\right|\right)<\varepsilon$ whenever n is this large.

Note the definition was of no use in finding a candidate for the limit. This had to be produced based on other considerations. The definition is for verifying beyond any doubt that something is the limit. It is also what must be referred to in establishing theorems which are good for finding limits.

Example 4.4.6 *Let $a_n = n^2$.*

Then in this case $\lim_{n\to\infty} a_n$ does not exist. This is so because a_n cannot be eventually close to anything.

Example 4.4.7 *Let $a_n = (-1)^n$.*

In this case, $\lim_{n\to\infty} (-1)^n$ does not exist. This follows from the definition. Let $\varepsilon = 1/2$. If there exists a limit, l, then eventually, for all n large enough, $|a_n - l| < 1/2$. However, $|a_n - a_{n+1}| = 2$ and so, $2 = |a_n - a_{n+1}| \leq |a_n - l| + |l - a_{n+1}| < 1/2 + 1/2 = 1$ which cannot hold. Therefore, there can be no limit for this sequence.

Theorem 4.4.8 *Suppose $\{a_n\}$ and $\{b_n\}$ are sequences, $a_n \in \mathbb{F}^p, b_n \in \mathbb{F}^p$ and that*

$$\lim_{n\to\infty} a_n = a \text{ and } \lim_{n\to\infty} b_n = b.$$

Also suppose x and y are in $\mathbb{F}$. Then

$$\lim_{n\to\infty} xa_n + yb_n = xa + yb \tag{4.3}$$

If $a_n \in \mathbb{F}, b_n \in \mathbb{F}^p$,

$$\lim_{n\to\infty} a_n b_n = ab \tag{4.4}$$

$$\lim_{n\to\infty} a_n^q = a^q \tag{4.5}$$

If $b \neq 0$ and $b \in \mathbb{F}$ and $a_n \in \mathbb{F}^p$,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}. \tag{4.6}$$

Proof: The first of these claims is left for you to do. To do the second, let $\varepsilon > 0$ be given and choose n_1 such that if $n \geq n_1$ then $|a_n - a| < 1$. Then for such n, the triangle inequality implies

$$\begin{aligned} \|a_n b_n - ab\| &\leq \|a_n b_n - a_n b\| + \|a_n b - ab\| \leq |a_n| \|b_n - b\| + \|b\| |a_n - a| \\ &\leq (|a| + 1) \|b_n - b\| + \|b\| |a_n - a| . \end{aligned}$$

Now let n_2 be large enough that for $n \geq n_2$,

$$\|b_n - b\| < \frac{\varepsilon}{2(|a|+1)}, \text{ and } |a_n - a| < \frac{\varepsilon}{2(\|b\|+1)}.$$

Such a number n_2 exists because of the definition of limit. Therefore, let $n_\varepsilon > \max(n_1, n_2)$. For $n \geq n_\varepsilon$,

$$\begin{aligned} \|a_n b_n - ab\| &\leq (|a| + 1) \|b_n - b\| + \|b\| |a_n - a| \\ &< (|a| + 1) \frac{\varepsilon}{2(|a|+1)} + \|b\| \frac{\varepsilon}{2(|b|+1)} \leq \varepsilon. \end{aligned}$$

This proves 4.4. Then 4.5 follows from this by induction in the above case where $b_n \in \mathbb{F}$.

Next consider 4.6. Let $\varepsilon > 0$ be given and let n_1 be so large that if $n \geq n_1$, $|b_n - b| < \frac{|b|}{2}$. Thus for such n,

$$\left\| \frac{a_n}{b_n} - \frac{a}{b} \right\| = \left\| \frac{a_n b - a b_n}{b b_n} \right\| \leq \frac{2}{|b|^2} \left[\|a_n b - ab\| + \|ab - ab_n\| \right]$$

$$\leq \frac{2}{|b|}\|a_n - a\| + \frac{2\|a\|}{|b|^2}\,|||b_n - b|||\,.$$

Now choose n_2 so large that if $n \geq n_2$, then

$$\|a_n - a\| < \frac{\varepsilon |b|}{4}, \text{ and } |b_n - b| < \frac{\varepsilon |b|^2}{4(\|a\|+1)}.$$

Letting $n_\varepsilon > \max(n_1, n_2)$, it follows that for $n \geq n_\varepsilon$,

$$\left\|\frac{a_n}{b_n} - \frac{a}{b}\right\| \leq \frac{2}{|b|}\|a_n - a\| + \frac{2\|a\|}{|b|^2}|b_n - b| < \frac{2}{|b|}\frac{\varepsilon |b|}{4} + \frac{2\|a\|}{|b|^2}\frac{\varepsilon |b|^2}{4(\|a\|+1)} < \varepsilon. \blacksquare$$

Another very useful theorem for finding limits is the squeezing theorem.

Theorem 4.4.9 *In case $a_n, b_n \in \mathbb{R}$, suppose $\lim_{n\to\infty} a_n = a = \lim_{n\to\infty} b_n$ and $a_n \leq c_n \leq b_n$ for all n large enough. Then $\lim_{n\to\infty} c_n = a$.*

Proof: Let $\varepsilon > 0$ be given and let n_1 be large enough that if $n \geq n_1$, $|a_n - a| < \varepsilon/2$ and $|b_n - a| < \varepsilon/2$. Then for such n,$|c_n - a| \leq |a_n - a| + |b_n - a| < \varepsilon$. The reason for this is that if $c_n \geq a$, then $|c_n - a| = c_n - a \leq b_n - a \leq |a_n - a| + |b_n - a|$ because $b_n \geq c_n$. On the other hand, if $c_n \leq a$, then

$$|c_n - a| = a - c_n \leq a - a_n \leq |a - a_n| + |b - b_n|. \blacksquare$$

As an example, consider the following.

Example 4.4.10 *Let $c_n \equiv (-1)^n \frac{1}{n}$ and let $b_n = \frac{1}{n}$, and $a_n = -\frac{1}{n}$. Then you may easily show that $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$. Since $a_n \leq c_n \leq b_n$, it follows $\lim_{n\to\infty} c_n = 0$ also.*

Theorem 4.4.11 $\lim_{n\to\infty} r^n = 0$. *Whenever $|r| < 1$. Here $r \in \mathbb{F}$.*

Proof: If $0 < r < 1$ if follows $r^{-1} > 1$. Why? Letting $\alpha = \frac{1}{r} - 1$, it follows $r = \frac{1}{1+\alpha}$. Therefore, by the binomial theorem, $0 < r^n = \frac{1}{(1+\alpha)^n} \leq \frac{1}{1+\alpha n}$. Therefore, $\lim_{n\to\infty} r^n = 0$ if $0 < r < 1$. In general, if $|r| < 1$, $|r^n| = |r|^n \to 0$ by the first part. ■

An important theorem is the one which states that if a sequence converges, so does every subsequence. You should review Definition 4.2.4 on Page 58 at this point.

Theorem 4.4.12 *Let $\{x_n\}$ be a sequence with $\lim_{n\to\infty} x_n = x$ and let $\{x_{n_k}\}$ be a subsequence. Then $\lim_{k\to\infty} x_{n_k} = x$.*

Proof: Let $\varepsilon > 0$ be given. Then there exists n_ε such that if $n > n_\varepsilon$, then $\|x_n - x\| < \varepsilon$. Suppose $k > n_\varepsilon$. Then $n_k \geq k > n_\varepsilon$ and so $\|x_{n_k} - x\| < \varepsilon$ showing $\lim_{k\to\infty} x_{n_k} = x$ as claimed. ■

Theorem 4.4.13 *Let $\{x_n\}$ be a sequence of real numbers and suppose each $x_n \leq l$ $(\geq l)$ for all n large enough, and $\lim_{n\to\infty} x_n = x$. Then $x \leq l\,(\geq l)$. More generally, suppose $\{x_n\}$ and $\{y_n\}$ are two sequences of real numbers such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$. Then if $x_n \leq y_n$ for all n sufficiently large, then $x \leq y$.*

Proof: I will show the second claim because it includes the first as a special case. Letting $\varepsilon > 0$ be given, for all n large enough, $|y - y_n| < \varepsilon$ so $y \geq y_n - \varepsilon$. Similarly, for n large enough, $x \leq x_n + \varepsilon$. Therefore,

$$y - x \geq y_n - \varepsilon - (x_n + \varepsilon) \geq (y_n - x_n) - 2\varepsilon \geq -2\varepsilon$$

Since ε is arbitrary, it follows that $y - x \geq 0$. ■

Another important observation is that if a sequence converges, then it must be bounded.

Proposition 4.4.14 *Suppose $x_n \to x$. Then $\|x_n\|$ is bounded by some $M < \infty$.*

Proof: There exists N such that if $n \geq N$, then $\|x - x_n\| < 1$. It follows from the triangle inequality, see Proposition 4.4.2, that for $n \geq N$, $\|x_n\| \leq 1 + \|x\|$. There are only finitely many x_k for $k < N$ and so for all k,

$$\|x_k\| \leq \max\{1 + \|x\|, \|x_k\| : k \leq N\} \equiv M < \infty. \blacksquare$$

4.5 Cauchy Sequences

A Cauchy sequence is one which "bunches up". This concept was developed by Bolzano and Cauchy. It is a fundamental idea in analysis.

Definition 4.5.1 *$\{a_n\}$ is a Cauchy sequence if for all $\varepsilon > 0$, there exists n_ε such that whenever $n, m \geq n_\varepsilon$, $|a_n - a_m| < \varepsilon$.*

A sequence is Cauchy means the terms are "bunching up to each other" as m, n get large.

Theorem 4.5.2 *The set of terms (values) of a Cauchy sequence in $\mathbb{F}^p$ is bounded.*

Proof: Let $\varepsilon = 1$ in the definition of a Cauchy sequence and let $n > n_1$. Then from the definition, $\|a_n - a_{n_1}\| < 1$. It follows from the triangle inequality that for all $n > n_1$, $\|a_n\| < 1 + \|a_{n_1}\|$ .Therefore, for all n, $\|a_n\| \leq 1 + \|a_{n_1}\| + \sum_{k=1}^{n_1} \|a_k\|$. ■

Theorem 4.5.3 *If a sequence $\{a_n\}$ in $\mathbb{F}^p$ converges, then the sequence is a Cauchy sequence.*

Proof: Let $\varepsilon > 0$ be given and suppose $a_n \to a$. Then from the definition of convergence, there exists n_ε such that if $n > n_\varepsilon$, it follows that $\|a_n - a\| < \frac{\varepsilon}{2}$. Therefore, if $m, n \geq n_\varepsilon + 1$, it follows that $\|a_n - a_m\| \leq \|a_n - a\| + \|a - a_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ showing that, since $\varepsilon > 0$ is arbitrary, $\{a_n\}$ is a Cauchy sequence. ■

The following theorem is very useful and is likely the most important property of Cauchy sequences. You know that if a sequence converges, then every subsequence converges to the same thing. However, you can have a sequence which does not converge, $a_n = (-1)^n$ for example which has a convergent subsequence, $n_k = 2k$ in this example. This won't happen with a Cauchy sequence.

Theorem 4.5.4 *Suppose $\{a_n\}$ is a Cauchy sequence in $\mathbb{F}^p$ and there exists a subsequence, $\{a_{n_k}\}$ which converges to a. Then $\{a_n\}$ also converges to a.*

Proof: Let $\varepsilon > 0$ be given. There exists N such that if $m, n > N$, then $\|a_m - a_n\| < \varepsilon/2$. Also there exists K such that if $k > K$, then $\|a - a_{n_k}\| < \varepsilon/2$. Then let $k > \max(K, N)$. Then for such k, $\|a_k - a\| \leq \|a_k - a_{n_k}\| + \|a_{n_k} - a\| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. ■

This theorem holds in all instances where it makes sense to speak of Cauchy sequences.

4.6 The Nested Interval Lemma

In Russia there is a kind of doll called a matrushka doll. You pick it up and notice it comes apart in the center. Separating the two halves you find an identical doll inside. Then you notice this inside doll also comes apart in the center. Separating the two halves, you find yet another identical doll inside. This goes on quite a while until the final doll is in one piece. The nested interval lemma is like a matrushka doll except the process never stops. It involves a sequence of intervals, the first containing the second, the second containing the third, the third containing the fourth and so on. The fundamental question is whether there exists a point in all the intervals. Sometimes there is such a point and this comes from completeness.

Lemma 4.6.1 *Let $I_k = \left[a^k, b^k\right]$and suppose that for all $k = 1,2,\cdots$, $I_k \supseteq I_{k+1}$. Then there exists a point, $c \in \mathbb{R}$ which is an element of every I_k. If the diameters (length) of these intervals, denoted as $\operatorname{diam}(I_k)$ converges to 0, then there is a unique point in the intersection of all these intervals.*

Proof: Since $I_k \supseteq I_{k+1}$, this implies

$$a^k \leq a^{k+1},\ b^k \geq b^{k+1}. \tag{4.7}$$

Consequently, if $k \leq l$,

$$a^l \leq a^l \leq b^l \leq b^k. \tag{4.8}$$

Now define $c \equiv \sup\left\{a^l : l = 1,2,\cdots\right\}$. By the first inequality in 4.7, and 4.8

$$a^k \leq c = \sup\left\{a^l : l = k, k+1,\cdots\right\} \leq b^k \tag{4.9}$$

for each $k = 1,2\cdots$. Thus $c \in I_k$ for every k and this proves the lemma. The reason for the last inequality in 4.9 is that from 4.8, b^k is an upper bound to $\left\{a^l : l = k, k+1,\cdots\right\}$. Therefore, it is at least as large as the least upper bound.

For the last claim, suppose there are two points x, y in the intersection. Then $|x-y| = r > 0$ but eventually the diameter of I_k is less than r. Thus it cannot contain both x and y. ■

This is really quite a remarkable result and may not seem so obvious. Consider the intervals $I_k \equiv (0, 1/k)$. Then there is no point which lies in all these intervals because no negative number can be in all the intervals and $1/k$ is smaller than a given positive number whenever k is large enough. Thus the only candidate for being in all the intervals is 0 and 0 has been left out of them all. The problem here is that the endpoints of the intervals were not included, contrary to the hypotheses of the above lemma in which all the intervals included the endpoints.

Corollary 4.6.2 *Let $R_n \equiv \prod_{k=1}^p \left[a_k^n, b_k^n\right]$ where $R_{n+1} \subseteq R_n$. Then $\cap_{n=1}^\infty R_n \neq \emptyset$. If the diameter of R_n defined as $\max\left\{b_k^n - a_k^n : k \leq p\right\}$ converges to 0, then there is exactly one point in this intersection.*

Proof: Since these rectangles R_k are nested, $\left[a_k^n, b_k^n\right] \supseteq \left[a_k^{n+1}, b_k^{n+1}\right]$ and so there exists $x_k \in \cap_n \left[a_k^n, b_k^n\right]$. Then $x \equiv (x_1, \cdots, x_p) \in \cap_n R_n$. In case the diameter of R_n converges to 0, if $x, y \in \cap R_n$, then $\|x-y\| \leq \max\left\{b_k^n - a_k^n, k \leq p\right\}$ and this converges to 0 as $n \to \infty$. Thus $x = y$. ■

4.7 Exercises

1. Find $\lim_{n\to\infty} \frac{n}{3n+4}$.

2. Find $\lim_{n\to\infty} \frac{3n^4+7n+1000}{n^4+1}$.

3. Find $\lim_{n\to\infty} \frac{2^n+7(5^n)}{4^n+2(5^n)}$.

4. Find $\lim_{n\to\infty} \sqrt{(n^2+6n)}-n$. **Hint:** Multiply and divide by $\sqrt{(n^2+6n)}+n$.

5. Find $\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{10^k}$.

6. Suppose $\{x_n+iy_n\}$ is a sequence of complex numbers which converges to the complex number $x+iy$. Show this happens if and only if $x_n \to x$ and $y_n \to y$.

7. For $|r|<1$, find $\lim_{n\to\infty} \sum_{k=0}^{n} r^k$. **Hint:** First show $\sum_{k=0}^{n} r^k = \frac{r^{n+1}}{r-1} - \frac{1}{r-1}$. Then recall Theorem 4.4.11.

8. Using the binomial theorem prove that for all $n \in \mathbb{N}$,
$$\left(1+\frac{1}{n}\right)^n \leq \left(1+\frac{1}{n+1}\right)^{n+1}.$$
Hint: Show first that $\binom{n}{k} = \frac{n\cdot(n-1)\cdots(n-k+1)}{k!}$. By the binomial theorem,
$$\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{n}\right)^k = \sum_{k=0}^{n} \frac{\overbrace{n\cdot(n-1)\cdots(n-k+1)}^{k \text{ factors}}}{k!n^k}.$$
Now consider the term $\frac{n\cdot(n-1)\cdots(n-k+1)}{k!n^k}$ and note that a similar term occurs in the binomial expansion for $\left(1+\frac{1}{n+1}\right)^{n+1}$ except you replace n with $n+1$ whereever this occurs. Argue the term got bigger and then note that in the binomial expansion for $\left(1+\frac{1}{n+1}\right)^{n+1}$, there are more terms.

9. Prove by induction that for all $k \geq 4$, $2^k \leq k!$

10. Use the Problems 21 and 8 to verify for all $n \in \mathbb{N}$, $\left(1+\frac{1}{n}\right)^n \leq 3$.

11. Prove $\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n$ exists and equals a number less than 3.

12. Using Problem 10, prove $n^{n+1} \geq (n+1)^n$ for all integers, $n \geq 3$.

13. Find $\lim_{n\to\infty} n\sin n$ if it exists. If it does not exist, explain why it does not.

14. Recall the axiom of completeness states that a set which is bounded above has a least upper bound and a set which is bounded below has a greatest lower bound. Show that a monotone decreasing sequence which is bounded below converges to its greatest lower bound. **Hint:** Let a denote the greatest lower bound and recall that because of this, it follows that for all $\varepsilon > 0$ there exist points of $\{a_n\}$ in $[a, a+\varepsilon]$.

15. Let $A_n = \sum_{k=2}^{n} \frac{1}{k(k-1)}$ for $n \geq 2$. Show $\lim_{n\to\infty} A_n$ exists and find the limit. **Hint:** Show there exists an upper bound to the A_n as follows.
$$\sum_{k=2}^{n} \frac{1}{k(k-1)} = \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k} \right) = 1 - \frac{1}{n} \leq 1.$$

16. Let $H_n = \sum_{k=1}^{n} \frac{1}{k^2}$ for $n \geq 2$. Show $\lim_{n\to\infty} H_n$ exists. **Hint:** Use the above problem to obtain the existence of an upper bound.

17. Let $I_n = (-1/n, 1/n)$ and let $J_n = (0, 2/n)$. The intervals, I_n and J_n are open intervals of length $2/n$. Find $\cap_{n=1}^{\infty} I_n$ and $\cap_{n=1}^{\infty} J_n$. Repeat the same problem for $I_n = (-1/n, 1/n]$ and $J_n = [0, 2/n)$.

18. Show the set of real numbers $[0,1]$ is not countable. That is, show that there can be no mapping from $\mathbb{N}$ onto $[0,1]$. **Hint:** Show that every sequence, the terms consisting only of 0 or 1 determines a unique point of $[0,1]$. Call this map γ. Show it is onto. Also show that there is a map from $[0,1]$ onto $\mathscr{S}$, the set of sequences of zeros and ones. This will involve the nested interval lemma. Thus there is a one to one and onto map α from $\mathscr{S}$ to $[0,1]$ by Corollary 3.2.5. Next show that there is a one to one and onto map from this set of sequences and $\mathscr{P}(\mathbb{N})$. Consider $\theta(\{a_n\}_{n=1}^{\infty}) = \{n : a_n = 1\}$. Now suppose that $f : \mathbb{N} \to [0,1]$ is onto. Then $\theta \circ \alpha^{-1} \circ f$ is onto $\mathscr{P}(\mathbb{N})$. Recall that there is no map from a set to its power set. Review why this is.

19. Show that if I and J are any two closed intervals, then there is a one to one and onto map from I to J. Thus from the above problem, no closed interval, however short can be countable.

4.8 Compactness

4.8.1 Sequential Compactness

First I will discuss the very important concept of sequential compactness. This is a property that some sets have. A set of numbers is sequentially compact if every sequence contained in the set has a subsequence which converges to a point **in the set**. It is unbelievably useful whenever you try to understand existence theorems.

Definition 4.8.1 *A set, $K \subseteq \mathbb{F}^p$ is sequentially compact if whenever $\{a_n\} \subseteq K$ is a sequence, there exists a subsequence, $\{a_{n_k}\}$ such that this subsequence converges to a point of K.*

The following theorem is part of the Heine Borel theorem.

Theorem 4.8.2 *Every closed interval $[a,b]$ is sequentially compact.*

Proof: Let $\{x_n\} \subseteq [a,b] \equiv I_0$. Consider the two intervals $\left[a, \frac{a+b}{2}\right]$ and $\left[\frac{a+b}{2}, b\right]$ each of which has length $(b-a)/2$. At least one of these intervals contains x_n for infinitely many values of n. Call this interval I_1. Now do for I_1 what was done for I_0. Split it in half and let I_2 be the interval which contains x_n for infinitely many values of n. Continue this way obtaining a sequence of nested intervals $I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \cdots$ where the length of I_n is $(b-a)/2^n$. Now pick n_1 such that $x_{n_1} \in I_1$, n_2 such that $n_2 > n_1$ and $x_{n_2} \in I_2, n_3$ such that

$n_3 > n_2$ and $x_{n_3} \in I_3$, etc. (This can be done because in each case the intervals contained x_n for infinitely many values of n.) By the nested interval lemma there exists a point c contained in all these intervals. Furthermore, $\left|x_{n_k} - c\right| < (b-a)\, 2^{-k}$ and so $\lim_{k\to\infty} x_{n_k} = c \in [a,b]$. ■

Corollary 4.8.3 *$R \equiv \prod_{k=1}^{p} [a_k, b_k]$ is sequentially compact in $\mathbb{R}^p$.*

Proof: Let $\{x_n\}_{n=1}^{\infty} \subseteq R$, $x_n = \left(x_1^n, \cdots, x_p^n\right)$. Then there is a subsequence, still denoted with n such that $\{x_1^n\}$ converges to a point $x_1 \in [a_1, b_1]$. Now there exists a further subsequence, still denoted with n such that x_2^n converges to $x_2 \in [a_2, b_2]$. Continuing to take subsequence, there is a subsequence, still denoted with n such that $x_k^n \to x_k \in [a_k, b_k]$ for each k. By Proposition 4.4.3, this shows that this subsequence converges to $x = (x_1, \cdots, x_p)$. ■

4.8.2 Closed and Open Sets

I have been using the terminology $[a,b]$ is a closed interval to mean it is an interval which contains the two endpoints. However, there is a more general notion of what it means to be closed. Similarly there is a general notion of what it means to be open.

Definition 4.8.4 *Let U be a set of points in $\mathbb{F}^p$. A point $p \in U$ is said to be an interior point if whenever $\|x - p\|$ is sufficiently small, it follows $x \in U$ also. The set of points, x which are closer to p than δ is denoted by*

$$B(p, \delta) \equiv \{x \in \mathbb{F} : \|x - p\| < \delta\}.$$

*This symbol, $B(p,\delta)$ is called an open ball of radius δ. Thus a point p is an interior point of U if there exists $\delta > 0$ such that $p \in B(p,\delta) \subseteq U$. An **open set** is one for which every point of the set is an interior point. **Closed sets** are those which are complements of open sets. Thus H is closed means H^C is open.*

Note the following:

Proposition 4.8.5 *If $U = H^C$ where H is closed, then U is open. Also $\emptyset$ and $\mathbb{F}^p$ are both open and closed.*

Proof: Note that $\mathbb{F}^p$ is open obviously. Also $\emptyset$ is obviously open because every point of $\emptyset$ is an interior point. Indeed, it has none so they all must be interior points. Therefore, $\mathbb{F}^p$ is also closed because it is the complement of an open set. Now $H = U^C$ and so, given that H is closed, then by definition, it must be the complement of an open set, but it is the complement of U and so U must be open. It follows that $\emptyset$ is open because it is the complement of a closed set $\mathbb{F}^p$. ■

Thus open sets are complements of closed sets and closed sets are complements of open sets. I will use this fact without comment whenever convenient.

What is an example of an open set? The simplest example is an open ball.

Proposition 4.8.6 *$B(p,\delta)$ is an open set and $D(p,r) \equiv \{x : \|x - p\| \leq r\}$ is a closed set.*

Proof: It is necessary to show every point of $B(p,\delta)$ is an interior point. Let $x \in B(p,\delta)$. Then let $r = \delta - \|x-p\|$. It follows $r > 0$ because it is given that $\|x-p\| < \delta$. Now consider $z \in B(x,r)$. From Proposition 4.4.2, the triangle inequality,

$$\|z-p\| \leq \|z-x\| + \|x-p\| < r + \|x-p\| = \delta - \|x-p\| + \|x-p\| = \delta$$

and so $z \in B(p,\delta)$. That is $B(x,r) \subseteq B(p,\delta)$. Since x was arbitrary, this has shown every point of the ball is an interior point. Thus the ball is an open set.

Consider the last assertion. If $y \notin D(p,r)$, then $\|y-p\| > r$ and you could consider $B(y, \|y-p\| - r)$. If $z \in B(y, \|y-p\| - r)$, then

$$\begin{aligned} \|z-p\| &= \|z-y+y-p\| \geq \|y-p\| - \|z-y\| \\ &> \|y-p\| - (\|y-p\| - r) = r \end{aligned}$$

and so $z \notin D(p,r)$ which shows that the complement of $D(p,r)$ is open so this set is closed. ■

Definition 4.8.7 *Let A be any nonempty set and let x be a point. Then x is said to be a limit point of A if for every $r > 0, B(x,r)$ contains a point of A which is not equal to x.*

The following proposition is fairly obvious from the above definition and will be used whenever convenient. It is equivalent to the above definition and so it can take the place of the above definition if desired.

Proposition 4.8.8 *A point x is a limit point of the nonempty set $A \subseteq \mathbb{F}^p$ if and only if every $B(x,r)$ contains infinitely many points of A, none of which are equal to x. In other words, there exists a sequence of distinct points of A none equal to x which converges to x.*

Proof: $\Leftarrow$ is obvious. Consider $\Rightarrow$. Let x be a limit point. Let $r_1 = 1$. Then $B(x,r_1)$ contains $a_1 \neq x$. If $\{a_1, \cdots, a_n\}$ have been chosen none equal to x and with no repeats in the list, let $0 < r_n < \min\left(\frac{1}{n}, \min\{\|a_i - x\|, i = 1,2,\cdots n\}\right)$. Then let $a_{n+1} \in B(x,r_n) \setminus \{x\}$. Thus a_{n+1} is not equal to any of the earlier a_k and every $B(x,r)$ contains $B(x,r_n)$ for all n large enough and hence it contains a_k for $k \geq n$ where the a_k are distinct, none equal to x. ■

Example 4.8.9 *Consider $A = \mathbb{N}$, the positive integers. Then none of the points of A is a limit point of A because if $n \in A, B(n, 1/10)$ contains no points of $\mathbb{N}$ which are not equal to n.*

Example 4.8.10 *Consider $A = (a,b)$, an open interval in $\mathbb{R}$. If $x \in (a,b)$, let*

$$r = \min(|x-a|, |x-b|).$$

Then $B(x,r) \subseteq A$ because if $|y-x| < r$, then

$$\begin{aligned} y-a &= y-x+x-a \geq x-a-|y-x| \\ &= |x-a| - |y-x| > |x-a| - r \geq 0 \end{aligned}$$

showing $y > a$. A similar argument which you should provide shows $y < b$. Thus $y \in (a,b)$ and x is an interior point. Since x was arbitrary, this shows every point of (a,b) is an interior point and so (a,b) is open.

Theorem 4.8.11 *The following are equivalent.*

1. *A is closed*
2. *If* $\{a_n\}_{n=1}^{\infty}$ *is a sequence of points of A and* $\lim_{n\to\infty} a_n = a$, *then* $a \in A$.
3. *A contains all of its limit points.*

If a is a limit point, then there is a sequence of distinct points of A none of which equal a which converges to a.

Proof: 1.$\Longleftrightarrow$2. Say A is closed and $a_n \to a$ where each $a_n \in A$. If $a \notin A$, then there exists $\varepsilon > 0$ such that $B(a,\varepsilon) \cap A = \emptyset$. But then a_n fails to converge to a so $a \in A$ after all. Conversely, if 2. holds and $x \notin A$, $B\left(x, \frac{1}{n}\right)$ must fail to contain any points of A for some $n \in \mathbb{N}$ because if not, you could pick $a_n \in B\left(x, \frac{1}{n}\right) \cap A$ and obtain $\lim_{n\to\infty} a_n = x$ which would give $x \in A$ by 2. Thus A^C is open and A is closed.

2. $\Rightarrow$ 3. Say a is a limit point of A. Then by Proposition 4.8.8 there is a sequence of distinct points of A $\{a_n\}$ with $a_n \to a$. By 2., $a \in A$.

3. $\Rightarrow$ 1. Given 3., why is A^C open? Let $x \in A^C$. By 3. x cannot be a limit point. Hence there exists $B(x,r)$ which contains at most finitely many points of A. Since $x \in A^C$, none of these are equal to x. Hence, making r still smaller, one can avoid all of these points. Thus the modified r has the property that $B(x,r)$ contains no points of A and so A is closed because its complement is open. The last claim is from Proposition 4.8.8. ■

Note that part of this theorem says that a set A having all its limit points is the same as saying that whenever a sequence of points of A converges to a point a, then it follows $a \in A$. In other words, closed is the same as being closed with respect to containing all limits of sequences of points of A.

Corollary 4.8.12 *Let A be a nonempty set and denote by* A' *the set of limit points of A. Then* $A \cup A'$ *is a closed set and it is the smallest closed set containing A. In fact,* $A \cup A' = \cap\{C : C \text{ is closed and } C \supseteq A\}$. *This set* $A \cup A'$ *is denoted as* $\overline{A}$.

Proof: Is it the case that $(A \cup A')^C$ is open? This is what needs to be shown if the given set is closed. Let $p \notin A \cup A'$. Then since p is neither in A nor a limit point of A, there exists $B(p,r)$ such that $B(p,r) \cap A = \emptyset$. Therefore, $B(p,r) \cap A' = \emptyset$ also. This is because if $z \in B(p,r) \cap A'$, then

$$B(z, r - \|p - z\|) \subseteq B(p,r)$$

and this smaller ball contains points of A since z is a limit point. This contradiction shows that $B(p,r) \cap A' = \emptyset$ as claimed. Hence $(A \cup A')^C$ is open because p was an arbitrary point of $(A \cup A')^C$. Hence $A \cup A'$ is closed as claimed.

It was just shown that $A \cup A' \supseteq \cap\{C : C \supseteq A\}$. Now suppose $C \supseteq A$ and C is closed. Then if p is a limit point of A, it follows from Theorem 4.8.11 that there exists a sequence of distinct points of A converging to p. Since C is closed, and these points of A are all in C, it follows that $p \in C$. Hence $C \supseteq A \cup A'$.Thus $A \cup A' \supseteq \cap\{C : C \supseteq A\} \supseteq A \cup A'$. ■

Theorem 4.8.13 *A set* $K \neq \emptyset$ *in* $\mathbb{R}^p$ *is sequentially compact if and only if it is closed and bounded. A set is bounded means it is contained in some ball having finite radius. If K is sequentially compact and if H is a closed subset of K then H is sequentially compact.*

Proof: $\Rightarrow$ Suppose K is sequentially compact. Why is it closed? Let $k_n \to k$ where each $k_n \in K$. Why is $k \in K$? Since K is sequentially compact, there is a subsequence $\{k_{n_j}\}$ such that $\lim_{j\to\infty} k_{n_j} = \hat{k} \in K$. However, the subsequence converges to k and so $k = \hat{k} \in K$. By Theorem 4.8.11, K is closed. Why is K bounded? If it were not, there would exist $\|k_n\| > n$ where $k_n \in K$ and $n \in \mathbb{N}$ which means this sequence could have no convergent subsequence because the subsequence would not even be bounded. See Theorems 4.5.3 and 4.5.2.

$\Leftarrow$Suppose that K is closed and bounded. Since S is bounded, there exists

$$R = \prod_{k=1}^{p} [a_k, b_k]$$

containing K. If $\{k_n\} \subseteq K$, then from Corollary 4.8.3, there exists a subsequence $\{k_{n_j}\}$ such that $\lim_{j\to\infty} k_{n_j} = k \in R$. However, K is closed and so in fact, $k \in K$.

The last claim follows from a repeat of the preceding argument. Just use K in place of R and H in place of K. Alternatively, if K is closed and bounded, then so is H, being a closed subset of K. ■

What about the sequentially compact sets in $\mathbb{C}^p$? This is actually a special case of Theorem 4.8.13. For $z \in \mathbb{C}^p$, $z = (z_1, \cdots, z_p)$ where $z_k = x^k + iy^k$. Thus

$$\left(x^1, y^1, x^2, y^2, \cdots, x^p, y^p\right) \equiv \theta z \in \mathbb{R}^{2p}$$

A set K is bounded in $\mathbb{C}^p$ if and only if $\{\theta z : z \in K\}$ is bounded in $\mathbb{R}^{2p}$. Also, $z_n \to z$ in $\mathbb{C}^p$ if and only if $\theta z_n \to \theta z$ in $\mathbb{R}^{2p}$. Now K is closed and bounded in $\mathbb{C}^p$ if and only if $\theta K \equiv \{\theta z : z \in K\}$ is closed and bounded in $\mathbb{R}^{2p}$and so K is closed and bounded in $\mathbb{C}^p$ if and only if θK is sequentially compact in $\mathbb{R}^{2p}$. Thus if $\{z_n\}$ is a sequence in K, there exists a subsequence, still denoted with n such that θz_n converges in $\mathbb{R}^{2p}$ if and only if z_n converges to some $z \in \mathbb{C}^p$. However, $z \in K$ because K is closed. Thus K is sequentially compact in $\mathbb{C}^p$.

Conversely, if K is sequentially compact, then it must be bounded since otherwise there would be a sequence $\{k_n\} \subseteq K$ with $\|k_n\| > n$ and so no subsequence can be Cauchy so no subsequence can converge. K must also be closed because if not, there would be $x \notin K$ and a sequence $\{k_n\} \subseteq K$ with $k_n \to x$. However, by sequential compactness, there is a subsequence $\{k_{n_k}\}_{k=1}^{\infty}$, $k_{n_k} \to k \in K$ and so $k = x \in K$ after all. This proves most of the following theorem.

Theorem 4.8.14 *Let $H \subseteq \mathbb{F}^p$. Then H is closed and bounded if and only if H is sequentially compact. A sequence $\{x_n\}$ is a Cauchy sequence in $\mathbb{F}^p$ if and only if it converges. In particular, $\mathbb{F}^p$ is complete, $p \geq 1$.*

Proof: Consider the last claim. If $\{z_n\}$ converges, then it is a Cauchy sequence by Theorem 4.5.3. Conversely, if $\{z_n\}$ is a Cauchy sequence, then it is bounded by Theorem 4.5.2 so it is contained in some closed and bounded subset of $\mathbb{F}^p$. Therefore, a subsequence converges to a point of this closed and bounded set. However, by Theorem 4.5.4, the original Cauchy sequence converges to this point. ■

4.8.3 Compactness and Open Coverings

In Theorem 4.8.13 it was shown that sequential compactness in $\mathbb{F}^p$ is the same as closed and bounded. Here we give the traditional definition of compactness and show that this is also equivalent to closed and bounded.

Definition 4.8.15 *A set K is called compact if whenever $\mathscr{C}$ is a collection of open sets such that $K \subseteq \cup\mathscr{C}$, there exists a finite subset of open sets $\{U_1, \cdots, U_m\} \subseteq \mathscr{C}$ such that $K \subseteq \cup_{i=1}^m U_i$. In words, it says that every open cover admits a finite subcover.*

Lemma 4.8.16 *If K is a compact set and H is a closed subset of K, then H is also compact.*

Proof: Let $\mathscr{C}$ be an open cover of H. Then $H^C, \mathscr{C}$ is an open cover of K. It follows that there are finitely many sets of $\mathscr{C}$, $\left\{U_j\right\}_{j=1}^m$ such that $H^C \cup \cup_{j=1}^m U_i \supseteq K$. Therefore, since no points of H are in the open set H^C, it follows that $\cup_{j=1}^m U_i \supseteq H$. ■

Now here is the main result, often called the Heine Borel theorem.

Theorem 4.8.17 *Let K be a nonempty set in $\mathbb{F}^p$. Then the following are equivalent.*

1. *K is compact*
2. *K is closed and bounded*
3. *K is sequentially compact.*

Proof: It was shown above in Theorem 4.8.14 that 2. $\iff$ 3. Consider 3. $\Rightarrow$ 1. If $\mathscr{C}$ is an open cover of K, then I claim there exists $\delta > 0$ such that if $k \in K$, then $B(k, \delta) \subseteq U$ for some $U \in \mathscr{C}$. This δ is called a Lebesgue number. If not, then there exists $k_n \in K$ such that $B\left(k_n, \frac{1}{n}\right)$ is not contained in any set of $\mathscr{C}$ because $1/n$ is not a Lebesgue number. Then by sequential compactness, there is a subsequence, still denoted by k_n which converges to $k \in K$. Now $B(k, \delta) \subseteq U$ for some $\delta > 0$ and some $U \in \mathscr{C}$. However, this is a contradiction because for n large, $\frac{1}{n} < \frac{\delta}{2}$ and $k_n \in B\left(k, \frac{\delta}{2}\right)$ so $B\left(k_n, \frac{1}{n}\right) \subseteq B(k, \delta) \subseteq U$ which is a contradiction. Consider $\{B(k, \delta) : k \in K\}$. Finitely many of these sets contain K in their union since otherwise, there would exist a sequence $\{k_n\}$ such that $\|k_n - k_m\| \geq \delta$ for all $m \neq n$ and so it cannot have any Cauchy subsequence. Hence K would fail to be compact. Thus $K \subseteq \cup_{i=1}^m B(k_i, \delta)$ for suitable finite set $\{k_i\}$. Pick $U_i \in \mathscr{C}$ with $U_i \supseteq B(k_i, \delta)$. Then $\{U_i\}_{i=1}^m$ is an open cover.

It remains to verify that 1.$\Rightarrow$ 2. Suppose that K is compact. Why is it closed and bounded? Suppose first it is not closed. Then there exists a limit point p which is not in K. If $x \in K$, then there exists open U_x containing x and V_x containing p such that $U_x \cap V_x = \emptyset$. Since K is compact, there are finitely many of these U_x which cover K. Say $\{U_{x_1}, \ldots, U_{x_n}\}$. Then let $U = \cup_i U_{x_i}, V = \cap V_{x_i}$, an open set. Hence $p \in V$ and V contains no points of K. Thus p is not a limit point after all. To see that K is bounded, pick $k_0 \in K$ and consider $\{B(k_0, n)\}_{n=1}^\infty$. This is an open cover of K and the sets are increasing so one of these balls covers K. Hence K is bounded. ■

4.8.4 Complete Separability

By Theorem 2.7.9, the rational numbers are dense in $\mathbb{R}$. They are also countable because there is an onto map from the Cartesian product of the two countable sets $\mathbb{Z}$ and $\mathbb{Z} \setminus \{0\}$ to the rationals. Indeed, if m/n is a rational number you consider the ordered pair (m, n) in $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ and let $f((m, n)) \equiv m/n$. Thus, it is possible to enumerate all rational numbers. Of course, as shown earlier, this means there exists a one to one mapping from $\mathbb{N}$ onto $\mathbb{Q}$ but this is not important here. The only thing which matters is that you can write $\mathbb{Q} = \{r_i\}_{i=1}^\infty$. Now the following is the important theorem.

Theorem 4.8.18 *Let $B(x,r)$ denote the interval $(x-r,x+r)$. It is the set of all real numbers y such that y is closer to x than r. Then there are countably many balls $\mathscr{B} \equiv \{B(x,r) : x \in \mathbb{Q}, r \in \mathbb{Q} \cap (0,\infty)\}$. Also every open set is the union of some collection of these balls.*

Proof: Let U be a nonempty open set and let $p \in U$. I need to show that $p \in B \subseteq U$ for some $B \in \mathscr{B}$. There exists $R > 0$ such that $p \in B(p,R) \subseteq U$. Let $x \in \mathbb{Q}$ such that $|p-x| < \frac{R}{10}$. This is possible because $\mathbb{Q}$ is dense. Then letting $\frac{R}{10} < r < \frac{R}{5}$ for $r \in \mathbb{Q}$, it follows that

$$|p-x| < \frac{R}{10} < r$$

and so $p \in B \equiv B(x,r) \in \mathscr{B}$. Also, if $z \in B(x,r)$, then $|z-p| \leq |z-x| + |x-p| < r + \frac{r}{10} < \frac{2R}{5} < R$ and so $p \in B \subseteq B(p,R) \subseteq U$ showing that U is indeed the union of some subset of $\mathscr{B}$. ■

When you have a countable set of open sets with the property that every open set is the union of a subset of this countable set, you call this countable set of open sets a countable basis. When this happens, you say the set is completely separable. . Thus $\mathbb{R}$ along with the usual way of finding distance using the absolute value of the difference of two real numbers is a completely separable set.

Definition 4.8.19 *There is also something called the Lindelöf property.*[4] *It says that if you have any set of open sets $\mathscr{C}$, then there is a countable subset of $\mathscr{C}$ denoted here as $\widehat{\mathscr{C}}$ such that $\cup \widehat{\mathscr{C}} = \cup \mathscr{C}$. Thus this property says that every open cover admits a countable subcover.*

Theorem 4.8.20 *$\mathbb{R}$ has the Lindelöf property.*

Proof: Let $\mathscr{B}$ consist of the open intervals having center a rational number and radius a positive rational number. Then if $\mathscr{C}$ is any collection of open sets, let $\widehat{\mathscr{B}}$ denote those balls of $\mathscr{B}$ which are contained in some set of $\mathscr{C}$. For each $B \in \widehat{\mathscr{B}}$, let $O(B)$ be one of the open sets from $\mathscr{C}$ which contains B. Then since every open set of $\mathscr{C}$ is the union of sets of $\mathscr{B}$, it follows that

$$\cup \mathscr{C} = \cup \widehat{\mathscr{B}} \subseteq \cup \left\{ O(B) : B \in \widehat{\mathscr{B}} \right\} \subseteq \cup \mathscr{C}$$

So let $\widehat{\mathscr{C}} \equiv \left\{ O(B) : B \in \widehat{\mathscr{B}} \right\}$. It is a countable set because $\widehat{\mathscr{B}}$ is countable, being a countable subset of a countable set $\mathscr{B}$ and the mapping $B \to O(B)$ is onto by definition. Note that the axiom of choice is used to select $O(B)$ from the set of open sets of $\mathscr{C}$ which contain B. ■

This is a very useful observation. It holds whenever you have a countable basis. Obviously much of what is being discussed applies to more general situations.

4.9 Exercises

1. Show the intersection of any collection of closed sets is closed and the union of any collection of open sets is open.

[4] Lindelöf was a Finnish mathematician who lived from 1870 to 1946. He did important work in complex analysis.

2. Show that if H is closed and U is open, then $H \setminus U$ is closed. Next show that $U \setminus H$ is open.
3. Show the finite intersection of any collection of open sets is open.
4. Show the finite union of any collection of closed sets is closed.
5. Suppose $\{H_n\}_{n=1}^N$ is a finite collection of sets and suppose x is a limit point of $\cup_{n=1}^N H_n$. Show x must be a limit point of at least one H_n.
6. Give an example of a set of closed sets whose union is not closed.
7. Give an example of a set of open sets whose intersection is not open.
8. Give an example of a set of open sets whose intersection is a closed interval.
9. Give an example of a set of closed sets whose union is open.
10. Give an example of a set of closed sets whose union is an open interval.
11. Give an example of a set of open sets whose intersection is closed.
12. Give an example of a set of open sets whose intersection is the natural numbers.
13. Explain why $\mathbb{F}$ and $\emptyset$ are sets which are both open and closed when considered as subsets of $\mathbb{F}$.
14. Let A be a nonempty set of points and let A' denote the set of limit points of A. Show $A \cup A'$ is closed. **Hint:** You must show the limit points of $A \cup A'$ are in $A \cup A'$. This is shown in the chapter but do it yourself.
15. Let U be any open set in $\mathbb{F}$. Show that every point of U is a limit point of U.
16. Suppose $\{K_n\}$ is a sequence of sequentially compact nonempty sets which have the property that $K_n \supseteq K_{n+1}$ for all n. Show there exists a point in the intersection of all these sets, denoted by $\cap_{n=1}^\infty K_n$.
17. Now suppose $\{K_n\}$ is a sequence of sequentially compact nonempty sets which have the finite intersection property, every finite subset of $\{K_n\}$ has nonempty intersection. Show there exists a point in $\cap_{n=1}^\infty K_n$.
18. Show that any finite union of sequentially compact sets is compact.
19. Start with the unit interval, $I_0 \equiv [0,1]$. In this interval, I_0, remove the following middle third open interval, $(1/3, 2/3)$ resulting in the two closed intervals, $I_1 = [0,1/3] \cup [2/3,1]$. Next delete the middle third of each of these intervals resulting in $I_2 = [0,1/9] \cup [2/9] \cup [2/3,5/9] \cup [8/9,1]$ and continue doing this forever. Show the intersection of all these I_n is nonempty. Letting $P = \cap_{n=1}^\infty I_n$ explain why every point of P is a limit point of P. Would the conclusion be any different if, instead of the middle third open interval, you took out an open interval of arbitrary length, each time leaving two closed intervals where there was one to begin with? This process produces something called the Cantor set. It is the basis for many pathological examples of unbelievably sick functions as well as being an essential ingredient in some extremely important theorems.

20. In Problem 19 in the case where the middle third is taken out, show the total length of open intervals removed equals 1. Thus what is left is very "short". For your information, the Cantor set is uncountable. In addition, it can be shown there exists a function which maps the Cantor set onto $[0,1]$, for example, although you could replace $[0,1]$ with the square $[0,1]\times[0,1]$ or more generally, any compact metric space, something you may study later.

21. Show that there exists an onto map from the Cantor set P just described onto $[0,1]$. Show that this is so even if you do not always take out the middle third, but instead an open interval of arbitrary length, leaving two closed intervals in place of one. It turns out that all of these Cantor sets are topologically the same meaning that there is a one to one onto and continuous mapping from one to another. **Hint:** Base your argument on the nested interval lemma. This will yield ideas which go somewhere.

22. Suppose $\{H_n\}$ is a sequence of sets with the property that for every point x, there exists $r>0$ such that $B(x,r)$ intersects only finitely many of the H_n. Such a collection of sets is called locally finite. Show that if the sets are all closed in addition to being locally finite, then the union of all these sets is also closed. This concept of local finiteness is of great significance although it will not be pursued further here.

23. Show that every uncountable set of points in $\mathbb{F}$ has a limit point. This is not necessarily true if you replace the word, uncountable with the word, infinite. Explain why.

24. In Section 4.8.4 generalize everything to $\mathbb{R}^p$. In this case, the countable dense subset will be $\mathbb{Q}^p$. Also explain why $(\mathbb{Q}+i\mathbb{Q})^p$ is countable and dense subset of $\mathbb{C}^p$ and why $\mathbb{C}^p$ is completely separable.

4.10 Cauchy Sequences and Completeness

You recall the definition of completeness which stated that every nonempty set of real numbers which is bounded above has a least upper bound and that every nonempty set of real numbers which is bounded below has a greatest lower bound and this is a property of the real line known as the completeness axiom. Geometrically, this involved filling in the holes. There is another way of describing completeness in terms of Cauchy sequences. Both of these concepts came during the first part of the nineteenth century and are due to Bolzano and Cauchy.

The next definition has to do with sequences which are real numbers.

Definition 4.10.1 *The sequence of real numbers, $\{a_n\}$, is monotone increasing if for all n, $a_n \leq a_{n+1}$. The sequence is monotone decreasing if for all n, $a_n \geq a_{n+1}$. People often leave off the word "monotone".*

If someone says a sequence is monotone, it usually means monotone increasing.

There exist different descriptions of completeness. An important result is the following theorem which gives a version of completeness in terms of Cauchy sequences. This is often more convenient to use than the earlier definition in terms of least upper bounds and greatest lower bounds because this version of completeness, although it is equivalent to the completeness axiom for the real line, also makes sense in many situations where Definition 2.10.1 on Page 27 does not make sense, $\mathbb{C}$ for example because by Problem 12 on Page 39 there is no way to place an order on $\mathbb{C}$. This is also the case whenever the sequence is of points in multiple dimensions.

It is the concept of completeness and the notion of limits which sets analysis apart from algebra. You will find that every existence theorem in analysis depends on the assumption that some space is complete. In case of $\mathbb{R}$ the least upper bound version corresponds to a statement about convergence of Cauchy sequences.

Theorem 4.10.2 *The following are equivalent.*

1. *Every Cauchy sequence in $\mathbb{R}$ converges*

2. *Every non-empty set of real numbers which is bounded above has a least upper bound.*

3. *Every nonempty set of real numbers which is bounded below has a greatest lower bound.*

Proof: 1.$\Rightarrow$ 2. First suppose every Cauchy sequence converges and let S be a nonempty set which is bounded above. Let b_1 be an upper bound. Pick $s_1 \in S$. If $s_1 = b_1$, the least upper bound has been found and equals b_1. If $(s_1 + b_1)/2$ is an upper bound to S, let this equal b_2. If not, there exists $b_1 > s_2 > (s_1 + b_1)/2$ so let $b_2 = b_1$ and s_2 be as just described. Now let b_2 and s_2 play the same role as s_1 and b_1 and do the same argument. This yields a sequence $\{s_n\}$ of points of S which is monotone increasing and another sequence of upper bounds, $\{b_n\}$ which is monotone decreasing and $|s_n - b_n| \leq 2^{-n+1}(b_1 - s_1)$. Therefore, if $m > n$

$$|b_n - b_m| \leq b_n - s_m \leq b_n - s_n \leq 2^{-n+1}(b_1 - s_1)$$

and so $\{b_n\}$ is a Cauchy sequence. Therefore, it converges to some number b. Then b must be an upper bound of S because if not, there would exist $s > b$ and then $b_n - b \geq s - b$ which would prevent $\{b_n\}$ from converging to b.

2.$\Rightarrow$ 3.,3. $\Rightarrow$ 2. The claim that every nonempty set of numbers bounded below has a greatest lower bound follows similarly. Alternatively, consider $-S \equiv \{-x : x \in S\}$ and apply what was just shown. If S is bounded below, then $-S$ is bounded above and so there exists a least upper bound for $-S$ called $-l$. Then l is a lower bound to S. If there is $b > l$ such that b is also a lower bound to S, then $-b$ would also be an upper bound to $-S$ and would be smaller than $-l$ which contradicts the definition of $-l$. Hence l is the greatest lower bound to S. To show 3. $\Rightarrow$ 2., also consider $-S$ and apply 3. to it similar to what was just done in showing 2.$\Rightarrow$ 3.

2.,3.$\Rightarrow$ 1.Now suppose the condition about existence of least upper bounds and greatest lower bounds. Let $\{a_n\}$ be a Cauchy sequence. Then by Theorem 4.5.2 $\{a_n\} \subseteq [a,b]$ for some numbers a,b. By Theorem 4.8.2 there is a subsequence, $\{a_{n_k}\}$ which converges to $x \in [a,b]$. By Theorem 4.5.4, the original sequence converges to x also. ■

Theorem 4.10.3 *If either of the above conditions for completeness holds, then whenever $\{a_n\}$ is a monotone increasing sequence which is bounded above, it converges and whenever $\{b_n\}$ is a monotone sequence which is bounded below, it converges.*

Proof: Let $a = \sup\{a_n : n \geq 1\}$ and let $\varepsilon > 0$ be given. Then from Proposition 2.10.3 on Page 27 there exists m such that $a - \varepsilon < a_m \leq a$. Since the sequence is increasing, it follows that for all $n \geq m$, $a - \varepsilon < a_n \leq a$. Thus $a = \lim_{n\to\infty} a_n$. The case of a decreasing sequence is similar. Alternatively, you could consider the sequence $\{-a_n\}$ and apply what was just shown to this decreasing sequence. ■

By Theorem 4.10.2 the following definition of completeness is equivalent to the original definition when both apply. However, note that convergence of Cauchy sequences does not depend on an order to it applies to much more general situations. Recall from Theorem 4.8.14 that $\mathbb{C}$ and $\mathbb{R}$ are complete. Just apply that theorem to the case where $p = 1$.

4.10.1 Decimals

You are all familiar with decimals. In the United States these are written in the form $.a_1 a_2 a_3 \cdots$ where the a_i are integers between 0 and 9.[5] Thus .23417432 is a number written as a decimal. You also recall the meaning of such notation in the case of a terminating decimal. For example, .234 is defined as $\frac{2}{10} + \frac{3}{10^2} + \frac{4}{10^3}$. Now what is meant by a nonterminating decimal?

Definition 4.10.4 *Let $.a_1 a_2 \cdots$ be a decimal. Define*

$$.a_1 a_2 \cdots \equiv \lim_{n\to\infty} \sum_{k=1}^{n} \frac{a_k}{10^k}.$$

Proposition 4.10.5 *The above definition makes sense. Also every number in $[0,1]$ can be written as such a decimal.*

Proof: Note the sequence $\left\{\sum_{k=1}^{n} \frac{a_k}{10^k}\right\}_{n=1}^{\infty}$ is an increasing sequence. Therefore, if there exists an upper bound, it follows from Theorem 4.10.3 that this sequence converges and so the definition is well defined.

$$\sum_{k=1}^{n} \frac{a_k}{10^k} \leq \sum_{k=1}^{n} \frac{9}{10^k} = 9 \sum_{k=1}^{n} \frac{1}{10^k}.$$

Now $\frac{9}{10}\left(\sum_{k=1}^{n} \frac{1}{10^k}\right) = \sum_{k=1}^{n} \frac{1}{10^k} - \frac{1}{10}\sum_{k=1}^{n} \frac{1}{10^k} = \sum_{k=1}^{n} \frac{1}{10^k} - \sum_{k=2}^{n+1} \frac{1}{10^k} = \frac{1}{10} - \frac{1}{10^{n+1}}$ and so $\sum_{k=1}^{n} \frac{1}{10^k} \leq \frac{10}{9}\left(\frac{1}{10} - \frac{1}{10^{n+1}}\right) \leq \frac{10}{9}\left(\frac{1}{10}\right) = \frac{1}{9}$. Therefore, since this holds for all n, it follows the above sequence is bounded above. It follows the limit exists.

Now suppose $x \in [0,1)$. Let $\frac{a_1}{10} \leq x < \frac{a_1+1}{10}$ where a_1 is an integer between 0 and 9. If integers $a_1, \cdots, a_n$ each between 0 and 9 have been obtained such that $\sum_{k=1}^{n} \frac{a_k}{10^k} \leq x < \sum_{k=1}^{n-1} \frac{a_k}{10^k} + \frac{a_n+1}{10^n}$ ($\sum_{k=1}^{0} \equiv 0$). Then from the above, $10^n\left(x - \sum_{k=1}^{n} \frac{a_k}{10^k}\right) < 1$ and so there exists a_{n+1} such that

$$\frac{a_{n+1}}{10} \leq 10^n\left(x - \sum_{k=1}^{n} \frac{a_k}{10^k}\right) < \frac{a_{n+1}+1}{10}$$

which shows that $\frac{a_{n+1}}{10^{n+1}} \leq \left(x - \sum_{k=1}^{n} \frac{a_k}{10^k}\right) < \frac{a_{n+1}+1}{10^{n+1}}$. Therefore,

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{a_k}{10^k}$$

[5] In France and Russia they use a comma instead of a period. This looks very strange but that is just the way they do it.

because the distance between the partial sum up to n and x is always no more than $1/10^n$. In case $x = 1$, just let each $a_n = 9$ and observe that the sum of the geometric series equals 1. ■

An amusing application of the above is in the following theorem. It gives an easy way to verify that the unit interval is uncountable.

Theorem 4.10.6 *The interval* $[0,1)$ *is not countable.*

Proof: Suppose it were. Then there would exist a list of all the numbers in this interval. Writing these as decimals,

$$\begin{aligned} x_1 &\equiv .a_{11}a_{12}a_{13}a_{14}a_{15}\cdots \\ x_2 &\equiv .a_{21}a_{22}a_{23}a_{14}a_{25}\cdots \\ x_3 &\equiv .a_{31}a_{32}a_{33}a_{34}a_{35}\cdots \\ &\vdots \end{aligned}$$

Consider the diagonal decimal, $.a_{11}a_{22}a_{33}a_{44}\cdots$. Now define a decimal expansion for another number in $[0,1)$ as follows. $y \equiv .b_1b_2b_3b_4\cdots$ where $|b_k - a_{kk}| \geq 4$. Then $|y - x_k| \geq \frac{4}{10^k}$.Thus y is not equal to any of the x_k which is a contradiction since $y \in [0,1)$. ■

4.10.2 lim sup and lim inf

Sometimes the limit of a sequence does not exist. For example, if $a_n = (-1)^n$, then $\lim_{n\to\infty} a_n$ does not exist. This is because the terms of the sequence are a distance of 1 apart. Therefore there can't exist a single number such that all the terms of the sequence are ultimately within $1/4$ of that number. The nice thing about lim sup and lim inf is that they always exist. First here is a simple lemma and definition. First review the definition of inf and sup on Page 27 along with the simple properties of these things. Also $\lim_{n\to\infty} a_n = \infty$ means that if $l \in \mathbb{R}$ is given, then for large enough $n, a_n > l$. A similar definition holds for $\lim_{n\to\infty} a_n = -\infty$.

Definition 4.10.7 *Denote by* $[-\infty,\infty]$ *the real line along with symbols* ∞ *and* $-\infty$*. It is understood that* ∞ *is larger than every real number and* $-\infty$ *is smaller than every real number. Then if* $\{A_n\}$ *is an increasing sequence of points of* $[-\infty,\infty]$, $\lim_{n\to\infty} A_n$ *equals* ∞ *if the only upper bound of the set* $\{A_n\}$ *is* ∞*. If* $\{A_n\}$ *is bounded above by a real number, then* $\lim_{n\to\infty} A_n$ *is defined in the usual way and equals the least upper bound of* $\{A_n\}$*. If* $\{A_n\}$ *is a decreasing sequence of points of* $[-\infty,\infty]$, $\lim_{n\to\infty} A_n$ *equals* $-\infty$ *if the only lower bound of the sequence* $\{A_n\}$ *is* $-\infty$*. If* $\{A_n\}$ *is bounded below by a real number, then* $\lim_{n\to\infty} A_n$ *is defined in the usual way and equals the greatest lower bound of* $\{A_n\}$*. More simply, if* $\{A_n\}$ *is increasing,* $\lim_{n\to\infty} A_n \equiv \sup\{A_n\}$*and if* $\{A_n\}$ *is decreasing then* $\lim_{n\to\infty} A_n \equiv \inf\{A_n\}$.

Lemma 4.10.8 *Let* $\{a_n\}$ *be a sequence of real numbers and let* $U_n \equiv \sup\{a_k : k \geq n\}$. *Then* $\{U_n\}$ *is a decreasing sequence. Also if* $L_n \equiv \inf\{a_k : k \geq n\}$, *then* $\{L_n\}$ *is an increasing sequence. Therefore,* $\lim_{n\to\infty} L_n$ *and* $\lim_{n\to\infty} U_n$ *both exist.*

Proof: From the definition, if $m \leq n$, $L_m \equiv \inf\{a_k : k \geq m\} \leq \inf\{a_k : k \geq n\} \equiv L_n$. Thus the L_n are increasing. If you take inf of a smaller set, it will be as large as inf of the larger set. Similarly the U_n are decreasing. Thus their limits exist as in the above definition. ■

From the lemma, the following definition makes sense.

Definition 4.10.9 *Let $\{a_n\}$ be any sequence of points of $[-\infty,\infty]$*

$$\limsup_{n\to\infty} a_n \equiv \lim_{n\to\infty} \sup\{a_k : k \geq n\}$$

$$\liminf_{n\to\infty} a_n \equiv \lim_{n\to\infty} \inf\{a_k : k \geq n\}.$$

Theorem 4.10.10 *Suppose $\{a_n\}$ is a sequence of real numbers and that*

$$\limsup_{n\to\infty} a_n \text{ and } \liminf_{n\to\infty} a_n$$

are both real numbers. Then $\lim_{n\to\infty} a_n$ exists if and only if $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$ and in this case, $\lim_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$.

Proof: First note that $\sup\{a_k : k \geq n\} \geq \inf\{a_k : k \geq n\}$ and so from Theorem 4.4.13,

$$\limsup_{n\to\infty} a_n \equiv \lim_{n\to\infty} \sup\{a_k : k \geq n\} \geq \lim_{n\to\infty} \inf\{a_k : k \geq n\} \equiv \liminf_{n\to\infty} a_n.$$

$\Rightarrow$ Suppose $\lim_{n\to\infty} a_n = a \in \mathbb{R}$. Then given $\varepsilon > 0$, there is N such that if $n \geq N$,

$$a - \varepsilon \leq a_n \leq a + \varepsilon$$

It follows that if $n \geq N, a - \varepsilon \leq L_n \leq U_n \leq a + \varepsilon$.Passing to a limit, it follows from Theorem 4.4.13

$$a - \varepsilon \leq \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n \leq a + \varepsilon$$

and so, since ε is arbitrary, lim inf and lim sup are equal to the limit a.

$\Leftarrow$ Suppose $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = a \in \mathbb{R}$. Then if n is large enough,

$$a - \varepsilon \leq L_n \leq a_n \leq U_n \leq a + \varepsilon$$

Since ε is arbitrary, $\lim_{n\to\infty} a_n = a$. ■

With the above theorem, here is how to define the limit of a sequence of points in $[-\infty,\infty]$, the new case being that a_n is allowed to be $\pm\infty$.

Definition 4.10.11 *Let $\{a_n\}$ be a sequence in $[-\infty,\infty]$. $\lim_{n\to\infty} a_n$ exists exactly when $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$, and in this case*

$$\lim_{n\to\infty} a_n \equiv \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n.$$

The significance of lim sup and lim inf, in addition to what was just discussed, is contained in the following theorem which follows quickly from the definition.

Theorem 4.10.12 *Suppose $\{a_n\}$ is a sequence of points of $[-\infty,\infty]$. Then let $\lambda = \limsup_{n\to\infty} a_n$. Then if $b > \lambda$, it follows there exists N such that whenever $n \geq N, a_n \leq b$. If $c < \lambda$, then $a_n > c$ for infinitely many values of n. Let $\gamma = \liminf_{n\to\infty} a_n$. Then if $d < \gamma$, it follows there exists N such that whenever $n \geq N, a_n \geq d$. If $e > \gamma$, it follows $a_n < e$ for infinitely many values of n.*

The proof of this theorem is left as an exercise for you. It follows directly from the definition and it is the sort of thing you must do yourself. Here is one other simple proposition.

Proposition 4.10.13 *Let* $\lim_{n\to\infty} a_n = a > 0$ *and suppose each* $b_n > 0$. *Then*

$$\limsup_{n\to\infty} a_n b_n = a \limsup_{n\to\infty} b_n.$$

Proof: This follows from the definition. Let $\lambda_n = \sup\{a_k b_k : k \geq n\}$. For all n large enough, $a_n > a - \varepsilon$ where ε is small enough that $a - \varepsilon > 0$. Therefore,

$$\lambda_n \geq \sup\{b_k : k \geq n\}(a - \varepsilon)$$

for all n large enough. Then $\limsup_{n\to\infty} a_n b_n = \lim_{n\to\infty} \lambda_n \equiv$

$$\limsup_{n\to\infty} a_n b_n \geq \lim_{n\to\infty} (\sup\{b_k : k \geq n\}(a - \varepsilon)) = (a - \varepsilon) \limsup_{n\to\infty} b_n$$

Similar reasoning shows $\limsup_{n\to\infty} a_n b_n \leq (a + \varepsilon) \limsup_{n\to\infty} b_n$. Now since $\varepsilon > 0$ is arbitrary, the conclusion follows. ■

4.10.3 Shrinking Diameters

It is useful to consider another version of the nested interval lemma. This involves a sequence of sets such that set $(n+1)$ is contained in set n and such that their diameters converge to 0. It turns out that if the sets are also closed, then often there exists a unique point in all of them. This is just a more general version of the nested interval theorem which holds in the context that the sets are not necessarily intervals.

Definition 4.10.14 *Let S be a nonempty set. Then* $\operatorname{diam}(S)$ *is defined as*

$$\operatorname{diam}(S) \equiv \sup\{|x - y| : x, y \in S\}.$$

This is called the diameter of S.

Theorem 4.10.15 *Let* $\{F_n\}_{n=1}^{\infty}$ *be a sequence of closed sets in* $\mathbb{F}^p$ *such that*

$$\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$$

and $F_n \supseteq F_{n+1}$ *for each n. Then there exists a unique* $p \in \cap_{k=1}^{\infty} F_k$.

Proof: Pick $p_k \in F_k$. This is always possible because by assumption each set is nonempty. Then $\{p_k\}_{k=m}^{\infty} \subseteq F_m$ and since the diameters converge to 0 it follows $\{p_k\}$ is a Cauchy sequence. Therefore, it converges to a point, p by completeness of $\mathbb{F}^p$. Since each F_k is closed, it must be that $p \in F_k$ for all k. Therefore, $p \in \cap_{k=1}^{\infty} F_k$. If $q \in \cap_{k=1}^{\infty} F_k$, then since both $p, q \in F_k$, $|p - q| \leq \operatorname{diam}(F_k)$. It follows since these diameters converge to 0, $|p - q| \leq \varepsilon$ for every ε. Hence $p = q$. ■

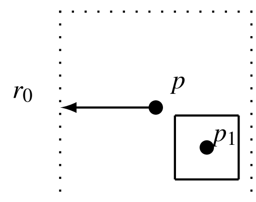
A sequence of sets, $\{G_n\}$ which satisfies $G_n \supseteq G_{n+1}$ for all n is called a nested sequence of sets.

The next theorem is a major result called Bair's theorem. In fact, you just need the context of a complete metric space but we are emphasizing $\mathbb{F}^p$ here.

Definition 4.10.16 *An open set* $U \subseteq \mathbb{F}^p$ *is dense if for every* $x \in \mathbb{F}^p$ *and* $r > 0, B(x, r) \cap U \neq \emptyset$.

Theorem 4.10.17 *Let $\{U_n\}$ be a sequence of dense open sets. Then $\cap_n U_n$ is dense.*

Proof: Let $p \in \mathbb{F}^p$ and let $r_0 > 0$. I 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 $r_1 < 2^{-1}$. This is possible because $U_1 \cap B(p,r_0)$ is an open set and so there exists r_1 such that $B(p_1,2r_1) \subseteq U_1 \cap B(p,r_0)$. But $B(p_1,r_1) \subseteq \overline{B(p_1,r_1)} \subseteq B(p_1,2r_1)$ because $\overline{B(p_1,r_1)} = \{x \in X : d(x,p) \leq r_1\}$. (Why?)



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 $r_n < 2^{-n}$,

$$\overline{B(p_n,r_n)} \subseteq U_1 \cap U_2 \cap ... \cap U_n \cap B(p,r_0),\ \overline{B(p_n,r_n)} \subseteq B(p_{n-1},r_{n-1}).$$

The sequence, $\{p_n\}$ is a Cauchy sequence because all terms of $\{p_k\}$ for $k \geq n$ are contained in $B(p_n,r_n)$, a set whose diameter is no larger than 2^{-n}. Since $\mathbb{F}^p$ is complete, (Theorem 4.8.14) there exists p_∞ such that $\lim_{n\to\infty} p_n = p_\infty$. Since all but finitely many terms of $\{p_n\}$ are in $\overline{B(p_m,r_m)}$, it follows that $p_\infty \in \overline{B(p_m,r_m)}$ for each m. Therefore,

$$p_\infty \in \cap_{m=1}^{\infty} \overline{B(p_m,r_m)} \subseteq \cap_{i=1}^{\infty} U_i \cap B(p,r_0). \blacksquare$$

The countable intersection of open sets is called a G_δ set.

4.11 The Euclidean Norm

For $a \equiv (a_1,\cdots,a_p) \in \mathbb{F}$, define $|a| \equiv \left(\sum_{k=1}^p |a_k|^2\right)^{1/2}$. Then it is obvious that $|\alpha a| = |\alpha|\,|a|$ whenever $\alpha \in \mathbb{F}$ and it is obvious that $|a| \geq 0$ and equals 0 if and only if $a = 0$, the zero vector. As to the triangle inequality 4.2, by the Cauchy Schwarz inequality,

$$\begin{aligned}
|x+y|^2 &\equiv \sum_{k=1}^p |x_k+y_k|^2 = \sum_{k=1}^p |x_k|^2 + \sum_{k=1}^p |y_k|^2 + 2\sum_{k=1}^p \operatorname{Re}(x_k\overline{y_k}) \\
&\leq \sum_{k=1}^p |x_k|^2 + \sum_{k=1}^p |y_k|^2 + 2\sum_{k=1}^p |x_k|\,|\overline{y_k}| \\
&\leq \sum_{k=1}^p |x_k|^2 + |y_k|^2 + 2\left(\sum_{k=1}^p |x_k|^2\right)^{1/2}\left(\sum_{k=1}^p |y_k|^2\right)^{1/2} = (|x|+|y|)^2
\end{aligned}$$

so $|x+y| \leq |x|+|y|$. Also, it is obvious that

$$n\,\|x\|^2 \geq |x|^2 \geq \|x\|^2 \text{ so } \sqrt{n}\,\|x\| \geq |x| \geq \|x\|$$

Thus, with this Euclidean norm $\mathbb{F}^p$ has the same Cauchy sequences, the same open and closed sets, and all the same theorems concerning compactness. Thus, from the point of view of analysis, there is no difference. The reason for the Euclidean norm is that it is geometrically better.

4.12 Exercises

1. Suppose $x = .3434343\overline{434}$ where the bar over the last 34 signifies that this repeats forever. In elementary school you were probably given the following procedure for finding the number x as a quotient of integers. First multiply by 100 to get $100x = 34.343434\overline{34}$ and then subtract to get $99x = 34$. From this you conclude that $x = 34/99$. Fully justify this procedure. **Hint:** $.343434\overline{34} = \lim_{n\to\infty} 34\sum_{k=1}^{n}\left(\frac{1}{100}\right)^k$ now use Problem 7 on Page 66.

2. Let $a \in [0,1]$. Show $a = .a_1a_2a_3...$ for some choice of integers in $\{0,1,2,\cdots,9\}$, $a_1,a_2,\cdots$ if it is possible to do this. Give an example where there may be more than one way to do this.

3. Show every rational number between 0 and 1 has a decimal expansion which either repeats or terminates.

4. Using Corollary 3.2.5, show that there exists a one to one and onto map θ from the natural numbers $\mathbb{N}$ onto $\mathbb{Q}$, the rational number. Denoting the resulting countable set of numbers as the sequence $\{r_n\}$, show that if x is any real number, there exists a subsequence from this sequence which converges to that number.

5. A number has decimal expansion $.010010001000010000001\cdots$. Show this is an irrational number.

6. Prove $\sqrt{2}$ is irrational. **Hint:** Suppose $\sqrt{2} = p/q$ where p,q are positive integers and the fraction is in lowest terms. Then $2q^2 = p^2$ and so p^2 is even. Explain why $p = 2r$ so p must be even. Next argue q must be even.

7. Show that between any two integers there exists an irrational number. Next show that between any two numbers there exists an irrational number. You can use the fact that $\sqrt{2}$ is irrational if you like.

8. Let a be a positive number and let $x_1 = b > 0$ where $b^2 > a$. Explain why there exists such a number, b. Now having defined x_n, define $x_{n+1} \equiv \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$. Verify that $\{x_n\}$ is a decreasing sequence and that it satisfies $x_n^2 \geq a$ for all n and is therefore, bounded below. Explain why $\lim_{n\to\infty} x_n$ exists. If x is this limit, show that $x^2 = a$. Explain how this shows that every positive real number has a square root. This is an example of a recursively defined sequence. Note this does not give a formula for x_n, just a rule which tells how to define x_{n+1} if x_n is known.

9. Let $a_1 = 0$ and suppose that $a_{n+1} = \frac{9}{9-a_n}$. Write a_2,a_3,a_4. Now prove that for all n, it follows that $a_n \leq \frac{9}{2} + \frac{3}{2}\sqrt{5}$. Find the limit of the sequence. **Hint:** You should prove these things by induction. Finally, to find the limit, let $n \to \infty$ in both sides and argue that the limit a, must satisfy $a = \frac{9}{9-a}$.

10. If $\lim_{n\to\infty} a_n = a$, does it follow that $\lim_{n\to\infty} |a_n| = |a|$? Prove or else give a counter example.

11. Show $\lim_{n\to\infty} \frac{n^5}{1.01^n} = 0$, $\lim_{n\to\infty} \frac{100^n}{n!} = 0$.

12. Suppose $\lim_{n\to\infty} x_n = x$. Show that then $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} x_k = x$. Give an example where $\lim_{n\to\infty} x_n$ does not exist but $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} x_k$ does.

13. Suppose $r \in (0,1)$. Show that $\lim_{n\to\infty} r^n = 0$. **Hint:** Use the binomial theorem. $r = \frac{1}{1+\delta}$ where $\delta > 0$. Therefore, $r^n = \frac{1}{(1+\delta)^n} < \frac{1}{1+n\delta}$, etc.

14. Prove $\lim_{n\to\infty} \sqrt[n]{n} = 1$. **Hint:** Let $e_n \equiv \sqrt[n]{n} - 1$ so that $(1+e_n)^n = n$. Now observe that $e_n > 0$ and use the binomial theorem to conclude $1 + ne_n + \frac{n(n-1)}{2} e_n^2 \leq n$. This nice approach to establishing this limit using only elementary algebra is in Rudin [24].

15. Find $\lim_{n\to\infty} (x^n + 5)^{1/n}$ for $x \geq 0$. There are two cases here, $x \leq 1$ and $x > 1$. Show that if $x > 1$, the limit is x while if $x \leq 1$ the limit equals 1. **Hint:** Use the argument of Problem 14. This interesting example is in [11].

16. Find $\limsup_{n\to\infty} (-1)^n$ and $\liminf_{n\to\infty} (-1)^n$. Explain your conclusions.

17. Give a careful proof of Theorem 4.10.12.

18. Let $\{a_n\}$ be a sequence in $(-\infty,\infty)$. Let $A_k \equiv \sup\{a_n : n \geq k\}$ so that
$$\lambda \equiv \limsup_{n\to\infty} a_n = \lim_{n\to\infty} A_n,$$
the A_n being a decreasing sequence.
    (a) Show that in all cases, there exists $B_n < A_n$ such that B_n is increasing and $\lim_{n\to\infty} B_n = \lambda$.
    (b) Explain why, in all cases there are infinitely many k such that $a_k \in [B_n, A_n]$. **Hint:** If for all $k \geq m > n$, $a_k \leq B_n$, then $a_k < B_m$ also and so $\sup\{a_k : k \geq m\} \leq B_m < A_m$ contrary to the definition of A_m.
    (c) Explain why there exists a subsequence $\{a_{n_k}\}$ such that $\lim_{k\to\infty} a_{n_k} = \lambda$.
    (d) Show that if $\gamma \in [-\infty,\infty]$ and there is a subsequence $\{a_{n_k}\}$ which has the property that $\lim_{k\to\infty} a_{n_k} = \gamma$, then $\gamma \leq \lambda$.

    This shows that $\limsup_{n\to\infty} a_n$ is the largest in $[-\infty,\infty]$ such that some subsequence converges to it. Would it all work if you only assumed that $\{a_n\}$ is not $-\infty$ for infinitely many n? What if $a_n = -\infty$ for all n large enough? Isn't this case fairly easy? The next few problems are similar.

19. Let $\lambda = \limsup_{n\to\infty} a_n$. Show there exists a subsequence, $\{a_{n_k}\}$ such that
$$\lim_{k\to\infty} a_{n_k} = \lambda.$$
Now consider the set S of all points in $[-\infty,\infty]$ such that for $s \in S$, some subsequence of $\{a_n\}$ converges to s. Show that S has a largest point and this point is $\limsup_{n\to\infty} a_n$.

20. Let $\{a_n\} \subseteq \mathbb{R}$ and suppose it is bounded above. Let
$$S \equiv \{x \in \mathbb{R} \text{ such that } x \leq a_n \text{ for infinitely many } n\}$$
Show that for each n, $\sup(S) \leq \sup\{a_k : k \geq n\}$. Why is $\sup(S) \leq \limsup_{n\to\infty} a_k$? Next explain why the two numbers are actually equal. Explain why such a sequence has a convergent subsequence. For the last part, see Problem 19 above.

21. Let $\lambda = \liminf_{n\to\infty} a_n$. Show there exists a subsequence, $\{a_{n_k}\}$ and $\lim_{k\to\infty} a_{n_k} = \lambda$. Now consider the set, S of all points in $[-\infty,\infty]$ such that for $s \in S$, some subsequence of $\{a_n\}$ converges to s. Show that S has a smallest point and this point is $\liminf_{n\to\infty} a_n$. Formulate a similar conclusion to Problem 20 in terms of $\liminf$ and a sequence which is bounded below.

22. Prove that if $a_n \leq b_n$ for all n sufficiently large that

$$\liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n, \ \limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n.$$

23. Prove that $\limsup_{n\to\infty} (-a_n) = -\liminf_{n\to\infty} a_n$.

24. Prove that if $a \geq 0$, then $\limsup_{n\to\infty} a a_n = a \limsup_{n\to\infty} a_n$ while if $a < 0$,

$$\limsup_{n\to\infty} a a_n = a \liminf_{n\to\infty} a_n.$$

25. Prove that if $\lim_{n\to\infty} b_n = b$, then $\limsup_{n\to\infty} (b_n + a_n) = b + \limsup_{n\to\infty} a_n$. Conjecture and prove a similar result for $\liminf$.

26. Give conditions under which the following inequalities hold.

$$\begin{aligned} \limsup_{n\to\infty} (a_n + b_n) &\leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \\ \liminf_{n\to\infty} (a_n + b_n) &\geq \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n. \end{aligned}$$

Hint: You need to consider whether the right hand sides make sense. Thus you can't consider $-\infty + \infty$.

27. Give an example of a nested sequence of nonempty sets whose diameters converge to 0 which have no point in their intersection.

28. Give an example of a nested sequence of nonempty sets S_n such that $S_n \supsetneq S_{n+1}$ whose intersection has more than one point. Next give an example of a nested sequence of nonempty sets S_n, $S_n \supsetneq S_{n+1}$ which has 2 points in their intersection.

29. For $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, suppose $\mathbb{F} = \cup_{n=1}^{\infty} H_n$ where each H_n is closed. Show that at least one of these must have nonempty interior. That is, one of them contains an open ball. You can use Theorem 4.10.17 if you like.

Chapter 5

Infinite Series of Numbers

5.1 Basic Considerations

Earlier in Definition 4.4.1 on Page 60 the notion of limit of a sequence was discussed. There is a very closely related concept called an infinite series which is dealt with in this section.

Definition 5.1.1 *Define $\sum_{k=m}^{\infty} a_k \equiv \lim_{n\to\infty} \sum_{k=m}^{n} a_k$ whenever the limit exists and is finite. In this case the series is said to converge. If the series does not converge, it is said to diverge. The sequence $\{\sum_{k=m}^{n} a_k\}_{n=m}^{\infty}$ in the above is called the sequence of partial sums. This is always the definition. Here it is understood that the a_k are in $\mathbb{F}$, either $\mathbb{R}$ or $\mathbb{C}$ but it is the same definition in more general situations.*

From this definition, it should be clear that infinite sums do not always make sense. Sometimes they do and sometimes they don't, depending on the behavior of the partial sums. As an example, consider $\sum_{k=1}^{\infty} (-1)^k$. The partial sums corresponding to this symbol alternate between -1 and 0. Therefore, there is no limit for the sequence of partial sums. It follows the symbol just written is meaningless and the infinite sum diverges.

Example 5.1.2 *Find the infinite sum, $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$.*

Note $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$ and so $\sum_{n=1}^{N} \frac{1}{n(n+1)} = \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) = -\frac{1}{N+1} + 1$. Therefore, $\lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n(n+1)} = \lim_{N\to\infty} \left(-\frac{1}{N+1} + 1\right) = 1$.

Proposition 5.1.3 *Let $a_k \geq 0$. Then $\{\sum_{k=m}^{n} a_k\}_{n=m}^{\infty}$ is an increasing sequence. If this sequence is bounded above, then $\sum_{k=m}^{\infty} a_k$ converges and its value equals*

$$\sup\left\{\sum_{k=m}^{n} a_k : n = m, m+1, \cdots\right\}.$$

When the sequence is not bounded above, $\sum_{k=m}^{\infty} a_k$ diverges.

Proof: It follows that $\{\sum_{k=m}^{n} a_k\}_{n=m}^{\infty}$ is an increasing sequence because

$$\sum_{k=m}^{n+1} a_k - \sum_{k=m}^{n} a_k = a_{n+1} \geq 0.$$

If it is bounded above, then by the form of completeness found in Theorem 4.10.2 on Page 76 it follows that the sequence of partial sums converges to

$$\sup\left\{\sum_{k=m}^{n} a_k : n = m, m+1, \cdots\right\}$$

If the sequence of partial sums is not bounded, then it is not a Cauchy sequence and so it does not converge. See Theorem 4.5.3 on Page 64. ■

In the case where $a_k \geq 0$, the above proposition shows there are only two alternatives available. Either the sequence of partial sums is bounded above or it is not bounded above.

In the first case convergence occurs and in the second case, the infinite series diverges. For this reason, people will sometimes write $\sum_{k=m}^{\infty} a_k < \infty$ to denote the case where convergence occurs and $\sum_{k=m}^{\infty} a_k = \infty$ for the case where divergence occurs. Be very careful you never think this way in the case where it is not true that all $a_k \geq 0$. For example, the partial sums of $\sum_{k=1}^{\infty} (-1)^k$ are bounded because they are all either -1 or 0 but the series does not converge.

One of the most important examples of a convergent series is the geometric series. This series is $\sum_{n=0}^{\infty} r^n$. The study of this series depends on simple high school algebra and Theorem 4.4.11 on Page 63. Let $S_n \equiv \sum_{k=0}^{n} r^k$. Then

$$S_n = \sum_{k=0}^{n} r^k, \ rS_n = \sum_{k=0}^{n} r^{k+1} = \sum_{k=1}^{n+1} r^k.$$

Therefore, subtracting the second equation from the first yields $(1-r) S_n = 1 - r^{n+1}$ and so a formula for S_n is available. In fact, if $r \neq 1, S_n = \frac{1-r^{n+1}}{1-r}$.By Theorem 4.4.11, $\lim_{n\to\infty} S_n = \frac{1}{1-r}$ in the case when $|r| < 1$. Now if $|r| \geq 1$, the limit clearly does not exist because S_n fails to be a Cauchy sequence (Why?). This shows the following.

Theorem 5.1.4 *The geometric series, $\sum_{n=0}^{\infty} r^n$ converges and equals $\frac{1}{1-r}$ if $|r| < 1$ and diverges if $|r| \geq 1$.*

If the series do converge, the following holds.

Theorem 5.1.5 *If $\sum_{k=m}^{\infty} a_k$ and $\sum_{k=m}^{\infty} b_k$ both converge and x, y are numbers, then*

$$\sum_{k=m}^{\infty} a_k = \sum_{k=m+j}^{\infty} a_{k-j} \tag{5.1}$$

$$\sum_{k=m}^{\infty} xa_k + yb_k = x \sum_{k=m}^{\infty} a_k + y \sum_{k=m}^{\infty} b_k \tag{5.2}$$

$$\left| \sum_{k=m}^{\infty} a_k \right| \leq \sum_{k=m}^{\infty} |a_k| \tag{5.3}$$

where in the last inequality, the last sum equals $+\infty$ if the partial sums are not bounded above.

Proof: The above theorem is really only a restatement of Theorem 4.4.8 on Page 62 and the above definitions of infinite series. Thus

$$\sum_{k=m}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=m}^{n} a_k = \lim_{n\to\infty} \sum_{k=m+j}^{n+j} a_{k-j} = \sum_{k=m+j}^{\infty} a_{k-j}.$$

To establish 5.2, use Theorem 4.4.8 on Page 62 to write

$$\begin{aligned} \sum_{k=m}^{\infty} xa_k + yb_k &= \lim_{n\to\infty} \sum_{k=m}^{n} xa_k + yb_k = \lim_{n\to\infty} \left(x \sum_{k=m}^{n} a_k + y \sum_{k=m}^{n} b_k \right) \\ &= x \sum_{k=m}^{\infty} a_k + y \sum_{k=m}^{\infty} b_k. \end{aligned}$$

Formula 5.3 follows from the observation that, from the triangle inequality,

$$\left|\sum_{k=m}^{n} a_k\right| \leq \sum_{k=m}^{n} |a_k| \leq \sum_{k=m}^{\infty} |a_k|$$

and so $\left|\sum_{k=m}^{\infty} a_k\right| = \lim_{n\to\infty} \left|\sum_{k=m}^{n} a_k\right| \leq \sum_{k=m}^{\infty} |a_k|$. ■

Recall that if $\lim_{n\to\infty} A_n = A$, then $\lim_{n\to\infty} |A_n| = |A|$.

Example 5.1.6 *Find* $\sum_{n=0}^{\infty} \left(\frac{5}{2^n} + \frac{6}{3^n}\right)$.

From the above theorem and Theorem 5.1.4,

$$\sum_{n=0}^{\infty} \left(\frac{5}{2^n} + \frac{6}{3^n}\right) = 5\sum_{n=0}^{\infty} \frac{1}{2^n} + 6\sum_{n=0}^{\infty} \frac{1}{3^n} = 5\frac{1}{1-(1/2)} + 6\frac{1}{1-(1/3)} = 19.$$

The following criterion is useful in checking convergence. All it is saying is that the series converges if and only if the sequence of partial sums is Cauchy. This is what the given criterion says. However, this is not new information.

Theorem 5.1.7 *Let $\{a_k\}$ be a sequence of points in $\mathbb{F}$. The sum $\sum_{k=m}^{\infty} a_k$ converges if and only if for all $\varepsilon > 0$, there exists n_ε such that if $q \geq p \geq n_\varepsilon$, then*

$$\left|\sum_{k=p}^{q} a_k\right| < \varepsilon. \tag{5.4}$$

Proof: Suppose first that the series converges. Then $\{\sum_{k=m}^{n} a_k\}_{n=m}^{\infty}$ is a Cauchy sequence by Theorem 4.5.3 on Page 64. Therefore, there exists $n_\varepsilon > m$ such that if $q \geq p-1 \geq n_\varepsilon > m$,

$$\left|\sum_{k=m}^{q} a_k - \sum_{k=m}^{p-1} a_k\right| = \left|\sum_{k=p}^{q} a_k\right| < \varepsilon. \tag{5.5}$$

Next suppose 5.4 holds. Then from 5.5 it follows upon letting p be replaced with $p+1$ that $\{\sum_{k=m}^{n} a_k\}_{n=m}^{\infty}$ is a Cauchy sequence and so, by Theorem 4.8.14, it converges. By the definition of infinite series, this shows the infinite sum converges as claimed. ■

5.2 Absolute Convergence

Definition 5.2.1 *The statement that a series $\sum_{k=m}^{\infty} a_k$ converges absolutely means $\sum_{k=m}^{\infty} |a_k|$ converges. If the series does converge but does not converge absolutely, then it is said to converge conditionally.*

Theorem 5.2.2 *If $\sum_{k=m}^{\infty} a_k$ converges absolutely, then it converges.*

Proof: Let $\varepsilon > 0$ be given. Then by assumption and Theorem 5.1.7, there exists n_ε such that whenever $q \geq p \geq n_\varepsilon$, $\sum_{k=p}^{q} |a_k| < \varepsilon$. Therefore, from the triangle inequality, $\varepsilon > \sum_{k=p}^{q} |a_k| \geq \left|\sum_{k=p}^{q} a_k\right|$. By Theorem 5.1.7, $\sum_{k=m}^{\infty} a_k$ converges. ■

In fact, the above theorem is really another version of the completeness axiom. Thus its validity implies completeness. You might try to show this.

One of the interesting things about absolutely convergent series is that you can "add them up" in any order and you will always get the same thing. This is the meaning of the following theorem. Of course there is no problem when you are dealing with finite sums thanks to the commutative law of addition. However, when you have infinite sums strange and wonderful things can happen because these involve a limit.

Theorem 5.2.3 *Let $\theta : \mathbb{N} \to \mathbb{N}$ be one to one and onto. Suppose $\sum_{k=1}^{\infty} a_k$ converges absolutely. Then $\sum_{k=1}^{\infty} a_{\theta(k)} = \sum_{k=1}^{\infty} a_k$.*

Proof: From absolute convergence, there exists M such that

$$\sum_{k=M+1}^{\infty} |a_k| \equiv \left(\sum_{k=1}^{\infty} |a_k| - \sum_{k=1}^{M} |a_k| \right) < \varepsilon.$$

Since θ is one to one and onto, there exists $N \geq M$ such that

$$\{1, 2, \cdots, M\} \subseteq \{\theta(1), \theta(2), \cdots, \theta(N)\}.$$

It follows that it is also the case that $\sum_{k=N+1}^{\infty} \left|a_{\theta(k)}\right| < \varepsilon$. This is because the partial sums of the above series are each dominated by a partial sum for $\sum_{k=M+1}^{\infty} |a_k|$ since every index $\theta(k)$ equals some n for $n \geq M+1$. Then since ε is arbitrary, this shows that the partial sums of $\sum a_{\theta(k)}$ are Cauchy. Hence, this series does converge and also

$$\left| \sum_{k=1}^{M} a_k - \sum_{k=1}^{N} a_{\theta(k)} \right| \leq \sum_{k=M+1}^{\infty} |a_k| < \varepsilon$$

Hence

$$\left| \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} a_{\theta(k)} \right| \leq \left| \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{M} a_k \right| + \left| \sum_{k=1}^{M} a_k - \sum_{k=1}^{N} a_{\theta(k)} \right|$$
$$+ \left| \sum_{k=1}^{N} a_{\theta(k)} - \sum_{k=1}^{\infty} a_{\theta(k)} \right| < \sum_{k=M+1}^{\infty} |a_k| + \varepsilon + \sum_{k=N+1}^{\infty} \left|a_{\theta(k)}\right| < 3\varepsilon$$

Since ε is arbitrary, this shows the two series are equal as claimed. ■

So what happens when series converge only conditionally?

Example 5.2.4 *Consider the series $\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$. Show that there is a rearrangement which converges to 7 although this series does converge. (In fact, it converges to $-\ln 2$ for those who remember calculus.)*

First of all consider why it converges. Notice that if S_n denotes the n^{th} partial sum, then

$$\begin{aligned}
S_{2n} - S_{2n-2} &= \frac{1}{2n} - \frac{1}{2n-1} < 0 \\
S_{2n+1} - S_{2n-1} &= -\frac{1}{2n+1} + \frac{1}{2n} > 0 \\
S_{2n} - S_{2n-1} &= \frac{1}{2n}
\end{aligned}$$

Thus the even partial sums are decreasing and the odd partial sums are increasing. The even partial sums are bounded below also. (Why?) Therefore, the limit of the even partial sums exists. However, it must be the same as the limit of the odd partial sums because of the last equality above. Thus $\lim_{n\to\infty} S_n$ exists and so the series converges. Now I will show later that $\sum_{k=1}^{\infty} \frac{1}{2k}$ and $\sum_{k=1}^{\infty} \frac{1}{2k-1}$ both diverge. Include enough even terms for the sum to exceed 7. Next add in enough odd terms so that the result will be less than 7. Next add enough even terms to exceed 7 and continue doing this. Since $1/k$ converges to 0, this rearrangement of the series must converge to 7. Of course you could also have picked 5 or -8 just as well. In fact, given any number, there is a rearrangement of this series which converges to this number.

Theorem 5.2.5 *(comparison test) Suppose $\{a_n\}$ and $\{b_n\}$ are sequences of non negative real numbers and suppose for all n sufficiently large, $a_n \leq b_n$. Then*

1. *If $\sum_{n=k}^{\infty} b_n$ converges, then $\sum_{n=m}^{\infty} a_n$ converges.*

2. *If $\sum_{n=k}^{\infty} a_n$ diverges, then $\sum_{n=m}^{\infty} b_n$ diverges.*

Proof: Consider the first claim. From the assumption, there exists n^* such that $n^* > \max(k,m)$ and for all $n \geq n^*$ $b_n \geq a_n$. Then if $p \geq n^*$,

$$\sum_{n=m}^{p} a_n \leq \sum_{n=m}^{n^*} a_n + \sum_{n=n^*+1}^{k} b_n \leq \sum_{n=m}^{n^*} a_n + \sum_{n=k}^{\infty} b_n.$$

Thus the sequence,$\{\sum_{n=m}^{p} a_n\}_{p=m}^{\infty}$ is bounded above and increasing. Therefore, it converges by completeness. The second claim is left as an exercise. ■

Example 5.2.6 *Determine the convergence of $\sum_{n=1}^{\infty} \frac{1}{n^2}$.*

For $n > 1$, $\frac{1}{n^2} \leq \frac{1}{n(n-1)}$. Now $\sum_{n=2}^{p} \frac{1}{n(n-1)} = \sum_{n=2}^{p} \left[\frac{1}{n-1} - \frac{1}{n}\right] = 1 - \frac{1}{p} \to 1$. Therefore, use the comparison test with $a_n = \frac{1}{n^2}$ and $b_n = \frac{1}{n(n-1)}$

A convenient way to implement the comparison test is to use the limit comparison test. This is considered next.

Theorem 5.2.7 *Let $a_n, b_n > 0$ and suppose for all n large enough,*

$$0 < a < \frac{a_n}{b_n} \leq \frac{a_n}{b_n} < b < \infty.$$

Then $\sum a_n$ and $\sum b_n$ converge or diverge together.

Proof: Let n^* be such that $n \geq n^*$, then $\frac{a_n}{b_n} > a$ and $\frac{a_n}{b_n} < b$ and so for all such n,$ab_n < a_n < bb_n$ and so the conclusion follows from the comparison test. ■

The following corollary follows right away from the definition of the limit.

Corollary 5.2.8 *Let $a_n, b_n > 0$ and suppose $\lim_{n\to\infty} \frac{a_n}{b_n} = \lambda \in (0,\infty)$. Then $\sum a_n$ and $\sum b_n$ converge or diverge together.*

Example 5.2.9 *Determine the convergence of $\sum_{k=1}^{\infty} \frac{1}{\sqrt{n^4+2n+7}}$.*

This series converges by the limit comparison test above. Compare with the series of Example 5.2.6.

$$\lim_{n\to\infty}\frac{\left(\frac{1}{n^2}\right)}{\left(\frac{1}{\sqrt{n^4+2n+7}}\right)}=\lim_{n\to\infty}\frac{\sqrt{n^4+2n+7}}{n^2}=\lim_{n\to\infty}\sqrt{1+\frac{2}{n^3}+\frac{7}{n^4}}=1.$$

Therefore, the series converges with the series of Example 5.2.6. How did I know what to compare with? I noticed that $\sqrt{n^4+2n+7}$ is essentially like $\sqrt{n^4}=n^2$ for large enough n. You see, the higher order term n^4 dominates the other terms in n^4+2n+7. Therefore, reasoning that $1/\sqrt{n^4+2n+7}$ is a lot like $1/n^2$ for large n, it was easy to see what to compare with. Of course this is not always easy and there is room for acquiring skill through practice.

To really exploit this limit comparison test, it is desirable to get lots of examples of series, some which converge and some which do not. The tool for obtaining these examples here will be the following wonderful theorem known as the Cauchy condensation test.

Theorem 5.2.10 *Let $a_n\geq 0$ and suppose the terms of the sequence $\{a_n\}$ are decreasing. Thus $a_n\geq a_{n+1}$ for all n. Then*

$$\sum_{n=1}^{\infty}a_n \text{ and } \sum_{n=0}^{\infty}2^n a_{2^n}$$

converge or diverge together.

Proof: This follows from the inequality of the following claim.
Claim:

$$\sum_{k=1}^{n}2^k a_{2^{k-1}}\geq\sum_{k=1}^{2^n}a_k\geq\sum_{k=0}^{n}2^{k-1}a_{2^k}. \tag{5.6}$$

Proof of the Claim: Note the claim is true for $n=1$. Suppose the claim is true for n. Then, since $2^{n+1}-2^n=2^n$, and the terms a_n, are decreasing,

$$\sum_{k=1}^{n+1}2^k a_{2^{k-1}}=2^{n+1}a_{2^n}+\sum_{k=1}^{n}2^k a_{2^{k-1}}\geq 2^{n+1}a_{2^n}+\sum_{k=1}^{2^n}a_k$$

$$\geq\sum_{k=1}^{2^{n+1}}a_k\geq 2^n a_{2^{n+1}}+\sum_{k=1}^{2^n}a_k\geq 2^n a_{2^{n+1}}+\sum_{k=0}^{n}2^{k-1}a_{2^k}=\sum_{k=0}^{n+1}2^{k-1}a_{2^k}.$$

By induction, the claim is valid. Then passing to a limit in 5.6

$$2\sum_{k=0}^{\infty}2^k a_{2^k}=\sum_{k=1}^{\infty}2^k a_{2^{k-1}}\geq\sum_{k=1}^{\infty}a_k\geq\sum_{k=0}^{\infty}2^{k-1}a_{2^k}=\frac{1}{2}\sum_{k=0}^{\infty}2^k a_{2^k}$$

Thus, if $\sum_{k=0}^{\infty}2^k a_{2^k}<\infty$ then the partial sums of $\sum_{k=1}^{\infty}a_k$ are bounded above by $\sum_{k=0}^{\infty}2^k a_{2^k}$ so these partial sums converge. If $\sum_{k=0}^{\infty}2^k a_{2^k}$ diverges, then

$$\infty=\lim_{n\to\infty}\frac{1}{2}\sum_{k=0}^{n}2^k a_{2^k}\leq\lim_{n\to\infty}\sum_{k=1}^{n}a_k$$

and so $\sum a_k$ also diverges. Thus the two series converge or diverge together. ■

Example 5.2.11 *Determine the convergence of* $\sum_{k=1}^{\infty} \frac{1}{k^p}$ *where* p *is a positive number. These are called the* p *series.*

Let $a_n = \frac{1}{n^p}$. Then $a_{2^n} = \left(\frac{1}{2^p}\right)^n$. From the Cauchy condensation test the two series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ and } \sum_{n=0}^{\infty} 2^n \left(\frac{1}{2^p}\right)^n = \sum_{n=0}^{\infty} \left(2^{(1-p)}\right)^n$$

converge or diverge together. If $p > 1$, the last series above is a geometric series having common ratio less than 1 and so it converges. If $p \leq 1$, it is still a geometric series but in this case the common ratio is either 1 or greater than 1 so the series diverges. It follows that the p series converges if $p > 1$ and diverges if $p \leq 1$. In particular, $\sum_{n=1}^{\infty} n^{-1}$ diverges while $\sum_{n=1}^{\infty} n^{-2}$ converges.

$$\begin{array}{lll} \sum_{n=1}^{\infty} \frac{1}{n^p} & p > 1 & \text{converges} \\ \\ \sum_{n=1}^{\infty} \frac{1}{n^p} & p \leq 1 & \text{diverges} \end{array}$$

Example 5.2.12 *Determine the convergence of* $\sum_{k=1}^{\infty} \frac{1}{\sqrt{n^2+100n}}$.

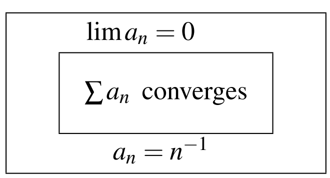
Use the limit comparison test. $\lim_{n\to\infty} \frac{\left(\frac{1}{n}\right)}{\left(\frac{1}{\sqrt{n^2+100n}}\right)} = 1$ and so this series diverges with $\sum_{k=1}^{\infty} \frac{1}{k}$.

Sometimes it is good to be able to say a series does not converge. The n^{th} term test gives such a condition which is sufficient for this. It is really a corollary of Theorem 5.1.7.

Theorem 5.2.13 *If* $\sum_{n=m}^{\infty} a_n$ *converges, then* $\lim_{n\to\infty} a_n = 0$.

Proof: Apply Theorem 5.1.7 to conclude that $\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sum_{k=n}^{n} a_k = 0$. ■

It is very important to observe that this theorem goes only in one direction. That is, you cannot conclude the series converges if $\lim_{n\to\infty} a_n = 0$. If this happens, you don't know anything from this information. Recall $\lim_{n\to\infty} n^{-1} = 0$ but $\sum_{n=1}^{\infty} n^{-1}$ diverges. The following picture is descriptive of the situation.



5.3 Exercises

1. Determine whether the following series converge and give reasons for your answers.

(a) $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2+n+1}}$

(b) $\sum_{n=1}^{\infty} \left(\sqrt{n+1}-\sqrt{n}\right)$

(c) $\sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!}$

(d) $\sum_{n=1}^{\infty} \frac{(2n)!}{(n!)^2}$

(e) $\sum_{n=1}^{\infty} \frac{1}{2n+2}$

(f) $\sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^n$

(g) $\sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{n^2}$

2. Determine whether the following series converge and give reasons for your answers.

(a) $\sum_{n=1}^{\infty} \frac{2^n+n}{n2^n}$

(b) $\sum_{n=1}^{\infty} \frac{2^n+n}{n^2 2^n}$

(c) $\sum_{n=1}^{\infty} \frac{n}{2n+1}$

(d) $\sum_{n=1}^{\infty} \frac{n^{100}}{1.01^n}$

3. Find the exact values of the following infinite series if they converge.

(a) $\sum_{k=3}^{\infty} \frac{1}{k(k-2)}$

(b) $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$

(c) $\sum_{k=3}^{\infty} \frac{1}{(k+1)(k-2)}$

(d) $\sum_{k=1}^{N} \left(\frac{1}{\sqrt{k}} - \frac{1}{\sqrt{k+1}}\right)$

4. Suppose $\sum_{k=1}^{\infty} a_k$ converges and each $a_k \geq 0$. Does it follow that $\sum_{k=1}^{\infty} a_k^2$ also converges?

5. Find a series which diverges using one test but converges using another if possible. If this is not possible, tell why.

6. If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ both converge and a_n, b_n are nonnegative, can you conclude the sum, $\sum_{n=1}^{\infty} a_n b_n$ converges?

7. If $\sum_{n=1}^{\infty} a_n$ converges and $a_n \geq 0$ for all n and b_n is bounded, can you conclude $\sum_{n=1}^{\infty} a_n b_n$ converges?

8. Determine the convergence of the series $\sum_{n=1}^{\infty} \left(\sum_{k=1}^{n} \frac{1}{k}\right)^{-n/2}$.

9. Is it possible there could exist a decreasing sequence of positive numbers, $\{a_n\}$ such that $\lim_{n\to\infty} a_n = 0$ but $\sum_{n=1}^{\infty} \left(1 - \frac{a_{n+1}}{a_n}\right)$ converges? (This seems to be a fairly difficult problem.)**Hint:** You might do something like this. Show

$$\lim_{x\to 1} \frac{1-x}{-\ln(x)} = \frac{1-x}{\ln(1/x)} = 1$$

Next use a limit comparison test with $\sum_{n=1}^{\infty} \ln\left(\frac{a_n}{a_{n+1}}\right)$ Go ahead and use what you learned in calculus about ln and any other techniques for finding limits. These things will be discussed better later on, but you have seen them in calculus so this is a little review.

10. Suppose $\sum a_n$ converges conditionally and each a_n is real. Show it is possible to add the series in some order such that the result converges to 13. Then show it is possible to add the series in another order so that the result converges to 7. Thus there is no generalization of the commutative law for conditionally convergent infinite series. **Hint:** To see how to proceed, consider Example 5.2.4.

5.4 More Tests for Convergence

5.4.1 Convergence Because of Cancellation

So far, the tests for convergence have been applied to non negative terms only. Sometimes, a series converges, not because the terms of the series get small fast enough, but because of cancellation taking place between positive and negative terms. A discussion of this involves some simple algebra.

Let $\{a_n\}$ and $\{b_n\}$ be sequences and let $A_n \equiv \sum_{k=1}^{n} a_k$, $A_{-1} \equiv A_0 \equiv 0$. Then if $p < q$

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} b_n (A_n - A_{n-1}) = \sum_{n=p}^{q} b_n A_n - \sum_{n=p}^{q} b_n A_{n-1}$$

$$= \sum_{n=p}^{q} b_n A_n - \sum_{n=p-1}^{q-1} b_{n+1} A_n = b_q A_q - b_p A_{p-1} + \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) \tag{5.7}$$

This formula is called the partial summation formula of Dirichlet.

Theorem 5.4.1 *(Dirichlet's test) Suppose $A_n \equiv \sum_{k=1}^{n} a_k$ is bounded and also that $\lim_{n\to\infty} b_n = 0$, with $b_n \geq b_{n+1}$ for all n. Then $\sum_{n=1}^{\infty} a_n b_n$ converges.*

Proof: This follows quickly from Theorem 5.1.7. Indeed, letting $|A_n| \leq C$, and using the partial summation formula above along with the assumption that the b_n are decreasing,

$$\left| \sum_{n=p}^{q} a_n b_n \right| = \left| b_q A_q - b_p A_{p-1} + \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) \right|$$

$$\leq C\left(\left|b_q\right| + \left|b_p\right|\right) + C \sum_{n=p}^{q-1} (b_n - b_{n+1}) = C\left(\left|b_q\right| + \left|b_p\right|\right) + C (b_p - b_q)$$

and by assumption, this last expression is small whenever p and q are sufficiently large. ■

Definition 5.4.2 *If $b_n > 0$ for all n, a series of the form $\sum_k (-1)^k b_k$ or $\sum_k (-1)^{k-1} b_k$ is known as an alternating series.*

The following corollary is known as the alternating series test.

Corollary 5.4.3 *(alternating series test) If $\lim_{n\to\infty} b_n = 0$, with $b_n \geq b_{n+1}$, then the series, $\sum_{n=1}^{\infty} (-1)^n b_n$ converges.*

Proof: Let $a_n = (-1)^n$. Then the partial sums of $\sum_n a_n$ are bounded and so Theorem 5.4.1 applies. ■

In the situation of Corollary 5.4.3 there is a convenient error estimate available.

Theorem 5.4.4 *Let $b_n > 0$ for all n such that $b_n \geq b_{n+1}$ for all n and $\lim_{n\to\infty} b_n = 0$. and consider either $\sum_{n=1}^{\infty} (-1)^n b_n$ or $\sum_{n=1}^{\infty} (-1)^{n-1} b_n$. Then*

$$\begin{aligned} \left| \sum_{n=1}^{\infty} (-1)^n b_n - \sum_{n=1}^{N} (-1)^n b_n \right| &\leq |b_{N+1}|, \\ \left| \sum_{n=1}^{\infty} (-1)^{n-1} b_n - \sum_{n=1}^{N} (-1)^{n-1} b_n \right| &\leq |b_{N+1}| \end{aligned}$$

See Problem 8 on Page 99 for an outline of the proof of this theorem along with another way to prove the alternating series test.

Example 5.4.5 *How many terms must I take in the sum,* $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n^2+1}$ *to be closer than* $\frac{1}{10}$ *to* $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n^2+1}$*?*

From Theorem 5.4.4, I need to find n such that $\frac{1}{n^2+1} \leq \frac{1}{10}$ and then $n-1$ is the desired value. Thus $n = 3$ and so $\left|\sum_{n=1}^{\infty}(-1)^n \frac{1}{n^2+1} - \sum_{n=1}^{2}(-1)^n \frac{1}{n^2+1}\right| \leq \frac{1}{10}$

Definition 5.4.6 *A series $\sum a_n$ is said to converge absolutely if $\sum |a_n|$ converges. It is said to converge conditionally if $\sum |a_n|$ fails to converge but $\sum a_n$ converges.*

Thus the alternating series or more general Dirichlet test can determine convergence of series which converge conditionally.

5.4.2 Ratio And Root Tests

A favorite test for convergence is the ratio test. This is discussed next. It is at the other extreme from the alternating series test, being completely oblivious to any sort of cancellation. It only gives absolute convergence or spectacular divergence.

Theorem 5.4.7 *Suppose $|a_n| > 0$ for all n and suppose $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = r$. Then*

$$\sum_{n=1}^{\infty} a_n \begin{cases} \text{diverges if } r > 1 \\ \text{converges absolutely if } r < 1 \\ \text{test fails if } r = 1 \end{cases}.$$

Proof: Suppose $r < 1$. Then there exists n_1 such that if $n \geq n_1$, then

$$0 < \left|\frac{a_{n+1}}{a_n}\right| < R$$

where $r < R < 1$. Then $|a_{n+1}| < R|a_n|$ for all such n. Therefore,

$$\left|a_{n_1+p}\right| < R\left|a_{n_1+p-1}\right| < R^2\left|a_{n_1+p-2}\right| < \cdots < R^p\left|a_{n_1}\right| \tag{5.8}$$

and so if $m > n$, then $|a_m| < R^{m-n_1}\left|a_{n_1}\right|$. By the comparison test and the theorem on geometric series, $\sum |a_n|$ converges. This proves the convergence part of the theorem.

To verify the divergence part, note that if $r > 1$, then 5.8 can be turned around for some $R > 1$. Showing $\lim_{n\to\infty}|a_n| = \infty$. Since the n^{th} term fails to converge to 0, it follows the series diverges.

To see the test fails if $r = 1$, consider $\sum n^{-1}$ and $\sum n^{-2}$. The first series diverges while the second one converges but in both cases, $r = 1$. (Be sure to check this last claim.) ■

The ratio test is very useful for many different examples but it is somewhat unsatisfactory mathematically. One reason for this is the assumption that $a_n \neq 0$, necessitated by the need to divide by a_n, and the other reason is the possibility that the limit might not exist. The next test, called the root test removes both of these objections. Before presenting this test, it is necessary to first prove the existence of the p^{th} root of any positive number. This was shown earlier in Theorem 2.11.2 but the following lemma gives an easier treatment of this issue based on theorems about sequences.

Lemma 5.4.8 *Let $\alpha > 0$ be any nonnegative number and let $p \in \mathbb{N}$. Then $\alpha^{1/p}$ exists. This is the unique positive number which when raised to the p^{th} power gives α.*

Proof: Consider the function $f(x) \equiv x^p - \alpha$. Then there exists b_1 such that $f(b_1) > 0$ and a_1 such that $f(a_1) < 0$. (Why?) Now cut the interval $[a_1, b_1]$ into two closed intervals of equal length. Let $[a_2, b_2]$ be one of these which has $f(a_2) f(b_2) \leq 0$. Now do for $[a_2, b_2]$ the same thing which was done to get $[a_2, b_2]$ from $[a_1, b_1]$. Continue this way obtaining a sequence of nested intervals $[a_k, b_k]$ with the property that

$$b_k - a_k = 2^{1-k}(b_1 - a_1).$$

By the nested interval theorem, there exists a unique point x in all these intervals. By Theorem 4.4.8

$$f(a_k) \to f(x),\ f(b_k) \to f(x).$$

Then from Theorem 4.4.13,

$$f(x) f(x) = \lim_{k\to\infty} f(a_k) f(b_k) \leq 0$$

Hence $f(x) = 0$. ■

Theorem 5.4.9 *Suppose $|a_n|^{1/n} < R < 1$ for all n sufficiently large. Then*

$$\sum_{n=1}^{\infty} a_n \text{ converges absolutely.}$$

If there are infinitely many values of n such that $|a_n|^{1/n} \geq 1$, then

$$\sum_{n=1}^{\infty} a_n \text{ diverges.}$$

Proof: Suppose first that $|a_n|^{1/n} < R < 1$ for all n sufficiently large. Say this holds for all $n \geq n_R$. Then for such n, $\sqrt[n]{|a_n|} < R$. Therefore, for such n, $|a_n| \leq R^n$ and so the comparison test with a geometric series applies and gives absolute convergence as claimed.

Next suppose $|a_n|^{1/n} \geq 1$ for infinitely many values of n. Then for those values of n, $|a_n| \geq 1$ and so the series fails to converge by the n^{th} term test. ■

Stated more succinctly the condition for the root test is this: Let $r = \limsup_{n\to\infty} |a_n|^{1/n}$ then

$$\sum_{k=m}^{\infty} a_k \begin{cases} \text{converges absolutely if } r < 1 \\ \text{test fails if } r = 1 \\ \text{diverges if } r > 1 \end{cases}$$

To see the test fails when $r = 1$, consider the same example given above, $\sum_n \frac{1}{n}$ and $\sum_n \frac{1}{n^2}$. Indeed, $\lim_{n\to\infty} n^{1/n} = 1$. To see this, let $e_n = n^{1/n} - 1$ so $(1 + e_n)^n = n$ By the binomial theorem, $1 + ne_n + \frac{n(n-1)}{2} e_n^2 \leq n$ and so $e_n^2 \leq \frac{2n}{n(n-1)}$ showing that $e_n \to 0$.

A special case occurs when the limit exists.

Corollary 5.4.10 *Suppose $\lim_{n\to\infty} |a_n|^{1/n}$ exists and equals r. Then*

$$\sum_{k=m}^{\infty} a_k \begin{cases} \textit{converges absolutely if } r < 1 \\ \textit{test fails if } r = 1 \\ \textit{diverges if } r > 1 \end{cases}$$

Proof: The first and last alternatives follow from Theorem 5.4.9. To see the test fails if $r = 1$, consider the two series $\sum_{n=1}^{\infty} \frac{1}{n}$ and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ both of which have $r = 1$ but having different convergence properties. The first diverges and the second converges. ■

5.5 Double Series

Sometimes it is required to consider double series which are of the form

$$\sum_{k=m}^{\infty} \sum_{j=m}^{\infty} a_{jk} \equiv \sum_{k=m}^{\infty} \left(\sum_{j=m}^{\infty} a_{jk} \right).$$

In other words, first sum on j yielding something which depends on k and then sum these. The major consideration for these double series is the question of when

$$\sum_{k=m}^{\infty} \sum_{j=m}^{\infty} a_{jk} = \sum_{j=m}^{\infty} \sum_{k=m}^{\infty} a_{jk}.$$

In other words, when does it make no difference which subscript is summed over first? In the case of finite sums there is no issue here. You can always write

$$\sum_{k=m}^{M} \sum_{j=m}^{N} a_{jk} = \sum_{j=m}^{N} \sum_{k=m}^{M} a_{jk}$$

because addition is commutative. However, there are limits involved with infinite sums and the interchange in order of summation involves taking limits in a different order. Therefore, it is not always true that it is permissible to interchange the two sums. A general rule of thumb is this: If something involves changing the order in which two limits are taken, you may not do it without agonizing over the question. In general, limits foul up algebra and also introduce things which are counter intuitive. Here is an example. This example is a little technical. It is placed here just to prove conclusively there is a question which needs to be considered.

Example 5.5.1 *Consider the following picture which depicts some of the ordered pairs (m,n) where m,n are positive integers.*

$$\begin{array}{ccccc} & & \vdots & & \\ 0 & 0 & c & 0 & -c \\ 0 & c & 0 & -c & 0 \\ b & 0 & -c & 0 & 0 \\ 0 & a & 0 & 0 & 0 \end{array} \cdots$$

The a,b,c are the values of a_{mn}. Thus $a_{nn} = 0$ for all $n \geq 1$, $a_{21} = a, a_{12} = b, a_{m(m+1)} = -c$ whenever $m > 1$, and $a_{m(m-1)} = c$ whenever $m > 2$. The numbers next to the point are the values of a_{mn}. You see $a_{nn} = 0$ for all n, $a_{21} = a, a_{12} = b, a_{mn} = c$ for (m,n) on the line $y = 1 + x$ whenever $m > 1$, and $a_{mn} = -c$ for all (m,n) on the line $y = x - 1$ whenever $m > 2$.

Then $\sum_{m=1}^{\infty} a_{mn} = a$ if $n = 1$, $\sum_{m=1}^{\infty} a_{mn} = b - c$ if $n = 2$ and if $n > 2, \sum_{m=1}^{\infty} a_{mn} = 0$. Therefore, $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{mn} = a + b - c$. Next observe that $\sum_{n=1}^{\infty} a_{mn} = b$ if $m = 1, \sum_{n=1}^{\infty} a_{mn} =$

$a+c$ if $m=2$, and $\sum_{n=1}^{\infty} a_{mn}=0$ if $m>2$. Therefore, $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}=b+a+c$ and so the two sums are different. Moreover, you can see that by assigning different values of a,b, and c, you can get an example for any two different numbers desired.

Don't become upset by this. It happens because, as indicated above, limits are taken in two different orders. An infinite sum always involves a limit and this illustrates why you must always remember this. This example in no way violates the commutative law of addition which has nothing to do with limits. Algebra is not analysis. Crazy things happen when you take limits. Intuition is routinely rendered useless.

However, it turns out that if $a_{ij}\geq 0$ for all i,j, then you can always interchange the order of summation. This is shown next and is based on the Lemma 2.10.5 which says you can intercange supremums.

Lemma 5.5.2 *If $\{A_n\}$ is an increasing sequence in $[-\infty,\infty]$, then* $\sup\{A_n\}=\lim_{n\to\infty}A_n$.

Proof: Let $\sup(\{A_n:n\in\mathbb{N}\})=r$. In the first case, suppose $r<\infty$. Then letting $\varepsilon>0$ be given, there exists n such that $A_n\in(r-\varepsilon,r]$. Since $\{A_n\}$ is increasing, it follows if $m>n$, then $r-\varepsilon<A_n\leq A_m\leq r$ and so $\lim_{n\to\infty}A_n=r$ as claimed. In the case where $r=\infty$, then if a is a real number, there exists n such that $A_n>a$. Since $\{A_k\}$ is increasing, it follows that if $m>n$, $A_m>a$. But this is what is meant by $\lim_{n\to\infty}A_n=\infty$. The other case is that $r=-\infty$. But in this case, $A_n=-\infty$ for all n and so $\lim_{n\to\infty}A_n=-\infty$. ■

Theorem 5.5.3 *Let $a_{ij}\geq 0$. Then $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{ij}=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}$.*

Proof: First note there is no trouble in defining these sums because the a_{ij} are all nonnegative. If a sum diverges, it only diverges to ∞ and so ∞ is the value of the sum. Next note that $\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}a_{ij}\geq\sup_n\sum_{j=r}^{\infty}\sum_{i=r}^{n}a_{ij}$ because for all j, $\sum_{i=r}^{\infty}a_{ij}\geq\sum_{i=r}^{n}a_{ij}$.Therefore, using Lemma 5.5.2,

$$\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}a_{ij}\geq\sup_n\sum_{j=r}^{\infty}\sum_{i=r}^{n}a_{ij}=\sup_n\lim_{m\to\infty}\sum_{j=r}^{m}\sum_{i=r}^{n}a_{ij}$$

$$\begin{aligned}
&= \sup_n\lim_{m\to\infty}\sum_{i=r}^{n}\sum_{j=r}^{m}a_{ij}=\sup_n\sum_{i=r}^{n}\lim_{m\to\infty}\sum_{j=r}^{m}a_{ij}\\
&= \sup_n\sum_{i=r}^{n}\sum_{j=r}^{\infty}a_{ij}=\lim_{n\to\infty}\sum_{i=r}^{n}\sum_{j=r}^{\infty}a_{ij}=\sum_{i=r}^{\infty}\sum_{j=r}^{\infty}a_{ij}
\end{aligned}$$

Interchanging the i and j in the above argument proves the theorem. ■

The following is the fundamental result on double sums.

Theorem 5.5.4 *Let $a_{ij}\in\mathbb{F}$ and suppose $\sum_{i=r}^{\infty}\sum_{j=r}^{\infty}|a_{ij}|<\infty$. Then $\sum_{i=r}^{\infty}\sum_{j=r}^{\infty}a_{ij}=\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}a_{ij}$ and every infinite sum encountered in the above equation converges.*

Proof: By Theorem 5.5.3 $\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}|a_{ij}|=\sum_{i=r}^{\infty}\sum_{j=r}^{\infty}|a_{ij}|<\infty$. Therefore, for each j, $\sum_{i=r}^{\infty}|a_{ij}|<\infty$ and for each i, $\sum_{j=r}^{\infty}|a_{ij}|<\infty$. By Theorem 5.2.2 on Page 87, both of the series $\sum_{i=r}^{\infty}a_{ij}$, $\sum_{j=r}^{\infty}a_{ij}$ converge, the first one for every j and the second for every i. Also, $\sum_{j=r}^{\infty}\left|\sum_{i=r}^{\infty}a_{ij}\right|\leq\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}|a_{ij}|<\infty$ and $\sum_{i=r}^{\infty}\left|\sum_{j=r}^{\infty}a_{ij}\right|\leq\sum_{i=r}^{\infty}\sum_{j=r}^{\infty}|a_{ij}|<\infty$ so by Theorem 5.2.2 again, $\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}a_{ij}$, $\sum_{i=r}^{\infty}\sum_{j=r}^{\infty}a_{ij}$ both exist. It only remains to verify they are equal. By similar reasoning you can replace a_{ij} with $\operatorname{Re} a_{ij}$ or with $\operatorname{Im} a_{ij}$ in the above and the two sums will exist.

The real part of a finite sum of complex numbers equals the sum of the real parts. Then passing to a limit, it follows $\operatorname{Re}\sum_{j=r}^{\infty}\sum_{i=r}^{\infty} a_{ij} = \sum_{j=r}^{\infty}\sum_{i=r}^{\infty} \operatorname{Re} a_{ij}$ and similarly, one can conclude that $\operatorname{Im}\sum_{i=r}^{\infty}\sum_{j=r}^{\infty} a_{ij} = \sum_{i=r}^{\infty}\sum_{j=r}^{\infty} \operatorname{Im} a_{ij}$. Note $0 \leq \left(\left|a_{ij}\right| + \operatorname{Re} a_{ij}\right) \leq 2\left|a_{ij}\right|$. Therefore, by Theorem 5.5.3 and Theorem 5.1.5 on Page 86

$$\sum_{j=r}^{\infty}\sum_{i=r}^{\infty}\left|a_{ij}\right| + \sum_{j=r}^{\infty}\sum_{i=r}^{\infty}\operatorname{Re} a_{ij} = \sum_{j=r}^{\infty}\sum_{i=r}^{\infty}\left(\left|a_{ij}\right| + \operatorname{Re} a_{ij}\right)$$

$$\begin{aligned} &= \sum_{i=r}^{\infty}\sum_{j=r}^{\infty}\left(\left|a_{ij}\right| + \operatorname{Re} a_{ij}\right) = \sum_{i=r}^{\infty}\sum_{j=r}^{\infty}\left|a_{ij}\right| + \sum_{i=r}^{\infty}\sum_{j=r}^{\infty}\operatorname{Re} a_{ij} \\ &= \sum_{j=r}^{\infty}\sum_{i=r}^{\infty}\left|a_{ij}\right| + \sum_{i=r}^{\infty}\sum_{j=r}^{\infty}\operatorname{Re} a_{ij} \end{aligned}$$

and so

$$\operatorname{Re}\sum_{j=r}^{\infty}\sum_{i=r}^{\infty} a_{ij} = \sum_{j=r}^{\infty}\sum_{i=r}^{\infty}\operatorname{Re} a_{ij} = \sum_{i=r}^{\infty}\sum_{j=r}^{\infty}\operatorname{Re} a_{ij} = \operatorname{Re}\sum_{i=r}^{\infty}\sum_{j=r}^{\infty} a_{ij}$$

Similar reasoning applies to the imaginary parts. Since the real and imaginary parts of the two series are equal, it follows the two series are equal. ■

One of the most important applications of this theorem is to the problem of multiplication of series.

Definition 5.5.5 *Let $\sum_{i=r}^{\infty} a_i$ and $\sum_{i=r}^{\infty} b_i$ be two series. For $n \geq r$, define*

$$c_n \equiv \sum_{k=r}^{n} a_k b_{n-k+r}.$$

The series $\sum_{n=r}^{\infty} c_n$ is called the Cauchy product of the two series.

It isn't hard to see where this comes from. Formally write the following in the case $r = 0$:

$$(a_0 + a_1 + a_2 + a_3 \cdots)(b_0 + b_1 + b_2 + b_3 \cdots)$$

and start multiplying in the usual way. This yields

$$a_0 b_0 + (a_0 b_1 + b_0 a_1) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \cdots$$

and you see the expressions in parentheses above are just the c_n for $n = 0, 1, 2, \cdots$. Therefore, it is reasonable to conjecture that $\sum_{i=r}^{\infty} a_i \sum_{j=r}^{\infty} b_j = \sum_{n=r}^{\infty} c_n$ and of course there would be no problem with this in the case of finite sums but in the case of infinite sums, it is necessary to prove a theorem. The following is a special case of Merten's theorem.

Theorem 5.5.6 *Suppose $\sum_{i=r}^{\infty} a_i$ and $\sum_{j=r}^{\infty} b_j$ both converge absolutely*[1]*. Then*

$$\left(\sum_{i=r}^{\infty} a_i\right)\left(\sum_{j=r}^{\infty} b_j\right) = \sum_{n=r}^{\infty} c_n$$

where $c_n = \sum_{k=r}^{n} a_k b_{n-k+r}$.

[1] Actually, it is only necessary to assume one of the series converges and the other converges absolutely. This is known as Merten's theorem and may be read in the 1974 book by Apostol listed in the bibliography.

Proof: Let $p_{nk} = 1$ if $r \leq k \leq n$ and $p_{nk} = 0$ if $k > n$. Then $c_n = \sum_{k=r}^{\infty} p_{nk} a_k b_{n-k+r}$. Also,

$$\sum_{k=r}^{\infty} \sum_{n=r}^{\infty} p_{nk} |a_k| |b_{n-k+r}| = \sum_{k=r}^{\infty} |a_k| \sum_{n=r}^{\infty} p_{nk} |b_{n-k+r}|$$

$$= \sum_{k=r}^{\infty} |a_k| \sum_{n=k}^{\infty} |b_{n-k+r}| = \sum_{k=r}^{\infty} |a_k| \sum_{n=k}^{\infty} \left| b_{n-(k-r)} \right| = \sum_{k=r}^{\infty} |a_k| \sum_{m=r}^{\infty} |b_m| < \infty.$$

Therefore, by Theorem 5.5.4

$$\sum_{n=r}^{\infty} c_n = \sum_{n=r}^{\infty} \sum_{k=r}^{n} a_k b_{n-k+r} = \sum_{n=r}^{\infty} \sum_{k=r}^{\infty} p_{nk} a_k b_{n-k+r}$$

$$= \sum_{k=r}^{\infty} a_k \sum_{n=r}^{\infty} p_{nk} b_{n-k+r} = \sum_{k=r}^{\infty} a_k \sum_{n=k}^{\infty} b_{n-k+r} = \sum_{k=r}^{\infty} a_k \sum_{m=r}^{\infty} b_m \blacksquare$$

5.6 Exercises

1. Determine whether the following series converge absolutely, conditionally, or not at all and give reasons for your answers.

   (a) $\sum_{n=1}^{\infty} (-1)^n \frac{2^n+n}{n2^n}$

   (b) $\sum_{n=1}^{\infty} (-1)^n \frac{2^n+n}{n^2 2^n}$

   (c) $\sum_{n=1}^{\infty} (-1)^n \frac{n}{2n+1}$

   (d) $\sum_{n=1}^{\infty} (-1)^n \frac{10^n}{n!}$

   (e) $\sum_{n=1}^{\infty} (-1)^n \frac{n^{100}}{1.01^n}$

   (f) $\sum_{n=1}^{\infty} (-1)^n \frac{3^n}{n^3}$

   (g) $\sum_{n=1}^{\infty} (-1)^n \frac{n^3}{3^n}$

   (h) $\sum_{n=1}^{\infty} (-1)^n \frac{n^3}{n!}$

   (i) $\sum_{n=1}^{\infty} (-1)^n \frac{n!}{n^{100}}$

2. Suppose $\sum_{n=1}^{\infty} a_n$ converges. Can the same thing be said about $\sum_{n=1}^{\infty} a_n^2$? Explain.

3. A person says a series converges conditionally by the ratio test. Explain why his statement is total nonsense.

4. A person says a series diverges by the alternating series test. Explain why his statement is total nonsense.

5. Find a series which diverges using one test but converges using another if possible. If this is not possible, tell why.

6. If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ both converge, can you conclude the sum, $\sum_{n=1}^{\infty} a_n b_n$ converges?

7. If $\sum_{n=1}^{\infty} a_n$ converges absolutely, and b_n is bounded, does $\sum_{n=1}^{\infty} a_n b_n$ always converge? What if it is only the case that $\sum_{n=1}^{\infty} a_n$ converges?

8. Prove Theorem 5.4.4. **Hint:** For $\sum_{n=1}^{\infty} (-1)^n b_n$, show the odd partial sums are all no larger than $\sum_{n=1}^{\infty} (-1)^n b_n$ and are increasing while the even partial sums are at least as large as $\sum_{n=1}^{\infty} (-1)^n b_n$ and are decreasing. Use this to give another proof of the alternating series test. If you have trouble, see most standard calculus books.

9. Use Theorem 5.4.4 in the following alternating series to tell how large n must be so that $\left|\sum_{k=1}^{\infty}(-1)^k a_k - \sum_{k=1}^{n}(-1)^k a_k\right|$ is no larger than the given number.

   (a) $\sum_{k=1}^{\infty}(-1)^k \frac{1}{k}$, .001

   (b) $\sum_{k=1}^{\infty}(-1)^k \frac{1}{k^2}$, .001

   (c) $\sum_{k=1}^{\infty}(-1)^{k-1} \frac{1}{\sqrt{k}}$, .001

10. Consider the series $\sum_{k=0}^{\infty}(-1)^n \frac{1}{\sqrt{n+1}}$. Show this series converges and so it makes sense to write $\left(\sum_{k=0}^{\infty}(-1)^n \frac{1}{\sqrt{n+1}}\right)^2$. What about the Cauchy product of this series? Does it even converge? What does this mean about using algebra on infinite sums as though they were finite sums?

11. Verify Theorem 5.5.6 on the two series $\sum_{k=0}^{\infty} 2^{-k}$ and $\sum_{k=0}^{\infty} 3^{-k}$.

12. You can define infinite series of complex numbers in exactly the same way as infinite series of real numbers. That is $w = \sum_{k=1}^{\infty} z_k$ means: For every $\varepsilon > 0$ there exists N such that if $n \geq N$, then $|w - \sum_{k=1}^{n} z_k| < \varepsilon$. Here the absolute value is the one which applies to complex numbers. That is, $|a+ib| = \sqrt{a^2+b^2}$. Show that if $\{a_n\}$ is a decreasing sequence of nonnegative numbers with the property that $\lim_{n\to\infty} a_n = 0$ and if ω is any complex number which is not equal to 1 but which satisfies $|\omega| = 1$, then $\sum_{n=1}^{\infty} \omega^n a_n$ must converge. Note a sequence of complex numbers, $\{a_n + ib_n\}$ converges to $a+ib$ if and only if $a_n \to a$ and $b_n \to b$. See Problem 6 on Page 66. There are quite a few things in this problem you should think about.

13. Suppose $\lim_{k\to\infty} s_k = s$. Show it follows $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} s_k = s$.

14. Using Problem 13 show that if $\sum_{j=1}^{\infty} \frac{a_j}{j}$ converges, then it follows $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} a_j = 0$.

15. Show that if $\{p_i\}_{i=1}^{\infty}$ are the prime numbers, then $\sum_{i=1}^{\infty} \frac{1}{p_i} = \infty$. That is, there are enough primes that the sum of their reciprocals diverges. **Hint:** Let $\pi(n)$ denote the number of primes less than equal to n, $\{p_1, \ldots, p_{\pi(n)}\}$. Then explain why

$$\sum_{k=1}^{n} \frac{1}{k} \leq \left(\sum_{k=1}^{n} \frac{1}{p_1^k}\right) \cdots \left(\sum_{k=1}^{n} \frac{1}{p_{\pi(n)}^k}\right) \leq \prod_{k=1}^{\pi(n)} \frac{1}{1-\frac{1}{p_k}} \leq \prod_{k=1}^{\pi(n)} e^{2/p_k} = e^{2\sum_{k=1}^{\pi(n)} \frac{1}{p_k}}$$

and consequently why $\lim_{n\to\infty} \pi(n) = \infty$ and $\sum_{i=1}^{\infty} \frac{1}{p_i} = \infty$.

16. Verify the allegation about the Euclidean norm $|x| \equiv \left(\sum_{k=1}^{p} |x_k|^2\right)^{1/2}$ that $\mathbb{F}^p$ with the Euclidean norm yields the same Cauchy sequences, compact sets, and open and closed sets as $\mathbb{F}^p$ with the norm $\|\cdot\|$.

17. Suppose S is an uncountable set and suppose $f(s)$ is a positive number for each $s \in S$. Also let $\hat{S}$ denote a finite subset of S. Show that

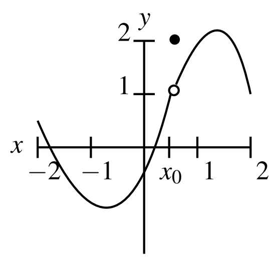
$$\sup\left\{\sum_{s\in\hat{S}} f(s) : \hat{S} \subseteq S\right\} = \infty$$

Chapter 6

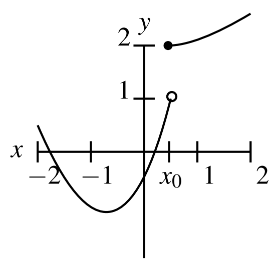
Continuous Functions

The concept of function is far too general to be useful in calculus. There are various ways to restrict the concept in order to study something interesting and the types of restrictions considered depend very much on what you find interesting. In Calculus, the most fundamental restriction made is to assume the functions are continuous. Continuous functions are those in which a sufficiently small change in x results in a small change in $f(x)$. They rule out things which could never happen physically. For example, it is not possible for a car to jump from one point to another instantly. Making this restriction precise turns out to be surprisingly difficult although many of the most important theorems about continuous functions seem intuitively clear.

Before giving the careful mathematical definitions, here are examples of graphs of functions which are not continuous at the point x_0.



You see, there is a hole in the picture of the graph of this function and instead of filling in the hole with the appropriate value, $f(x_0)$ is too large. This is called a removable discontinuity because the problem can be fixed by redefining the function at the point x_0. Here is another example.



You see from this picture that there is no way to get rid of the jump in the graph of this function by simply redefining the value of the function at x_0. That is why it is called a nonremovable discontinuity or jump discontinuity. Now that pictures have been given of what it is desired to eliminate, it is time to give the precise definition.

The definition which follows, due to Bolzano, Cauchy[1] Bolzano, and Weierstrass and

[1]Augustin Louis Cauchy 1789-1857 was the son of a lawyer who was married to an aristocrat. He was born in France just after the fall of the Bastille and his family fled the reign of terror and hid in the countryside till it was over. Cauchy was educated at first by his father who taught him Greek and Latin. Eventually Cauchy learned many languages.

Weierstrass[2] is the precise way to exclude the sort of behavior described above and all statements about continuous functions must ultimately rest on this definition from now on or something which is equivalent to it. I am going to present this in the context of functions which are defined on $D(f) \subseteq \mathbb{F}^p$ having values in $\mathbb{F}^q$ where p, q are positive integers because it is no harder. However, in most of the applications in this book, $D(f)$ will be in $\mathbb{R}$ or $\mathbb{C}$.

Definition 6.0.1 *A function $f : D(f) \subseteq \mathbb{F}^p \to \mathbb{F}^q$ is continuous at $x \in D(f)$ if for each $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $y \in D(f)$ and $\|y - x\| < \delta$ it follows that $\|f(x) - f(y)\| < \varepsilon$. A function f is continuous if it is continuous at every point of $D(f)$.*

If f has values in $\mathbb{F}^p$, it is of the form $x \to (f_1(x), \cdots, f_p(x))$ where the f_i are real valued functions.

In sloppy English this definition says roughly the following: A function f is continuous at x when it is possible to make $f(y)$ as close as desired to $f(x)$ provided y is taken close enough to x. In fact this statement in words is pretty much the way Cauchy described it. The completely rigorous definition above is usually ascribed to Weierstrass.

If you are like me, you may find the following equivalent description of continuity easier to remember and use. I don't have a very good reason why this is so, but it seems to be the case, at least for many people. I will use either definition whenever convenient.

Theorem 6.0.2 *A function f is continuous if and only if whenever $x_n \to x$ with $x_n, x \in D(f)$, it follows that $f(x_n) \to f(x)$. In words, convergent sequences get taken to convergent sequences.*

Proof: $\Rightarrow$ Suppose $x_n \to x$ as described. I need to verify that $f(x_n) \to f(x)$. I know that for any $\varepsilon > 0$ there exists a suitable δ such that the conditions of continuity hold. I also know that, since $x_n \to x$, eventually, for all n large enough, $\|x_n - x\| < \delta$. Therefore, for all n large enough, $\|f(x) - f(x_n)\| < \varepsilon$, but this is the definition of what it means to say that $f(x_n) \to f(x)$.

After the reign of terror, the family returned to Paris and Cauchy studied at the university to be an engineer but became a mathematician although he made fundamental contributions to physics and engineering. Cauchy was one of the most prolific mathematicians who ever lived. He wrote several hundred papers which fill 24 volumes. He also did research on many topics in mechanics and physics including elasticity, optics and astronomy. More than anyone else, Cauchy invented the subject of complex analysis. He is also credited with giving the first rigorous use of continuity in terms of ε, δ arguments in some of his work, although he clung to the notion of infinitesimals. He might have his name associated with more important topics in mathematics and engineering than any other person. He was a devout Catholic, a royalist, adhering to the Bourbons, and a man of integrity and principle, according to his understanding.

He married in 1818 and lived for 12 years with his wife and two daughters in Paris till the revolution of 1830. Cauchy was a "Legitimist" and refused to take the oath of allegiance to the new ruler, Louis Philippe because Louis was not sufficiently Bourbon, and ended up leaving his family and going into exile for 8 years. It wasn't the last time that he refused to take such an oath.

Notwithstanding his great achievements he was not a popular teacher.

[2] Wilhelm Theodor Weierstrass 1815-1897 brought calculus to essentially the state it is in now. When he was a secondary school teacher, he wrote a paper which was so profound that he was granted a doctor's degree. He made fundamental contributions to partial differential equations, complex analysis, calculus of variations, and many other topics. He also discovered some pathological examples such as nowhere differentiable continuous functions. Cauchy and Bolzano were the first to use the ε δ definition presented here but this rigorous definition is associated more with Weierstrass. Cauchy clung to the notion of infinitesimals and Bolzano's work was not readily available. The need for rigor in the subject of calculus was only realized over a long period of time and this definition is part of a trend which went on during the nineteenth century to define exactly what was meant.

$\Leftarrow$ Suppose the sequence condition holds. Why is f continuous at x? If it isn't, then there exists $\varepsilon > 0$ for which there is no suitable definition from the definition of continuity. Hence $1/n$ is not a suitable δ for this ε. It follows that there exists x_n such that $\|x_n - x\| < 1/n$ and yet $\|f(x_n) - f(x)\| \geq \varepsilon$. But then $x_n \to x$ and $f(x_n) \nrightarrow f(x)$ where the symbol $\nrightarrow$ indicates that $f(x_n)$ does not converge to $f(x)$. Hence f must be continuous at x after all. ■

This definition or its equivalent formulation rules out the sorts of graphs drawn above.

Consider the second nonremovable discontinuity. The removable discontinuity case is similar. You could let $x_n \to x_0$ where each $x_n < x_0$ and the limit of $f(x_n)$ will fill in the hole at the bottom of the graph although the actual value of the function at $f(x_0)$ is larger. Thus $f(x_n) \nrightarrow f(x_0)$ so f is not continuous at x_0.

Notice that the concept of continuity as described in the definition is a point property. That is to say it is a property which a function may or may not have at a single point. Here is an example.

Example 6.0.3 *Let*

$$f(x) = \begin{cases} x \text{ if } x \text{ is rational} \\ 0 \text{ if } x \text{ is irrational} \end{cases}.$$

This function is continuous at $x = 0$ and nowhere else.

If $x_n \to 0$, then $|f(x_n)| \leq |x_n|$ and $|x_n| \to 0$ so it follows that $f(x_n) \to 0 \equiv f(0)$ and so the function is continuous at 0. However, if $x \neq 0$ and rational, you could consider a sequence of irrational numbers converging to x, $\{x_n\}$ and $f(x_n) = 0 \to 0 \neq f(x)$. If x is irrational, you could pick a sequence of rational numbers $\{x_n\}$ converging to x and so $f(x_n) = x_n \to x \neq f(x)$. Here is another example.

Example 6.0.4 *Show the function $f(x) = -5x + 10$ is continuous at $x = -3$.*

To do this, note first that $f(-3) = 25$ and it is desired to verify the conditions for continuity. Consider the following. $|-5x + 10 - (25)| = 5|x - (-3)|$.

This allows one to find a suitable δ. If $\varepsilon > 0$ is given, let $0 < \delta \leq \frac{1}{5}\varepsilon$. Then if $0 < |x - (-3)| < \delta$, it follows from this inequality that $|-5x + 10 - (25)| = 5|x - (-3)| < 5\frac{1}{5}\varepsilon = \varepsilon$.

Sometimes the determination of δ in the verification of continuity can be a little more involved. Here is another example.

Example 6.0.5 *Show the function $f(x) = \sqrt{2x + 12}$ is continuous at $x = 5$.*

First note $f(5) = \sqrt{22}$. Now consider: $\left|\sqrt{2x + 12} - \sqrt{22}\right| = \left|\frac{2x+12-22}{\sqrt{2x+12}+\sqrt{22}}\right|$

$$= \frac{2}{\sqrt{2x + 12} + \sqrt{22}}|x - 5| \leq \frac{1}{11}\sqrt{22}|x - 5|$$

whenever $|x - 5| < 1$ because for such x, $\sqrt{2x + 12} > 0$. Now let $\varepsilon > 0$ be given. Choose δ such that $0 < \delta \leq \min\left(1, \frac{\varepsilon\sqrt{22}}{2}\right)$. Then if $|x - 5| < \delta$, all the inequalities above hold and

$$\left|\sqrt{2x + 12} - \sqrt{22}\right| \leq \frac{2}{\sqrt{22}}|x - 5| < \frac{2}{\sqrt{22}}\frac{\varepsilon\sqrt{22}}{2} = \varepsilon.$$

Example 6.0.6 *Show $f(x) = -3x^2 + 7$ is continuous at $x = 7$.*

Suppose $x_n \to x$. Then by the theorem on limits, Theorem 4.4.8, $-3x_n^2 + 7 \to -3x^2 + 7$ and so this function is continuous at x. In particular, it is continuous at 7.

Proposition 6.0.7 *For $x \in \mathbb{F}^p$, and $S \subseteq \mathbb{F}^p, S \neq \emptyset$, let*

$$\inf\{\|x - s\| : s \in S\} \equiv \operatorname{dist}(x, S)$$

Then

$$|\operatorname{dist}(x, S) - \operatorname{dist}(y, S)| \leq \|x - y\| \tag{$*$}$$

so $\operatorname{dist} : \mathbb{F}^p \to \mathbb{R}$ *is continuous.*

Proof: One of $\operatorname{dist}(x, S), \operatorname{dist}(y, S)$ is larger. Say $\operatorname{dist}(y, S) \geq \operatorname{dist}(x, S)$. Then pick $\hat{s} \in S$ such that $\operatorname{dist}(x, S) + \varepsilon > \|x - \hat{s}\|$. Then

$$\begin{aligned} |\operatorname{dist}(x, S) - \operatorname{dist}(y, S)| &= \operatorname{dist}(y, S) - \operatorname{dist}(x, S) \leq \|y - \hat{s}\| - (\|x - \hat{s}\| - \varepsilon) \\ &\leq \|y - x\| + \|x - \hat{s}\| - (\|x - \hat{s}\| - \varepsilon) \leq \|y - x\| + \varepsilon \end{aligned}$$

If $\operatorname{dist}(x, S) > \operatorname{dist}(y, S)$, reverse x, y in the argument. Since ε is arbitrary, this shows $*$. Then letting $\delta = \varepsilon$ in the definition for continuity shows $x \to \operatorname{dist}(x, S)$ is continuous. ■

The following is a useful theorem which makes it easy to recognize many examples of continuous functions.

Theorem 6.0.8 *The following assertions are valid*

1. *The function $af + bg$ is continuous at x when f, g are continuous at $x \in D(f) \cap D(g)$ and $a, b \in \mathbb{F}$.*

2. *If f has values in $\mathbb{F}^q$ and g has values in $\mathbb{F}$ are each continuous at x, then fg is continuous at x. If, in addition to this, $g(x) \neq 0$, then f/g is continuous at x.*

3. *If f is continuous at x, $f(x) \in D(g) \subseteq \mathbb{F}^p$, and g is continuous at $f(x)$,then $g \circ f$ is continuous at x.*

4. *The function $f : \mathbb{F}^p \to \mathbb{R}$, given by $f(x) = \|x\|$ is continuous.*

Proof: All of these follow immediately from the theorem about limits of sequences, Theorem 4.4.8, and the equivalent definition of continuity given above, Theorem 6.0.2. For example, consider the third claim about the composition of continuous functions. Suppose $x_n \to x$. Then by continuity of $f, f(x_n) \to f(x)$ and now, by continuity of $g, g(f(x_n)) \to g(f(x))$. The other claims are similar. ■

6.1 Continuity at Every Point of $D(f)$

Next is a useful property of continuous functions which has to do with preserving order.

Theorem 6.1.1 *Suppose $f : D(f) \to \mathbb{R}$ is continuous at $x \in D(f)$ and suppose $f(x_n) \leq l$ $(\geq l)$ for all n sufficiently large, where $\{x_n\}$ is a sequence of points of $D(f)$ which converges to x. Then $f(x) \leq l\ (\geq l)$.*

Proof: Since $f(x_n) \leq l$ and f is continuous at x, it follows from Theorem 4.4.13 and Theorem 6.0.2, $f(x) = \lim_{n\to\infty} f(x_n) \leq l$. The other case is entirely similar. ■

Now here is the main result about inverse images and continuity at every point.

Theorem 6.1.2 *A function f is continuous at every point of its domain $D(f)$ if and only if either of the two equivalent conditions hold. For*

$$f^{-1}(S) \equiv \{x \in D(f) : f(x) \in S\},$$

1. *$f^{-1}(V) = U \cap D(f)$ for some U open whenever V is open.*

2. *$f^{-1}(H) = C \cap D(f)$ for some C closed whenever H is closed.*

Proof: Continuous $\Rightarrow$ 1. Let $x \in f^{-1}(V)$ so $f(x) \in B(f(x), \varepsilon_x) \subseteq V$ for some $\varepsilon_x > 0$. Then by continuity, there is $\delta_x > 0$ such that $f(B(x,\delta_x) \cap D(f)) \subseteq B(f(x), \varepsilon_x) \subseteq V$ and so, letting $U \equiv \cup_{x\in f^{-1}(V)} B(x,\delta_x)$, it follows that $f^{-1}(V) = U \cap D(f)$.

1. $\Rightarrow$ Continuous. You could pick a particular open set $B(f(x), \varepsilon) \equiv V$ and then $f^{-1}(V) = U \cap D(f)$ for open U and so if $x \in f^{-1}(V)$, then there is δ_x such that $B(x,\delta_x) \subseteq U$ and so if $y \in B(x,\delta_x) \cap D(f)$, then $f(y) \in B(f(x), \varepsilon)$ which is the standard definition of continuity at x.

1. $\Rightarrow$ 2. Let H be closed. Then H^C is open and so $f^{-1}(H^C) = U \cap D(f)$ for U open. But then $f^{-1}(H) = U^C \cap D(f)$ where U^C is closed. $(D(f) = f^{-1}(H) \cup f^{-1}(H^C))$

2. $\Rightarrow$ 1. Let U be open. Then U^C is closed and so $f^{-1}(U^C) = C \cap D(f)$ for a closed set C. But then $C^C \cap D(f) = f^{-1}(U)$ where C^C is open. ■

This suggests the following definition.

Definition 6.1.3 *Let S be a nonempty set. Then one can define relatively open and relatively closed subsets of S as follows. A set $O \subseteq S$ is relatively open if $O = U \cap S$ where U is open. A set $K \subseteq S$ is relatively closed if there is a closed set C such that $K = S \cap C$.*

In words, the above theorem says that a function is continuous at every point of its domain if and only if inverse images of open sets are relatively open if and only if inverse images of closed sets are relatively closed.

6.2 Exercises

1. Let $f(x) = 2x + 7$. Show f is continuous at every point x. **Hint:** You need to let $\varepsilon > 0$ be given. In this case, you should try $\delta \leq \varepsilon/2$. Note that if one δ works in the definition, then so does any smaller δ.

2. Suppose $D(f) = [0,1] \cup \{9\}$ and $f(x) = x$ on $[0,1]$ while $f(9) = 5$. Is f continuous at the point, 9? Use whichever definition of continuity you like.

3. Let $f(x) = x^2 + 1$. Show f is continuous at $x = 3$. **Hint:** Consider the following which comes from algebra. $|f(x) - f(3)| = \left|x^2 + 1 - (9+1)\right| = |x+3||x-3|$. Thus if $|x-3| < 1$, it follows from the triangle inequality, $|x| < 1 + 3 = 4$ and so $|f(x) - f(3)| < 4|x-3|$. Complete the argument by letting $\delta \leq \min(1, \varepsilon/4)$. The symbol, min means to take the minimum of the two numbers in the parenthesis.

4. Let $f(x) = 2x^2 + 1$. Show f is continuous at $x = 1$.

5. Let $f(x) = x^2 + 2x$. Show f is continuous at $x = 2$. Then show it is continuous at every point.

6. Let $f(x) = |2x+3|$. Show f is continuous at every point. **Hint:** Review the two versions of the triangle inequality for absolute values.

7. Let $f(x) = \frac{1}{x^2+1}$. Show f is continuous at every value of x.

8. If $x \in \mathbb{R}$, show there exists a sequence of rational numbers, $\{x_n\}$ such that $x_n \to x$ and a sequence of irrational numbers, $\{x_n'\}$ such that $x_n' \to x$. Now consider the following function.
$$f(x) = \begin{cases} 1 \text{ if } x \text{ is rational} \\ 0 \text{ if } x \text{ is irrational} \end{cases}.$$
Show using the sequential version of continuity in Theorem 6.0.2 that f is discontinuous at every point.

9. If $x \in \mathbb{R}$, show there exists a sequence of rational numbers, $\{x_n\}$ such that $x_n \to x$ and a sequence of irrational numbers, $\{x_n'\}$ such that $x_n' \to x$. Now consider the following function.
$$f(x) = \begin{cases} x \text{ if } x \text{ is rational} \\ 0 \text{ if } x \text{ is irrational} \end{cases}.$$
Show using the sequential version of continuity in Theorem 6.0.2 that f is continuous at 0 and nowhere else.

10. Suppose y is irrational and $y_n \to y$ where y_n is rational. Say $y_n = p_n/q_n$. Show that $\lim_{n\to\infty} q_n = \infty$. Now consider the function
$$f(x) \equiv \begin{cases} 0 \text{ if } x \text{ is irrational} \\ \frac{1}{q} \text{ if } x = \frac{p}{q} \text{ where the fraction is in lowest terms} \end{cases}$$
Show that f is continuous at each irrational number and discontinuous at every nonzero rational number.

11. Suppose f is a function defined on $\mathbb{R}$. Define
$$\omega_\delta f(x) \equiv \sup\{|f(y) - f(z)| : y, z \in B(x, \delta)\}$$
Note that these are decreasing in δ. Let $\omega f(x) \equiv \inf_{\delta>0} \omega_\delta f(x)$. Explain why f is continuous at x if and only if $\omega f(x) = 0$. Next show that
$$\{x : \omega f(x) = 0\} = \cap_{m=1}^\infty \cup_{n=1}^\infty \left\{x : \omega_{(1/n)} f(x) < \frac{1}{m}\right\}$$
Now show that $\cup_{n=1}^\infty \left\{x : \omega_{(1/n)} f(x) < \frac{1}{m}\right\}$ is an open set. Explain why the set of points where f is continuous must always be a G_δ set. Recall that a G_δ set is the countable intersection of open sets.

12. Show that the set of rational numbers is **not** a G_δ set. That is, there is no sequence of open sets whose intersection is the rational numbers. Extend to show that no countable dense set can be G_δ. Using Problem 11, show that there is no function which is continuous at a countable dense set of numbers but discontinuous at every other number.

13. Use the sequential definition of continuity described above to give an easy proof of Theorem 6.0.8.

14. Let $f(x) = \sqrt{x}$ show f is continuous at every value of x in its domain. For now, assume $\sqrt{x}$ exists for all positive x. **Hint:** You might want to make use of the identity, $\sqrt{x} - \sqrt{y} = \frac{x-y}{\sqrt{x}+\sqrt{y}}$ at some point in your argument.

15. Using Theorem 6.0.8, show all polynomials are continuous and that a rational function is continuous at every point of its domain. **Hint:** First show the function given as $f(x) = x$ is continuous and then use the Theorem 6.0.8. What about the case where x can be in $\mathbb{F}$? Does the same conclusion hold?

16. Let $f(x) = \begin{cases} 1 \text{ if } x \in \mathbb{Q} \\ 0 \text{ if } x \notin \mathbb{Q} \end{cases}$ and consider $g(x) = f(x)\left(x - x^3\right)$. Determine where g is continuous and explain your answer.

17. Suppose f is any function whose domain is the integers. Thus $D(f) = \mathbb{Z}$, the set of whole numbers, $\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots$. Then f is continuous. Why? **Hint:** In the definition of continuity, what if you let $\delta = \frac{1}{4}$? Would this δ work for a given $\varepsilon > 0$? This shows that the idea that a continuous function is one for which you can draw the graph without taking the pencil off the paper is a lot of nonsense.

18. Give an example of a function f which is not continuous at some point but $|f|$ is continuous at that point.

19. Find two functions which fail to be continuous but whose product is continuous.

20. Find two functions which fail to be continuous but whose sum is continuous.

21. Find two functions which fail to be continuous but whose quotient is continuous.

22. Suppose f is a function defined on $\mathbb{R}$ and f is continuous at 0. Suppose also that $f(x+y) = f(x) + f(y)$. Show that if this is so, then f must be continuous at every value of $x \in \mathbb{R}$. Next show that for every rational number, r, $f(r) = rf(1)$. Finally explain why $f(r) = rf(1)$ for every r a real number. **Hint:** To do this last part, you need to use the density of the rational numbers and continuity of f.

23. Show that if r is an irrational number and $\frac{p_n}{q_n} \to r$ where p_n, q_n are positive integers, then it must be that $p_n \to \infty$ and $q_n \to \infty$. **Hint:** If not, extract a convergent subsequence for p_n and q_n argue that to which these converge must be integers. Hence r would end up being rational.

6.3 The Extreme Values Theorem

The extreme values theorem says continuous functions achieve their maximum and minimum provided they are defined on a sequentially compact set.

Example 6.3.1 *Let $f(x) = 1/x$ for $x \in (0,1)$.*

Clearly, f is not bounded. Does this violate the conclusion of the above lemma? It does not because the end points of the interval involved are not in the interval. The same function defined on $[.000001, 1)$ would have been bounded although in this case the boundedness of the function would not follow from the above lemma because it fails to include the right endpoint.

The next theorem is known as the max min theorem or extreme value theorem.

Theorem 6.3.2 *Let $K \subseteq \mathbb{F}^p$ be sequentially compact and let $f : K \to \mathbb{R}$ be continuous. Then f achieves its maximum and its minimum on K. This means there exist, $x_1, x_2 \in K$ such that for all $x \in K, f(x_1) \leq f(x) \leq f(x_2)$.*

Proof: Let $\lambda \equiv \sup\{f(x) : x \in K\}$. Then if $l < \lambda$, there exists $x \in K$ such that $f(x) > l$ since otherwise, λ is not as defined since l would be a smaller upper bound. Thus there exists a sequence $\{x_n\} \in K$ such that $\lim_{n\to\infty} f(x_n) = \lambda$. This is called a maximizing sequence. Since K is sequentially compact, there exists a subsequence $\{x_{n_k}\}$ which converges to $x \in K$. Therefore, $\lambda = \lim_{k\to\infty} f(x_{n_k}) = f(x)$ so f achieves its maximum value. A similar argument using a minimizing sequence and $\eta \equiv \inf\{f(x) : x \in K\}$ shows f achieves its minimum value on K. ■

In fact a continuous function takes compact sets to compact sets. This is another of those big theorems which tends to hold whenever it makes sense. Therefore, I will be vague about the domain and range of the function f.

Theorem 6.3.3 *Let $D(f) \supseteq K$ where K is a compact set. Then $f(K)$ is also compact.*

Proof: Suppose $\mathscr{C}$ is an open cover of $f(K)$. Then by Theorem 6.1.2, since f is continuous, it satisfies the inverse image of open sets being open condition. For $U \in \mathscr{C}$,

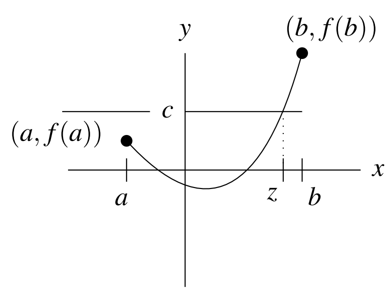
$$f^{-1}(U) = O_U \cap D(f), \text{ where } O_U \text{ is open}$$

Thus $\{O_U : U \in \mathscr{C}\}$ is an open cover of K. Hence there exist $\{O_{U_1}, \cdots, O_{U_n}\}$ each open whose union contains K. It follows that $\{U_1, \cdots, U_n\}$ is an open cover of $f(K)$. ■

You could also do the following: If $\{f(x_n)\}$ is a sequence in $f(K)$, then there is a subsequence $\{x_{n_k}\}$ such that $\lim_{k\to\infty} x_{n_k} = x \in K$ by compactness of K. Hence by continuity, $f(x_{n_k}) \to f(x) \in f(K)$ so $f(K)$ is sequentially compact. By Theorem 4.8.17, $f(K)$ is compact.

6.4 The Intermediate Value Theorem

The next big theorem is called the intermediate value theorem and the following picture illustrates its conclusion. It gives the existence of a certain point. This theorem is due to Bolzano around 1817. He identified completeness of $\mathbb{R}$ as the reason for its validity.



You see in the picture there is a horizontal line, $y = c$ and a continuous function which starts off less than c at the point a and ends up greater than c at point b. The intermediate value theorem says there is some point between a and b shown in the picture as z such that the value of the function at this point equals c. It may seem this is obvious but without completeness the conclusion of the theorem cannot be drawn. Nevertheless, the above picture makes this theorem very easy to believe.

Proposition 6.4.1 *Suppose $f : [a,b] \to \mathbb{R}$ is continuous and suppose*

$$f(a) f(b) \leq 0.$$

Then there exists $x \in [a,b]$ such that $f(x) = 0$.

Proof: When we have an interval $[a_n, b_n]$ in this argument, c_n will be the midpoint $(a_n + b_n)/2$. Let $a_0 = a, b_0 = b$. If $[a_n, b_n]$ has been chosen such that $f(a_n) f(b_n) \leq 0$, consider $[a_n, c_n]$ and $[c_n, b_n]$. Either $f(a_n) f(c_n) \leq 0$ or $f(c_n) f(b_n) \leq 0$ since if both products are positive, then $f(a_n)$ and $f(b_n)$ are either both positive or both negative contradicting $f(a_n) f(b_n) \leq 0$. Pick one of the intervals for which the product is non-positive. Let the left endpoint be a_{n+1} and the right endpoint be b_{n+1} so $f(a_{n+1}) f(b_{n+1}) \leq 0$. Now these nested intervals have exactly one point in their intersection because they have diameters converging to 0. Call it x. Then $(f(x))^2 = \lim_{n\to\infty} f(a_n) f(b_n) \leq 0$. This is by Theorem 6.1.1. Thus $f(x) = 0$. ■

It is easy to generalize this Proposition.

Theorem 6.4.2 *Suppose $f : [a,b] \to \mathbb{R}$ is continuous and suppose either $f(a) < c < f(b)$ or $f(a) > c > f(b)$. Then there exists $x \in (a,b)$ such that $f(x) = 0$.*

Proof: Apply the above proposition to $g(x) \equiv f(x) - c$ obtaining a point $x \in (a,b)$ with $g(x) = f(x) - c = 0$. ■

Lemma 6.4.3 *Let $\phi : [a,b] \to \mathbb{R}$ be a continuous function and suppose ϕ is one to one, written as $1-1$ on (a,b). Then ϕ is either strictly increasing or strictly decreasing on $[a,b]$.*

Proof: First it is shown that ϕ is either strictly increasing or strictly decreasing on (a,b).

If ϕ is not strictly decreasing on (a,b), then there exists $x_1 < y_1$, $x_1, y_1 \in (a,b)$ such that

$$(\phi(y_1) - \phi(x_1))(y_1 - x_1) > 0.$$

If for some other pair of points, $x_2 < y_2$ with $x_2, y_2 \in (a,b)$, the above inequality does not hold, then since ϕ is $1-1$,

$$(\phi(y_2) - \phi(x_2))(y_2 - x_2) < 0.$$

Let $x_t \equiv t x_1 + (1-t) x_2$ and $y_t \equiv t y_1 + (1-t) y_2$. Then $x_t < y_t$ for all $t \in [0,1]$ because

$$t x_1 \leq t y_1 \text{ and } (1-t) x_2 \leq (1-t) y_2$$

with strict inequality holding for at least one of these inequalities since not both t and $(1-t)$ can equal zero. Now define

$$h(t) \equiv (\phi(y_t) - \phi(x_t))(y_t - x_t).$$

Since h is continuous and $h(0) < 0$, while $h(1) > 0$, there exists $t \in (0,1)$ such that $h(t) = 0$. Therefore, both x_t and y_t are points of (a,b) and $\phi(y_t) - \phi(x_t) = 0$ contradicting the assumption that ϕ is one to one. It follows ϕ is either strictly increasing or strictly decreasing on (a,b).

This property of being either strictly increasing or strictly decreasing on (a,b) carries over to $[a,b]$ by the continuity of ϕ. Suppose ϕ is strictly increasing on (a,b). (A similar argument holds for ϕ strictly decreasing on (a,b).) If $x > a$, then let z_n be a decreasing sequence of points of (a,x) converging to a. Then by continuity of ϕ at a,

$$\phi(a) = \lim_{n\to\infty} \phi(z_n) \leq \phi(z_1) < \phi(x).$$

Therefore, $\phi(a) < \phi(x)$ whenever $x \in (a,b)$. Similarly $\phi(b) > \phi(x)$ for all $x \in (a,b)$. ■

Corollary 6.4.4 *Let $f : (a,b) \to \mathbb{R}$ be one to one and continuous. Then $f((a,b))$ is an open interval, (c,d) and $f^{-1} : (c,d) \to (a,b)$ is continuous.*

Proof: Since f is either strictly increasing or strictly decreasing, it maps open intervals to open intervals. Letting I be an open interval, $\left(f^{-1}\right)^{-1}(I) = f(I)$ which is an open interval. Therefore, if V is open, $\left(f^{-1}\right)^{-1}(V) = \left(f^{-1}\right)^{-1}(\cup_{x\in V} I_x) = \cup_{x\in V} \left(f^{-1}\right)^{-1}(I_x)$ which is an open set because it is the union of open sets. Here $x \in I_x \subseteq V$ and the open interval I_x exists because V is open. By Theorem 6.1.2, f^{-1} is continuous. ■

6.5 Connected Sets

Some sets are connected and some are not. The term means roughly that the set is in one "one piece". The concept is a little tricky because it is defined in terms of not being something else. In some of the theorems below, I will be vague about where the sets involved in the discussion are because it is often the case that it doesn't matter. However, you can think of the sets as being in $\mathbb{F}^p$ where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. First recall the following definition.

Definition 6.5.1 *Let S be a set. Then $\bar{S}$, called the **closure of** S consists of $S \cup S'$ where S' denotes the set of limit points of S.*

Recall Corollary 4.8.12 which says that $S \cup S'$, denoted as $\overline{S}$ is the intersection of all closed sets which contain S and is a closed set.

Note that it is obvious from the above definition that if $S \subseteq T$, then $\bar{S} \subseteq \bar{T}$.

Definition 6.5.2 *A set S is said to be **separated** if it is of the form*

$$S = A \cup B, \text{ where } \bar{A} \cap B = \bar{B} \cap A = \emptyset$$

*A set S is **connected** if it is not separated.*

Example 6.5.3 *Consider $S = [0,1) \cup (1,2]$. This is separated. Therefore, it is not connected.*

To see this, note that $\overline{[0,1)} = [0,1]$ which has empty intersection with $(1,2]$. Similarly $\overline{(1,2]} = [1,2]$ and has empty intersection with $[0,1)$.

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

Theorem 6.5.4 *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 $B \cap U$. 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. ■

How do connected sets relate to continuous real valued functions?

Theorem 6.5.5 *Let $f : X \to \mathbb{R}$ be continuous where X is connected. Then $f(X)$ is also connected.*

Proof: To do this you 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 ball U of radius ε for some $\varepsilon > 0$ containing $f(z)$ such that $U \cap A = \emptyset$. But then, the continuity of f and the definition of continuity imply that there exists $\delta > 0$ such that $f(B(z,\delta)) \subseteq U$. Therefore z is not a limit point of $f^{-1}(A)$. Since z was arbitrary, it follows that $f^{-1}(B)$ 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. ■

On $\mathbb{R}$ the connected sets are pretty easy to describe. 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 6.5.6 *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 $p, q \in C$. You 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. ■

Another useful idea is that of connected components. 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 6.5.7 *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 6.5.8 *Let C_p be a connected component of a set S. Then C_p is a connected set and if $C_p \cap C_q \neq \emptyset$, then $C_p = C_q$.*

Proof: Let $\mathscr{C}$ 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 $\mathscr{C}$ must also be contained in A since otherwise, as in Theorem 6.5.4, the set would be separated. But this implies B is empty. Therefore, C_p is connected. From this, and Theorem 6.5.4, the second assertion of the theorem is proved. ■

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

Probably the most useful application of this is to the case where you have an open set and consider its connected components.

Theorem 6.5.9 *Let U be an open set on $\mathbb{R}$. Then each connected component is open. Thus U is an at most countable union of disjoint open intervals.*

Proof: Let C be a connected component of U. Let $x \in C$. Since U is open, there exists $\delta > 0$ such that $(x-\delta, x+\delta) \subseteq U$. Hence this open interval is also contained in C because it is connected and shares a point with C which equals the union of all connected sets containing x. Thus each component is both open and connected and is therefore, an open interval. Each of these disjoint open intervals contains a rational number. Therefore, there are countably many of them because there are countably many rational numbers. ■

That the rational numbers are at most countable is easy to see. You know the integers are countable because they are the union of two countable sets. Thus $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ is countable because of Theorem 3.2.7. Now let $\theta : \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \to \mathbb{Q}$ be defined as $\theta(m,n) \equiv \frac{m}{n}$. This is onto. Hence $\mathbb{Q}$ is at most countable. This is sufficient to conclude there are at most countably many of these open intervals.

To emphasize what the above theorem shows, it states that every open set in $\mathbb{R}$ is the countable union of open intervals. It is really nice to be able to say this.

6.6 Exercises

1. Give an example of a continuous function defined on $(0,1)$ which does not achieve its maximum on $(0,1)$.

2. Give an example of a continuous function defined on $(0,1)$ which is bounded but which does not achieve either its maximum or its minimum.

3. Give an example of a discontinuous function defined on $[0,1]$ which is bounded but does not achieve either its maximum or its minimum.

4. Give an example of a continuous function defined on $[0,1) \cup (1,2]$ which is positive at 2, negative at 0 but is not equal to zero for any value of x.

5. Let $f(x) = x^5 + ax^4 + bx^3 + cx^2 + dx + e$ where a, b, c, d, and e are numbers. Show there exists real x such that $f(x) = 0$.

6. Give an example of a function which is one to one but neither strictly increasing nor strictly decreasing.

7. Show that the function $f(x) = x^n - a$, where n is a positive integer and a is a number, is continuous.

8. Use the intermediate value theorem on the function $f(x) = x^7 - 8$ to show $\sqrt[7]{8}$ must exist. State and prove a general theorem about n^{th} roots of positive numbers.

9. Prove $\sqrt{2}$ is irrational. **Hint:** Suppose $\sqrt{2} = p/q$ where p, q are positive integers and the fraction is in lowest terms. Then $2q^2 = p^2$ and so p^2 is even. Explain why $p = 2r$ so p must be even. Next argue q must be even.

10. Let $f(x) = x - \sqrt{2}$ for $x \in \mathbb{Q}$, the rational numbers. Show that even though $f(0) < 0$ and $f(2) > 0$, there is no point in $\mathbb{Q}$ where $f(x) = 0$. Does this contradict the intermediate value theorem? Explain.

11. It has been known since the time of Pythagoras that $\sqrt{2}$ is irrational. If you throw out all the irrational numbers, show that the conclusion of the intermediate value theorem could no longer be obtained. That is, show there exists a function which starts off less than zero and ends up larger than zero and yet there is no number where the function equals zero. **Hint:** Try $f(x) = x^2 - 2$. You supply the details.

12. A circular hula hoop lies partly in the shade and partly in the hot sun. Show there exist two points on the hula hoop which are at opposite sides of the hoop which have the same temperature. **Hint:** Imagine this is a circle and points are located by specifying their angle, θ from a fixed diameter. Then letting $T(\theta)$ be the temperature in the hoop, $T(\theta + 2\pi) = T(\theta)$. You need to have $T(\theta) = T(\theta + \pi)$ for some θ. Assume T is a continuous function of θ.

13. A car starts off on a long trip with a full tank of gas. The driver intends to drive the car till it runs out of gas. Show that at some time the number of miles the car has gone exactly equals the number of gallons of gas in the tank.

14. Suppose f is a continuous function defined on $[0,1]$ which maps $[0,1]$ into $[0,1]$. Show there exists $x \in [0,1]$ such that $x = f(x)$. **Hint:** Consider $h(x) \equiv x - f(x)$ and the intermediate value theorem. This is a one dimensional version of the Brouwer fixed point theorem.

15. Let f be a continuous function on $[0,1]$ such that $f(0) = f(1)$. Let n be a positive integer larger than 2. Show there must exist $c \in \left[0, 1 - \frac{1}{n}\right]$ such that $f\left(c + \frac{1}{n}\right) = f(c)$. **Hint:** Consider $h(x) \equiv f\left(x + \frac{1}{n}\right) - f(x)$. Consider the subintervals $\left[\frac{k-1}{n}, \frac{k}{n}\right]$ for $k = 1, \cdots, n-1$. You want to show that h equals zero on one of these intervals. If h changes sign between two successive intervals, then you are done. Assume then, that this does not happen. Say h remains positive. Argue that $f(0) < f\left(\frac{n-1}{n}\right)$. Thus $f\left(\frac{n-1}{n}\right) > f(1) = f\left(\frac{n-1}{n} + \frac{1}{n}\right)$. It follows that $h\left(1 - \frac{1}{n}\right) < 0$ but $h\left(1 - \frac{2}{n}\right) > 0$.

16. Use Theorem 6.5.5 and the characterization of connected sets in $\mathbb{R}$ to give a quick proof of the intermediate value theorem.

17. A set is said to be totally disconnected if each component consists of a single point. Show that the Cantor set is totally disconnected but that every point is a limit point of the set. **Hint:** Show it contains no intervals other than single points.

18. A perfect set is a non empty closed set such that every point is a limit point. Show that no perfect set in $\mathbb{R}$ can be countable. **Hint:** You might want to use the fact that the set of infinite sequences of 0 and 1 is uncountable. Show that there is a one to one mapping from this set of sequences onto a subset of the perfect set.

19. Suppose $f : K \to \mathbb{R}$ where K is a compact set and f is continuous. Show that f achieves its maximum and minimum by using Theorem 6.3.3 and the characterization of compact sets in $\mathbb{R}$ given earlier which said that such a set is closed and bounded. **Hint:** You need to show that a closed and bounded set in $\mathbb{R}$ has a largest value and a smallest value.

6.7 Uniform Continuity

There is a theorem about the integral of a continuous function which requires the notion of uniform continuity. Uniform continuity is discussed in this section. Consider the function $f(x) = \frac{1}{x}$ for $x \in (0,1)$. This is a continuous function because, by Theorem 6.0.8, it is continuous at every point of $(0,1)$. However, for a given $\varepsilon > 0$, the δ needed in the ε, δ definition of continuity becomes very small as x gets close to 0. The notion of uniform continuity involves being able to choose a single δ which works on the whole domain of f. Here is the definition.

Definition 6.7.1 *Let f be a function. Then f is uniformly continuous if for every $\varepsilon > 0$, there exists a δ* ***depending only on*** *ε such that if $|x-y| < \delta$ then $|f(x) - f(y)| < \varepsilon$.*

It is an amazing fact that under certain conditions continuity implies uniform continuity.

Theorem 6.7.2 *Let $f : K \to \mathbb{F}^q$ be continuous where K is a sequentially compact set in $\mathbb{F}^p$. Then f is uniformly continuous on K.*

Proof: If this is not true, there exists $\varepsilon > 0$ such that for every $\delta > 0$ there exists a pair of points, x_δ and y_δ such that even though $\|x_\delta - y_\delta\| < \delta$, $\|f(x_\delta) - f(y_\delta)\| \geq \varepsilon$. Taking a succession of values for δ equal to $1, 1/2, 1/3, \cdots$, and letting the exceptional pair of points for $\delta = 1/n$ be denoted by x_n and y_n,

$$\|x_n - y_n\| < \frac{1}{n}, \|f(x_n) - f(y_n)\| \geq \varepsilon.$$

Now since K is sequentially compact, there exists a subsequence, $\{x_{n_k}\}$ such that $x_{n_k} \to z \in K$. Now $n_k \geq k$ and so $\|x_{n_k} - y_{n_k}\| < \frac{1}{k}$. Consequently, $y_{n_k} \to z$ also. x_{n_k} is like a person walking toward a certain point and y_{n_k} is like a dog on a leash which is constantly getting shorter. Obviously y_{n_k} must also move toward the point also. Indeed,

$$\|y_{n_k} - z\| \leq \|y_{n_k} - x_{n_k}\| + \|x_{n_k} - z\| \leq \frac{1}{k} + \|x_{n_k} - z\|$$

and the right side converges to 0 as $k \to \infty$.

By continuity of f and Theorem 6.1.1,

$$0 = \|f(z) - f(z)\| = \lim_{k\to\infty} \|f(x_{n_k}) - f(y_{n_k})\| \geq \varepsilon,$$

an obvious contradiction. Therefore, the theorem must be true. ■

The following corollary follows from this theorem and Theorems 4.8.14, 4.8.13 which give closed and bounded sets are sequentially compact.

Corollary 6.7.3 *Suppose K is any closed and bounded set in $\mathbb{F}^p$. Then if f is continuous on K, it follows that f is uniformly continuous on K.*

6.8 Exercises

1. A function $f : \mathbb{F}^p \to \mathbb{F}^q$ is Holder continuous if there exists a constant, K such that

$$\|f(x) - f(y)\| \leq K \|x - y\|^{\alpha}$$

for all $x, y \in D$. Show every Holder continuous function is uniformly continuous. When $\alpha = 1$, this is called a Lipschitz function or Lipschitz continuous function.

2. Let $x \to \text{dist}(x, S)$ be defined in Proposition 6.0.7. Show it is uniformly continuous on $\mathbb{F}^p$.

3. If $\|x_n - y_n\| \to 0$ and $x_n \to z$, show that $y_n \to z$ also. This was used in the proof of Theorem 6.7.2.

4. Consider $f : (1, \infty) \to \mathbb{R}$ given by $f(x) = \frac{1}{x}$. Show f is uniformly continuous even though the set on which f is defined is not sequentially compact.

5. If f is uniformly continuous, does it follow that $|f|$ is also uniformly continuous? If $|f|$ is uniformly continuous does it follow that f is uniformly continuous? Answer the same questions with "uniformly continuous" replaced with "continuous". Explain why.

6. Let $f : D \to \mathbb{R}$ be a function. This function is said to be lower semicontinuous[3] at $x \in D$ if for any sequence $\{x_n\} \subseteq D$ which converges to x it follows

$$f(x) \leq \liminf_{n \to \infty} f(x_n).$$

Suppose D is sequentially compact and f is lower semicontinuous at every point of D. Show that then f achieves its minimum on D.

7. Let $f : D \to \mathbb{R}$ be a function. This function is said to be upper semicontinuous at $x \in D$ if for any sequence $\{x_n\} \subseteq D$ which converges to x it follows

$$f(x) \geq \limsup_{n \to \infty} f(x_n).$$

Suppose D is sequentially compact and f is upper semicontinuous at every point of D. Show that then f achieves its maximum on D.

8. Show that a real valued function is continuous if and only if it is both upper and lower semicontinuous.

9. Give an example of a lower semicontinuous function which is not continuous and an example of an upper semicontinuous function which is not continuous.

[3] The notion of lower semicontinuity is very important for functions which are defined on infinite dimensional sets. In more general settings, one formulates the concept differently.

10. Suppose $\{f_\alpha : \alpha \in \Lambda\}$ is a collection of continuous functions. Let
$$F(x) \equiv \inf\{f_\alpha(x) : \alpha \in \Lambda\}$$
Show F is an upper semicontinuous function. Next let
$$G(x) \equiv \sup\{f_\alpha(x) : \alpha \in \Lambda\}$$
Show G is a lower semicontinuous function.

11. Let f be a function. $\text{epi}(f)$ is defined as
$$\{(x,y) : y \geq f(x)\}.$$
It is called the epigraph of f. We say $\text{epi}(f)$ is closed if whenever $(x_n, y_n) \in \text{epi}(f)$ and $x_n \to x$ and $y_n \to y$, it follows $(x,y) \in \text{epi}(f)$. Show f is lower semicontinuous if and only if $\text{epi}(f)$ is closed. What would be the corresponding result equivalent to upper semicontinuous?

12. Suppose $K \subseteq \mathbb{F}^p$ is a compact set and $f : K \to \mathbb{F}^q$ is continuous and one to one. Show that $f^{-1} : f(K) \to K$ is continuous.

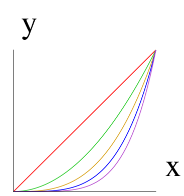
6.9 Sequences and Series of Functions

When you understand sequences and series of numbers it is easy to consider sequences and series of functions.

Definition 6.9.1 *A sequence of functions is a map defined on $\mathbb{N}$ or some set of integers larger than or equal to a given integer, m which has values which are functions. It is written in the form $\{f_n\}_{n=m}^{\infty}$ where f_n is a function. It is assumed also that the domain of all these functions is the same.*

In the above, where do the functions have values? Are they real valued functions? Are they complex valued functions? Are they functions which have values in $\mathbb{R}^n$? It turns out it does not matter very much and the same definition holds. However, if you like, you can think of them as having values in $\mathbb{F}$. This is the main case of interest here.

Example 6.9.2 *Suppose $f_n(x) = x^n$ for $x \in [0,1]$. Here is a graph of the functions $f(x) = x, x^2, x^3, x^4, x^5$.*



Definition 6.9.3 *Let $\{f_n\}$ be a sequence of functions. Then the sequence converges pointwise to a function f if for all $x \in D$, the domain of the functions in the sequence,*
$$f(x) = \lim_{n\to\infty} f_n(x)$$
This is always the definition regardless of where the f_n have their values.

Thus you consider for each $x \in D$ the sequence $\{f_n(x)\}$ and if this sequence converges for each $x \in D$, the thing it converges to is called $f(x)$.

Example 6.9.4 *In Example 6.9.2 find* $\lim_{n\to\infty} f_n$.

For $x \in [0,1), \lim_{n\to\infty} x^n = f_n(x) = 0$. At $x = 1, f_n(1) = 1$ for all n so $\lim_{n\to\infty} f_n(1) = 1$. Therefore, this sequence of functions converges pointwise to the function $f(x)$ given by $f(x) = 0$ if $0 \leq x < 1$ and $f(1) = 1$. However, given small $\varepsilon > 0$, and n, there is always some x such that $|f(x) - f_n(x)| > \varepsilon$. Just pick x less than 1 but close to 1. Then $f(x) = 0$ but $f_n(x)$ will be close to 1.

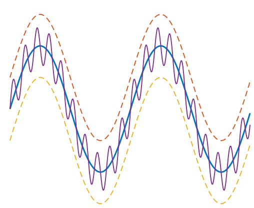
Pointwise convergence is a very inferior thing but sometimes it is all you can get. It's undesirability is illustrated by Example 6.9.4. The limit function is not continuous although each f_n is continuous. Now here is another example of a sequence of functions.

Example 6.9.5 *Let* $f_n(x) = \frac{1}{n}\sin\left(n^2 x\right)$.

In this example, $|f_n(x)| \leq \frac{1}{n}$ so this function is close to 0 for all x at once provided n is large enough. There is a difference between the two examples just given. They both involve pointwise convergence, but in the second example, the pointwise convergence happens for all x at once. In this example, you have **uniform convergence**.

Definition 6.9.6 *Let $\{f_n\}$ be a sequence of functions defined on D. Then $\{f_n\}$ is said to* ***converge uniformly*** *to f if it converges pointwise to f and for every $\varepsilon > 0$ there exists N such that for all $n \geq N$,* $\sup_{x\in D} |f(x) - f_n(x)| < \varepsilon$

The following picture illustrates the above definition.



The dashed lines define a small tube centered about the graph of f and the graph of the function f_n fits in this tube for all n sufficiently large. In the picture, the function f is being approximated by f_n which is very wriggly.

The reason uniform convergence is desirable is that it drags continuity along with it and imparts this property to the limit function.

Theorem 6.9.7 *Let $\{f_n\}$ be a sequence of functions defined on D which are continuous at z and suppose this sequence converges uniformly to f. Then f is also continuous at z. If each f_n is uniformly continuous on D, then f is also uniformly continuous on D.*

Proof: Let $\varepsilon > 0$ be given and pick $z \in D$. By uniform convergence, there exists N such that if $n > N$, then for all $x \in D$,

$$|f(x) - f_n(x)| < \varepsilon/3. \tag{6.1}$$

Pick such an n. By assumption, f_n is continuous at z. Therefore, there exists $\delta > 0$ such that if $|z - x| < \delta$ then $|f_n(x) - f_n(z)| < \varepsilon/3$. It follows that for $|x - z| < \delta$,

$$\begin{aligned} |f(x) - f(z)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(z)| + |f_n(z) - f(z)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \end{aligned}$$

which shows that since ε was arbitrary, f is continuous at z.

In the case where each f_n is uniformly continuous, and using the same f_n for which 6.1 holds, there exists a $\delta > 0$ such that if $|y-z| < \delta$, then $|f_n(z) - f_n(y)| < \varepsilon/3$. Then for $|y-z| < \delta$,

$$\begin{aligned} |f(y) - f(z)| &\leq |f(y) - f_n(y)| + |f_n(y) - f_n(z)| + |f_n(z) - f(z)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \end{aligned}$$

This shows uniform continuity of f. ■

Definition 6.9.8 *Let $\{f_n\}$ be a sequence of functions defined on D. Then the sequence is said to be uniformly Cauchy if for every $\varepsilon > 0$ there exists N such that whenever $m, n \geq N$, $\sup_{x \in D} |f_m(x) - f_n(x)| < \varepsilon$.*

Then the following theorem follows easily.

Theorem 6.9.9 *Let $\{f_n\}$ be a uniformly Cauchy sequence of $\mathbb{F}$ valued functions defined on D. Then there exists f defined on D such that $\{f_n\}$ converges uniformly to f.*

Proof: For each $x \in D$, $\{f_n(x)\}$ is a Cauchy sequence. Therefore, it converges to some number because of completeness of $\mathbb{F}$. (Recall that completeness is the same as saying every Cauchy sequence converges.) Denote by $f(x)$ this number. Let $\varepsilon > 0$ be given and let N be such that if $n, m \geq N$, $|f_m(x) - f_n(x)| < \varepsilon/2$ for all $x \in D$. Then for any $x \in D$, pick $n \geq N$ and it follows from Theorem 4.4.13

$$|f(x) - f_n(x)| = \lim_{m \to \infty} |f_m(x) - f_n(x)| \leq \varepsilon/2 < \varepsilon. \blacksquare$$

Corollary 6.9.10 *Let $\{f_n\}$ be a uniformly Cauchy sequence of functions continuous on D. Then there exists f defined on D such that $\{f_n\}$ converges uniformly to f and f is continuous. Also, if each f_n is uniformly continuous, then so is f.*

Proof: This follows from Theorem 6.9.9 and Theorem 6.9.7. ■

Here is one more fairly obvious theorem.

Theorem 6.9.11 *Let $\{f_n\}$ be a sequence of functions defined on D. Then it converges pointwise if and only if the sequence $\{f_n(x)\}$ is a Cauchy sequence for every $x \in D$. It converges uniformly if and only if $\{f_n\}$ is a uniformly Cauchy sequence.*

Proof: If the sequence converges pointwise, then by Theorem 4.5.3 the sequence $\{f_n(x)\}$ is a Cauchy sequence for each $x \in D$. Conversely, if $\{f_n(x)\}$ is a Cauchy sequence for each $x \in D$, then since f_n has values in $\mathbb{F}$, and $\mathbb{F}$ is complete, it follows the sequence $\{f_n(x)\}$ converges for each $x \in D$. (Recall that completeness is the same as saying every Cauchy sequence converges.)

Now suppose $\{f_n\}$ is uniformly Cauchy. Then from Theorem 6.9.9 there exists f such that $\{f_n\}$ converges uniformly on D to f. Conversely, if $\{f_n\}$ converges uniformly to f on D, then if $\varepsilon > 0$ is given, there exists N such that if $n \geq N$,

$$|f(x) - f_n(x)| < \varepsilon/2$$

for every $x \in D$. Then if $m, n \geq N$ and $x \in D$,

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Thus $\{f_n\}$ is uniformly Cauchy. ■

Note that the above theorem would hold just as well if the functions had values in any complete space meaning that Cauchy sequences converge. As before, once you understand sequences, it is no problem to consider series.

Definition 6.9.12 *Let $\{f_n\}$ be a sequence of functions defined on D. Then*

$$\left(\sum_{k=1}^{\infty} f_k\right)(x) \equiv \lim_{n\to\infty} \sum_{k=1}^{n} f_k(x) \tag{6.2}$$

whenever the limit exists. Thus there is a new function denoted by

$$f(x) \equiv \sum_{k=1}^{\infty} f_k(x) \tag{6.3}$$

If for all $x \in D$, the limit in 6.2 exists, then 6.3 is said to converge pointwise. $\sum_{k=1}^{\infty} f_k$ is said to converge uniformly on D if the sequence of partial sums, $\{\sum_{k=1}^{n} f_k\}$ converges uniformly. If the indices for the functions start at some other value than 1, you make the obvious modification to the above definition as was done earlier with series of numbers.

Theorem 6.9.13 *Let $\{f_n\}$ be a sequence of functions defined on D. The series $\sum_{k=1}^{\infty} f_k$ converges pointwise if and only if for each $\varepsilon > 0$ and $x \in D$, there exists $N_{\varepsilon,x}$ which may depend on x as well as ε such that when $q > p \geq N_{\varepsilon,x}$,*

$$\left|\sum_{k=p}^{q} f_k(x)\right| < \varepsilon$$

The series $\sum_{k=1}^{\infty} f_k$ converges uniformly on D if for every $\varepsilon > 0$ there exists N_ε such that if $q > p \geq N_\varepsilon$ then

$$\sup_{x\in D}\left|\sum_{k=p}^{q} f_k(x)\right| < \varepsilon \tag{6.4}$$

Proof: The first part follows from Theorem 5.1.7. The second part follows from observing the condition is equivalent to the sequence of partial sums forming a uniformly Cauchy sequence and then by Theorem 6.9.11, these partial sums converge uniformly to a function which is the definition of $\sum_{k=1}^{\infty} f_k$. ■

Is there an easy way to recognize when 6.4 happens? Yes, there is. It is called the Weierstrass M test.

Theorem 6.9.14 *Let $\{f_n\}$ be a sequence of functions defined on D. Suppose there exists M_n such that $\sup\{|f_n(x)| : x \in D\} < M_n$ and $\sum_{n=1}^{\infty} M_n$ converges. Then $\sum_{n=1}^{\infty} f_n$ converges uniformly on D.*

Proof: Let $z \in D$. Then letting $m < n$,

$$\left|\sum_{k=1}^{n} f_k(z) - \sum_{k=1}^{m} f_k(z)\right| \leq \sum_{k=m+1}^{n} |f_k(z)| \leq \sum_{k=m+1}^{\infty} M_k < \varepsilon$$

whenever m is large enough because of the assumption that $\sum_{n=1}^{\infty} M_n$ converges. Therefore, the sequence of partial sums is uniformly Cauchy on D and therefore, converges uniformly to $\sum_{k=1}^{\infty} f_k$ on D. ■

Theorem 6.9.15 *If $\{f_n\}$ is a sequence of functions defined on D which are continuous at z and $\sum_{k=1}^{\infty} f_k$ converges uniformly, then the function $\sum_{k=1}^{\infty} f_k$ must also be continuous at z.*

Proof: This follows from Theorem 6.9.7 applied to the sequence of partial sums of the above series which is assumed to converge uniformly to the function $\sum_{k=1}^{\infty} f_k$. ■

6.10 Weierstrass Approximation

It turns out that if f is a continuous real valued function defined on an interval, $[a,b]$ then there exists a sequence of polynomials, $\{p_n\}$ such that the sequence converges uniformly to f on $[a,b]$. I will first show this is true for the interval $[0,1]$ and then verify it is true on any closed and bounded interval. First here is a little lemma which is interesting for its own sake in probability. It is actually an estimate for the variance of a binomial distribution.

Lemma 6.10.1 *The following estimate holds for $x \in [0,1]$ and $m \geq 2$.*

$$\sum_{k=0}^{m} \binom{m}{k} (k - mx)^2 x^k (1-x)^{m-k} \leq \frac{1}{4} m$$

Proof: Here are some observations. $\sum_{k=0}^{m} \binom{m}{k} k x^k (1-x)^{m-k} =$

$$\begin{aligned} & mx \sum_{k=1}^{m} \frac{(m-1)!}{(k-1)!\,((m-1)-(k-1))!} x^{k-1} (1-x)^{(m-1)-(k-1)} \\ = & \ mx \sum_{k=0}^{m-1} \binom{m-1}{k} x^k (1-x)^{m-1-k} = mx \end{aligned}$$

$$\begin{aligned} & \sum_{k=0}^{m} \binom{m}{k} k(k-1) x^k (1-x)^{m-k} \\ = & \ m(m-1) x^2 \sum_{k=2}^{m} \frac{(m-2)!}{(k-2)!\,(m-2-(k-2))!} x^{k-2} (1-x)^{(m-2)-(k-2)} \\ = & \ m(m-1) x^2 \sum_{k=0}^{m-2} \binom{m-2}{k} x^k (1-x)^{(m-2)-k} = m(m-1) x^2 \end{aligned}$$

Now $(k-mx)^2 = k^2 - 2kmx + m^2x^2 = k(k-1) + k(1-2mx) + m^2x^2$. From the above and the binomial theorem, $\sum_{k=0}^{m} \binom{m}{k} (k-mx)^2 x^k (1-x)^{m-k} =$

$$\sum_{k=0}^{m} \binom{m}{k} k(k-1) x^k (1-x)^{m-k} + (1-2mx) \sum_{k=0}^{m} \binom{m}{k} k x^k (1-x)^{m-k}$$

$$+ m^2x^2 \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} = m(m-1)x^2 + (1-2mx)\,mx + m^2x^2$$

$$= mx(1-x) \leq m\frac{1}{4} \ ■$$

Now let f be a continuous function defined on $[0,1]$. Let p_n be the polynomial defined by

$$p_n(x) \equiv \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}. \tag{6.5}$$

Theorem 6.10.2 *The sequence of polynomials in 6.5 converges uniformly to f on $[0,1]$. These polynomials are called the Bernstein polynomials.*

Proof: By the binomial theorem,

$$f(x) = f(x)\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}f(x)x^k(1-x)^{n-k}$$

and so by the triangle inequality

$$|f(x) - p_n(x)| \leq \sum_{k=0}^{n}\binom{n}{k}\left|f\left(\frac{k}{n}\right) - f(x)\right|x^k(1-x)^{n-k}$$

At this point you break the sum into two pieces, those values of k such that k/n is close to x and those values for k such that k/n is not so close to x. Thus

$$\begin{aligned} |f(x) - p_n(x)| \leq & \sum_{|x-(k/n)|<\delta}\binom{n}{k}\left|f\left(\frac{k}{n}\right) - f(x)\right|x^k(1-x)^{n-k} \\ & + \sum_{|x-(k/n)|\geq\delta}\binom{n}{k}\left|f\left(\frac{k}{n}\right) - f(x)\right|x^k(1-x)^{n-k} \end{aligned} \tag{6.6}$$

where δ is a positive number chosen in an auspicious manner about to be described. Since f is continuous on $[0,1]$, it follows from Theorems 4.8.2 and 6.7.2 that f is uniformly continuous. Therefore, letting $\varepsilon > 0$, there exists $\delta > 0$ such that if $|x-y| < \delta$, then $|f(x) - f(y)| < \varepsilon/2$. This is the auspicious choice for δ. Also, by Lemma 6.3.2 $|f(x)|$ for $x \in [0,1]$ is bounded by some number M. Thus 6.6 implies that for $x \in [0,1]$,

$$\begin{aligned} |f(x) - p_n(x)| & \leq \sum_{|x-(k/n)|<\delta}\binom{n}{k}\frac{\varepsilon}{2}x^k(1-x)^{n-k} \\ & \quad +2M\sum_{|nx-k|\geq n\delta}\binom{n}{k}x^k(1-x)^{n-k} \\ & \leq \frac{\varepsilon}{2} + 2M\sum_{|nx-k|\geq n\delta}\binom{n}{k}\frac{(k-nx)^2}{n^2\delta^2}x^k(1-x)^{n-k} \\ & \leq \frac{\varepsilon}{2} + \frac{2M}{n^2\delta^2}\sum_{k=0}^{n}\binom{n}{k}(k-nx)^2x^k(1-x)^{n-k} \end{aligned}$$

Now by Lemma 6.10.1 there is an estimate for the last sum. Using this estimate yields for all $x \in [0,1]$,

$$|f(x) - p_n(x)| \leq \frac{\varepsilon}{2} + \frac{2M}{n^2\delta^2}\frac{n}{4} = \frac{\varepsilon}{2} + \frac{M}{2n\delta^2}.$$

Therefore, whenever n is sufficiently large that $\frac{M}{2n\delta^2} < \frac{\varepsilon}{2}$,it follows that for all n this large and $x \in [0,1]$,

$$|f(x) - p_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \blacksquare$$

Now this theorem has been done, it is easy to extend to continuous functions defined on $[a,b]$. This yields the celebrated Weierstrass approximation theorem. Also note that this would hold just as well if the functions had values in $\mathbb{C}$ or even $\mathbb{C}^n$ provided you had a norm defined on $\mathbb{C}^n$. In fact, it would hold if the functions have values in any normed space, a vector space which has a norm. These Bernstein polynomials are very remarkable.

Theorem 6.10.3 *Suppose f is a continuous function defined on $[a,b]$. Then there exists a sequence of polynomials, $\{p_n\}$ which converges uniformly to f on $[a,b]$.*

Proof: For $t \in [0,1]$, let $h(t) = a + (b-a)t$. Thus h maps $[0,1]$ one to one and onto $[a,b]$. Thus $f \circ h$ is a continuous function defined on $[0,1]$. It follows there exists a sequence of polynomials $\{p_n\}$ defined on $[0,1]$ which converges uniformly to $f \circ h$ on $[0,1]$. Thus for every $\varepsilon > 0$ there exists N_ε such that if $n \geq N_\varepsilon$, then for all $t \in [0,1]$, $|f \circ h(t) - p_n(t)| < \varepsilon$. However, h is onto and one to one and so for all $x \in [a,b]$, $\left|f(x) - p_n\left(h^{-1}(x)\right)\right| < \varepsilon$. Now note that the function $x \to p_n\left(h^{-1}(x)\right)$ is a polynomial because $h^{-1}(x) = \frac{x-a}{b-a}$. More specifically, if $p_n(t) = \sum_{k=0}^m a_k t^k$ it follows

$$p_n\left(h^{-1}(x)\right) = \sum_{k=0}^m a_k \left(\frac{x-a}{b-a}\right)^k$$

which is clearly another polynomial. ■

6.11 Ascoli Arzela Theorem

This is a major result which plays the role of the Heine Borel theorem for the set of continuous functions. I will give the version which holds on an interval, although this theorem holds in much more general settings. First is a definition of what it means for a collection of functions to be equicontinuous. In words, this happens when they are all uniformly continuous simultaneously.

Definition 6.11.1 *Let $S \subseteq C([0,T])$ where $C([0,T])$ denotes the set of functions which are continuous on the interval $[0,T]$. Thus S is a set of functions. Then S is said to be equicontinuous if whenever $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $f \in S$ and $|x-y| < \delta$, it follows*

$$|f(x) - f(y)| < \varepsilon$$

The set of functions is said to be uniformly bounded if there is a positive number M such that for all $f \in S$,

$$\sup\{|f(x)| : x \in [0,T]\} \leq M$$

Then the Ascoli Arzela theorem says the following in which it is assumed the functions have values in $\mathbb{F}$ although this could be generalized.

Theorem 6.11.2 *Let $\{f_n\}_{n=1}^\infty \subseteq C([0,T])$ be uniformly bounded and equicontinuous. Then there exists a uniformly Cauchy subsequence.*

Proof: Let $\varepsilon > 0$ be given and let δ correspond to $\varepsilon/4$ in the definition of equicontinuity. Let $0 = x_0 < x_1 < \cdots < x_n = T$ where these points are uniformly spaced and the distance between successive points is $T/n < \delta$. Then the points $\{f_n(x_0)\}_{n=1}^\infty$ is a bounded set in $\mathbb{F}$. By the Heine Borel theorem, there is a convergent subsequence $\left\{f_{k(0)}(x_0)\right\}_{k(0)=1}^\infty$. Thus $\{k(0)\}$ denotes a strictly increasing sequence of integers. Then the same theorem implies there is a convergent subsequence of this one, denoted as $k(1)$ such that $\lim_{k(1)\to\infty} f_{k(1)}(x_0)$ and $\lim_{k(1)\to\infty} f_{k(1)}(x_1)$ both exist. Then take a subsequence of $\left\{f_{k(1)}\right\}$ called $k(2)$ such that for $x_i = x_0, x_1, x_2$, $\lim_{k(2)\to\infty} f_{k(2)}(x_i)$ exists. This can be done because if a sequence converges then every subsequence converges also. Continue this way. Denote by $\{k\}$ the last of these subsequences. Thus for each x_i of these equally spaced points of the interval, $\lim_{k\to\infty} f_k(x_i)$ converges. Thus there exists m such that if $k, l \geq m$, then for each of these $x_i, i = 1, ..., n$,

$$|f_k(x_i) - f_l(x_i)| < \frac{\varepsilon}{4}$$

Let $x \in [0, T]$ be arbitrary. Then there is x_i such that $x_i \leq x < x_{i+1}$. Hence, for $k, l \geq m$,

$$|f_k(x) - f_l(x)| \leq |f_k(x) - f_k(x_i)| + |f_k(x_i) - f_l(x_i)| + |f_l(x_i) - f_l(x)|$$

By the assumption of equicontinuity, this implies

$$\begin{aligned} |f_k(x) - f_l(x)| &\leq |f_k(x) - f_k(x_i)| + |f_k(x_i) - f_l(x_i)| + |f_l(x_i) - f_l(x)| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon \end{aligned}$$

This has shown that for every $\varepsilon > 0$ there exists a subsequence $\{f_k\}$ with the property that $\sup_{x\in[0,T]} |f_k(x) - f_l(x)| < \varepsilon$ provided k, l are large enough. The argument also applies with no change to a given subsequence in place of the original sequence of functions. That is, for any subsequence of the original one, there is a further subsequence which satisfies the above condition. In what follows $\{f_{ik}\}_{k=1}^{\infty}$ will denote a subsequence of $\left\{f_{(i-1)k}\right\}_{k=1}^{\infty}$. Let $\varepsilon_i = 1/2^i$ so that $\lim_{i\to\infty} \varepsilon_i = 0$. Then let $\{f_{ik}\}_{k=1}^{\infty}$ denote a subsequence which corresponds to ε_i in the above construction. Consider the following diagram.

$$\begin{array}{c} f_{11}, f_{12}, f_{13}, f_{14}, \cdots \\ f_{21}, f_{22}, f_{23}, f_{24}, \cdots \\ f_{31}, f_{32}, f_{33}, f_{34}, \cdots \\ \vdots \end{array}$$

The Cantor diagonal sequence is $f_k = f_{kk}$ in the above. That is, it is the sequence

$$f_{11}, f_{22}, f_{33}, f_{44}, \cdots$$

Then from the construction, $f_j, f_{j+1}, f_{j+2}, \cdots$ is a subsequence of $\left\{f_{jk}\right\}_{k=1}^{\infty}$. Therefore, there exists m such that $k, l > m$,

$$\sup_{x\in[0,T]} |f_k(x) - f_l(x)| < \varepsilon_j$$

However, these ε_j converge to 0 and this shows that the diagonal sequence $\left\{f_j\right\}_{j=1}^{\infty}$ just described is a uniformly Cauchy sequence. ■

The process of obtaining this subsequence is called the Cantor diagonal process and occurs in other situations.

From this follows an easy corollary.

Corollary 6.11.3 *Let $\{f_n\}_{n=1}^{\infty} \subseteq C([0,T])$ be uniformly bounded and equicontinuous. Then there exists a subsequence which converges uniformly to a continuous function f defined on $[0,T]$.*

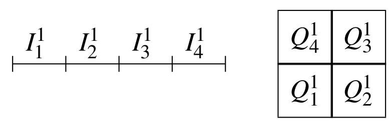
Proof: From Theorem 6.9.11 the uniformly Cauchy subsequence from the Ascoli Arzela theorem above converges uniformly to a function f. Now by Theorem 6.9.7, this function f is also continuous because, by this theorem, uniform convergence takes continuity with it and imparts it to the limit function. ■

This theorem and corollary are major results in the theory of differential equations. There are also infinite dimensional generalizations which have had great usefulness in the theory of nonlinear partial differential equations.

6.12 Space Filling Continuous Curves

When you have a function $\theta : [a,b] \to \mathbb{R}^p$, which is continuous, then the set of points obtained $\theta([a,b])$ is called a continuous curve. One of the horrifying examples which came out in the late nineteenth and early twentieth centuries was a space filling curve. Peano was the first to find one of these but there are many related results, all counter intuitive. Here is a simple example of how these things can occur.

Consider the following picture of a square Q having sides equal to 1 subdivided into four equal squares and an interval divided into four equal sub-intervals as shown.



Let $\theta_1\left(I_i^1\right) = Q_i^1$ as shown. Now the important thing about this is that if two of the intervals I_k^1, I_l^1 are adjacent, then so are are the squares Q_k^1 and Q_l^1. Subdivide I_i^1 into four equal intervals and Q_i^1 into four equal squares in exactly the same way. Denote the resulting intervals by I_i^2 and Q_i^2 so there are now 16 of these intervals and squares, still with the property that if I_k^2, and I_l^2 are adjacent, then so are Q_k^2 and Q_l^2. Continue this way such that each I_l^{n+1} is contained in some I_k^n and then $\theta_n\left(I_k^n\right) \supseteq \theta_{n+1}\left(I_l^{n+1}\right)$ and adjacent adjacent intervals are mapped to adjacent squares. The diameter of the union of two adjacent squares in the n^{th} stage of the construction is $2\left(2^{-n}\right)$.

Now if $x \in I$, the original interval, $x = \cap_{n=1}^{\infty} I_{k_n}^n$ for a sequence of nested intervals $I_{k_n}^{n+1}$. Define $\theta(x) \equiv \cap_{n=1}^{\infty} \theta_n\left(I_{k_n}^n\right)$. This is well defined because if you have two sequences of intervals having x equal to their intersection, then since the adjacent intervals go to adjacent squares, the diameter of $\theta_n\left(I_l^n\right) \cup \theta_n\left(I_k^n\right)$ for x in both I_k^n and I_l^n is no more than $2\left(2^{-n}\right)$ showing that the definitions of $\theta(x)$ do not differ by more than this for each n. This is also why θ is continuous. If $x_r \to x$, and $n \in \mathbb{N}$, eventually, for r large enough, $|x_r - x| < 5^{-n}$. It follows that both x, x_r are in a single I_k^n or else there are adjacent intervals I_k^n, I_l^n such that x is in one and x_r is in the other. Hence $\theta(x)$ is in some $\theta_n\left(I_k^n\right)$ and $\theta(x_r)$ is in either $\theta_n\left(I_l^n\right)$ an adjacent square or in $\theta_n\left(I_k^n\right)$. Either way, the construction implies that $\|\theta(x) - \theta(x_r)\| < 2\left(2^{-n}\right)$. Thus θ is continuous. It is clear that θ is onto because if $y \in Q$, then there is a nested sequence of squares $\left\{Q_{k_n}^n\right\}_{n=1}^{\infty}$ with $y = \cap_n Q_{k_n}^n$ and then $\theta\left(\cap_n I_{k_n}^n\right) \equiv \theta(x) = y$. However, θ is not one to one.

The last assertion about θ not being one to one relates to the fact that if θ were one to one, then its inverse would also be continuous and the interval and the square would be what is called homeomorphic. However, if you remove a point from the middle of the interval, the result is not connected but if you remove a point from the square, the result is connected. This could not happen because, as discussed earlier, the continuous image of a connected set is connected.

Proposition 6.12.1 *Let Q be a square. Then there exists a continuous mapping θ which maps the unit interval $[0,1]$ onto Q.*

Note that Q could have been an n dimensional cube. You would just need to subdivide intervals into more equal pieces. This shows that there are space filling continuous curves where $\theta([0,1]) = Q$ where Q is some sort of square or box, etc. Not surprisingly, there are generalizations. One generalization is to something called chainable continua.

6.13 Tietze Extension Theorem

This is about taking a real valued continuous function defined on a closed set in $\mathbb{F}^p$ and extending it to a continuous function which is defined on all of $\mathbb{F}^p$. First, review Lemma 6.0.7.

Lemma 6.13.1 *Let H,K be two nonempty disjoint closed subsets of $\mathbb{F}^p$. Then there exists a continuous function, $g : \mathbb{F}^p \to [-1/3, 1/3]$ such that*

$$g(H) = -1/3, \ \ g(K) = 1/3, g(\mathbb{F}^p) \subseteq [-1/3, 1/3].$$

Proof: Let $f(x) \equiv \frac{\operatorname{dist}(x,H)}{\operatorname{dist}(x,H)+\operatorname{dist}(x,K)}$. The denominator is never equal to zero because if $\operatorname{dist}(x,H) = 0$, then $x \in H$ because H is closed. (To see this, pick $h_k \in B(x,1/k) \cap H$. Then $h_k \to x$ and since H is closed, $x \in H$.) Similarly, if $\operatorname{dist}(x,K) = 0$, then $x \in K$ and so the denominator is never zero as claimed. Hence f is continuous and from its definition, $f = 0$ on H and $f = 1$ on K. Now let $g(x) \equiv \frac{2}{3}\left(f(x) - \frac{1}{2}\right)$. Then g has the desired properties. ■

Definition 6.13.2 *For $f : M \subseteq \mathbb{F}^p \to \mathbb{R}$, define $\|f\|_M$ as $\sup\{|f(x)| : x \in M\}$.*

Lemma 6.13.3 *Suppose M is a closed set in $\mathbb{F}^p$ and suppose $f : M \to [-1,1]$ is continuous at every point of M. Then there exists a function, g which is defined and continuous on all of $\mathbb{F}^p$ such that $\|f-g\|_M < \frac{2}{3}$, $g(\mathbb{F}^p) \subseteq [-1/3, 1/3]$.*

Proof: Let $H = f^{-1}([-1,-1/3]), K = f^{-1}([1/3,1])$. Thus H and K are disjoint closed subsets of M. Suppose first H,K are both nonempty. Then by Lemma 6.13.1 there exists g such that g is a continuous function defined on all of $\mathbb{F}^p$ and $g(H) = -1/3$, $g(K) = 1/3$, and $g(\mathbb{F}^p) \subseteq [-1/3, 1/3]$. It follows $\|f-g\|_M < 2/3$. If $H = \emptyset$, then f has all its values in $[-1/3, 1]$ and so letting $g \equiv 1/3$, the desired condition is obtained. If $K = \emptyset$, let $g \equiv -1/3$. ■

Lemma 6.13.4 *Suppose M is a closed set in $\mathbb{F}^p$ and suppose $f : M \to [-1,1]$ is continuous at every point of M. Then there exists a function g which is defined and continuous on all of $\mathbb{F}^p$ such that $g = f$ on M and g has its values in $[-1,1]$.*

Proof: Using Lemma 6.13.3, let $g_1(\mathbb{F}^p) \subseteq [-1/3, 1/3]$ and $\|f - g_1\|_M \leq \frac{2}{3}$. Suppose $g_1, \cdots, g_m$ have been chosen such that $g_j(\mathbb{F}^p) \subseteq [-1/3, 1/3]$ and

$$\left\| f - \sum_{i=1}^{m} \left(\frac{2}{3}\right)^{i-1} g_i \right\|_M < \left(\frac{2}{3}\right)^m. \tag{6.7}$$

This has been done for $m = 1$. Then

$$\left\| \left(\frac{3}{2}\right)^m \left(f - \sum_{i=1}^m \left(\frac{2}{3}\right)^{i-1} g_i\right) \right\|_M \leq 1$$

and so $\left(\frac{3}{2}\right)^m \left(f - \sum_{i=1}^m \left(\frac{2}{3}\right)^{i-1} g_i\right)$ can play the role of f in the first step of the proof. Therefore, there exists g_{m+1} defined and continuous on all of $\mathbb{F}^p$ such that its values are in $[-1/3, 1/3]$ and

$$\left\| \left(\frac{3}{2}\right)^m \left(f - \sum_{i=1}^m \left(\frac{2}{3}\right)^{i-1} g_i\right) - g_{m+1} \right\|_M \leq \frac{2}{3}.$$

Hence

$$\left\| \left(f - \sum_{i=1}^m \left(\frac{2}{3}\right)^{i-1} g_i\right) - \left(\frac{2}{3}\right)^m g_{m+1} \right\|_M \leq \left(\frac{2}{3}\right)^{m+1}.$$

It follows there exists a sequence, $\{g_i\}$ such that each has its values in $[-1/3, 1/3]$ and for every m 6.7 holds. Then let $g(x) \equiv \sum_{i=1}^\infty \left(\frac{2}{3}\right)^{i-1} g_i(x)$. It follows

$$|g(x)| \leq \left| \sum_{i=1}^\infty \left(\frac{2}{3}\right)^{i-1} g_i(x) \right| \leq \sum_{i=1}^m \left(\frac{2}{3}\right)^{i-1} \frac{1}{3} \leq 1$$

and $\left| \left(\frac{2}{3}\right)^{i-1} g_i(x) \right| \leq \left(\frac{2}{3}\right)^{i-1} \frac{1}{3}$ so the Weierstrass M test applies and shows convergence is uniform. Therefore g must be continuous by Theorem 6.9.7. The estimate 6.7 implies $f = g$ on M. ■

The following is the Tietze extension theorem.

Theorem 6.13.5 *Let M be a closed nonempty subset of $\mathbb{F}^p$ and let $f : M \to [a,b]$ be continuous at every point of M. Then there exists a function, g continuous on all of $\mathbb{F}^p$ which coincides with f on M such that $g(\mathbb{F}^p) \subseteq [a,b]$.*

Proof: Let $f_1(x) = 1 + \frac{2}{b-a}(f(x) - b)$. Then f_1 satisfies the conditions of Lemma 6.13.4 and so there exists $g_1 : \mathbb{F}^p \to [-1,1]$ such that g is continuous on $\mathbb{F}^p$ and equals f_1 on M. Let $g(x) = (g_1(x) - 1)\left(\frac{b-a}{2}\right) + b$. This works. ■

6.14 Exercises

1. Suppose $\{f_n\}$ is a sequence of decreasing positive functions defined on $[0, \infty)$ which converges pointwise to 0 for every $x \in [0, \infty)$. Can it be concluded that this sequence converges uniformly to 0 on $[0, \infty)$? Now replace $[0, \infty)$ with $(0, \infty)$. What can be said in this case assuming pointwise convergence still holds?

2. If $\{f_n\}$ and $\{g_n\}$ are sequences of functions defined on D which converge uniformly, show that if a, b are constants, then $af_n + bg_n$ also converges uniformly. If there exists a constant, M such that $|f_n(x)|, |g_n(x)| < M$ for all n and for all $x \in D$, show $\{f_n g_n\}$ converges uniformly. Let $f_n(x) \equiv 1/x$ for $x \in (0,1)$ and let $g_n(x) \equiv (n-1)/n$. Show $\{f_n\}$ converges uniformly on $(0,1)$ and $\{g_n\}$ converges uniformly but $\{f_n g_n\}$ fails to converge uniformly.

3. Show that if $x > 0, \sum_{k=0}^{\infty} \frac{x^k}{k!}$ converges uniformly on any interval of finite length.

4. Let $x \geq 0$ and consider the sequence $\left\{ \left(1 + \frac{x}{n}\right)^n \right\}$. Show this is an increasing sequence and is bounded above by $\sum_{k=0}^{\infty} \frac{x^k}{k!}$.

5. Show for every x, y real, $\sum_{k=0}^{\infty} \frac{(x+y)^k}{k!}$ converges and is $\left(\sum_{k=0}^{\infty} \frac{y^k}{k!}\right)\left(\sum_{k=0}^{\infty} \frac{x^k}{k!}\right)$.

6. Consider the series $\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$. Show this series converges uniformly on any interval of the form $[-M, M]$.

7. Formulate a theorem for a series of functions which will allow you to conclude the infinite series is uniformly continuous based on reasonable assumptions about the functions in the sum.

8. Find an example of a sequence of continuous functions such that each function is nonnegative and each function has a maximum value equal to 1 but the sequence of functions converges to 0 pointwise on $(0, \infty)$.

9. Suppose $\{f_n\}$ is a sequence of real valued functions which converges uniformly to a continuous function f. Can it be concluded the functions f_n are continuous? Explain.

10. Let $h(x)$ be a bounded continuous function. Show the function $f(x) = \sum_{n=1}^{\infty} \frac{h(nx)}{n^2}$ is continuous.

11. Let S be a any countable subset of $\mathbb{R}$. Show there exists a function f defined on $\mathbb{R}$ which is discontinuous at every point of S but continuous everywhere else. **Hint:** This is real easy if you do the right thing. It involves Theorem 6.9.15 and the Weierstrass M test.

12. By Theorem 6.10.3 there exists a sequence of polynomials converging uniformly to $f(x) = |x|$ on the interval $[-1, 1]$. Show there exists a sequence of polynomials, $\{p_n\}$ converging uniformly to f on $[-1, 1]$ which has the additional property that for all $n, p_n(0) = 0$.

13. If f is any continuous function defined on $[a, b]$, show there exists a series of the form $\sum_{k=1}^{\infty} p_k$, where each p_k is a polynomial, which converges uniformly to f on $[a, b]$. **Hint:** You should use the Weierstrass approximation theorem to obtain a sequence of polynomials. Then arrange it so the limit of this sequence is an infinite sum.

14. Sometimes a series may converge uniformly without the Weierstrass M test being applicable. Show $\sum_{n=1}^{\infty} (-1)^n \frac{x^2+n}{n^2}$ converges uniformly on $[0, 1]$ but does not converge absolutely for any $x \in \mathbb{R}$. To do this, it might help to use the partial summation formula, 5.7.

15. Suppose you have a collection of functions $S \subseteq C([0, T])$ which satisfy

$$\max_{x \in [0,T]} |f(x)| < M, \quad \sup_{0 \leq x < y \leq T} \frac{|f(x) - f(y)|}{|x - y|^{\gamma}} < K$$

where $\gamma \leq 1$. Show there is a uniformly convergent subsequence of S which converges uniformly to some continuous function. The second condition on f is called a Holder condition and such functions are said to be Holder continuous. These functions are denoted as $C^{0,\gamma}([0,T])$ and this little problem shows that the embedding of $C^{0,\gamma}([0,T])$ into $C([0,T])$ is compact.

16. Suppose $f \in C([a,b])$ and $\int_a^b f(x)x^n dx = 0$ for every $n \geq 0$, such that n is an integer. Show that then $f(x) = 0$ for all x. I am assuming you know about the integral from beginning calculus. This will be developed later in more generality. **Hint:** Use Weierstrass approximation theorem.

17. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and T periodic. Thus $f(x+kT) = f(x)$ for $k \in \mathbb{Z}$. Show that there exists a sequence of polynomials $\{p_n\}$ converging uniformly to f on $[-T/2, T/2]$ such that $p_n(T/2) = p_n(-T/2)$. **Hint:** Say $\{\hat{p}_n\}$ converges uniformly to f on $[-T/2, T/2]$. Consider
$$p_n(x) \equiv f(T/2) + \hat{p}_n(x) - \left[\left(\frac{\hat{p}_n(T/2) - \hat{p}_n(-T/2)}{T}\right)(x+T/2) + \hat{p}_n(-T/2)\right]$$
If $\tilde{p}_n$ denotes the T periodic extension of p_n, explain why $\tilde{p}_n$ converges uniformly to f on $\mathbb{R}$.

18. Is it possible to get a continuous onto function $f : [0,1] \to P$ where P is the Cantor set?

19. Show there exists a continuous function $\theta : P \to [0,1]$ such that θ is onto where P is the Cantor set. Show there exists a continuous function which maps the Cantor set onto $[0,1] \times [0,1]$. You might recall that the Cantor set does not even contain any intervals so this is very surprising.

20. For P the Cantor set, show there is a continuous, onto function $f : P \to [0,1] \times [0,1]$. This may be a little easier than what is in the chapter. Can you extend f to all of $[0,1]$ with $f(x) \in [0,1] \times [0,1]$? **Hint:** You might use Tietze extension theorem on the components of f.

21. Let K be a nonempty compact subset in $\mathbb{F}^p$. For P the Cantor set, there is a continuous function $f : P \to K$ which is onto. Try and show this. Will it be possible to extend f to all of $[0,1]$ if K is not connected? **Hint:** Try and show that for every n, there are finitely many closed balls having radius $1/n$ whose union contains K. Then for B one of these closed balls, you could consider $K \cap B$ as another compact set.

22. If $f : K \to \mathbb{R}^q$ is continuous and one to one and K is compact, show that $f^{-1} : f(K) \to K$ must be continuous.

Chapter 7

The Derivative

Some functions have them and some don't. Some have them at some points and not at others. This chapter is on the derivative. Functions which have derivatives are somehow better than those which don't. To begin with it is necessary to discuss the concept of a limit of a function. This is a harder concept than continuity and it is also harder than the concept of the limit of a sequence or series although that is similar. One cannot make any rational sense of the concept of derivative without an understanding of limits of a function. This is the main reason for considering the notion of limit.

7.1 Limit of a Function

For now, I will continue considering functions defined on a subset $D(f)$ of $\mathbb{F}^p$ having values in $\mathbb{F}^q$.

Definition 7.1.1 *Let x be a limit point of $D(f)$. Then $\lim_{y\to x} f(y) = L$ means that for every $\varepsilon > 0$ there is $\delta > 0$ such that whenever $0 < \|y-x\| < \delta$, with $y \in D(f)$, then $\|f(y) - L\| < \varepsilon$.*

Definition 7.1.2 *Let x be a limit point of $D(f)$ and let f have values in $\mathbb{R}$. Then $\lim_{y\to x} f(y) = \infty$ means that for l, there is $\delta > 0$ such that if $0 < \|y-x\| < \delta$, with $y \in D(f)$, then $f(y) > l$. A similar definition holds to define $\lim_{y\to x} f(y) = -\infty$. If $D(f)$ contains an interval (a,∞), then $\lim_{x\to\infty} f(x) = L \in \mathbb{F}^p$ means that for every $\varepsilon > 0$ there exists l such that if $x > l$, then $\|f(x) - L\| < \varepsilon$. $\lim_{x\to-\infty} f(x) = L$ is defined similarly.*

I will leave for the reader the appropriate definition in the case that $x \to \infty$ and $f(x) \to \infty$ and other such cases.

Theorem 7.1.3 *If $\lim_{y\to x} f(y) = L$ and $\lim_{y\to x} f(y) = L_1$, then $L = L_1$. Uniqueness also holds for one sided limits and for limits as $x \to \infty$ or $x \to -\infty$.*

Proof: Let $\varepsilon > 0$ be given. There exists $\delta > 0$ such that if $0 < |y-x| < \delta$, then $|f(y) - L| < \varepsilon$, $|f(y) - L_1| < \varepsilon$. Therefore, for such y,(It exists because x is a limit point) $|L - L_1| \leq |L - f(y)| + |f(y) - L_1| < \varepsilon + \varepsilon = 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary, this shows $L = L_1$. The argument is exactly the same in the case of one sided limits. You simply need to have some y close enough to x on one side and in the case of limits at $\pm\infty$, you use y such that $|y|$ is sufficiently large. ■

In the special case that f is defined near a point x, we sometimes speak of left and right limits by restricting the domain to be either those $y < x$ or those $y > x$. When this is done, one writes $\lim_{y\to x+} f(y)$ or $\lim_{y\to x-} f(y)$.

The first thing to do is to give an easier to use description in terms of sequences.

Proposition 7.1.4 *Let x be a limit point of $D(f)$. Then $\lim_{y\to x} f(y) = L \in \mathbb{F}^q$ if and only if whenever $x_n \to x$ for each $x_n \neq x$, the x_n distinct points, it follows that $f(x_n) \to L$.*

Proof: $\Rightarrow$ Let $x_n \to x$ where no x_n equals x. Let $\varepsilon > 0$ be given. By assumption, $|f(y) - L| < \varepsilon$ whenever $0 < |y-x| < \delta$ for some δ. However, for all n large enough, $0 < |x_n - x| < \delta$ and so $|f(x_n) - L| < \varepsilon$. Hence $f(x_n) \to L$.

$\Leftarrow$ Suppose the condition on the sequences holds. If the condition for the limit does not hold, then there exists $\varepsilon > 0$ such that no matter how small δ, there will be $0 <$

$|y-x|<\delta, y \in D(f)$, and yet $|f(y)-L| \geq \varepsilon$. Now let $\delta_1 = 1$. There exists $x_1 \neq x$ with $x_1 \in B(x,\delta_1) \cap D(f)$ and $|f(x_1)-L| \geq \varepsilon$. Let $\delta_2 \equiv \min\left(\frac{1}{2}, \frac{1}{2}|x-x_1|\right)$. Now pick $x_2 \in B(x,\delta_2), x_2 \neq x$ such that $|f(x_2)-L| \geq \varepsilon$. Let $\delta_3 \equiv \min\left(\frac{1}{2^3}, \frac{1}{2}|x-x_1|, \frac{1}{2}|x-x_2|\right)$ and pick $x_3 \in B(x,\delta_3)$ with $|f(x_3)-L| \geq \varepsilon, x_3 \neq x$. Continue this way to generate a sequence of distinct points $\{x_n\}$, none equal to x which converges to x. Then $L = \lim_{n\to\infty} f(x_n)$ because of the condition on limits of the sequence so eventually $|L-f(x_n)| < \varepsilon$, contrary to the construction of the x_n. ■

The value of a function at x is irrelevant to the value of the limit at x! This must always be kept in mind. In fact, it is not necessary for f to even be defined at the limit point. All interesting limits are this way. You do not evaluate interesting limits by computing $f(x)$! It may be the case that $f(x)$ is right but this is merely a happy coincidence when it occurs and as explained below in Theorem 7.1.8, this is sometimes equivalent to f being continuous at x.

Theorem 7.1.5 *In this theorem, x is always a limit point of $D(f)$. Suppose that both $\lim_{y\to x} f(y) = L$ and $\lim_{y\to x} g(y) = K$ where K and L are numbers, not $\pm\infty$. Then if a, b are numbers,*

$$\lim_{y\to x}(af(y)+bg(y)) = aL+bK, \tag{7.1}$$

$$\lim_{y\to x} fg(y) = LK \tag{7.2}$$

and if $K \neq 0$,

$$\lim_{y\to x}\frac{f(y)}{g(y)} = \frac{L}{K}. \tag{7.3}$$

Also, if h is a continuous function defined in some interval containing L, then

$$\lim_{y\to x} h\circ f(y) = h(L). \tag{7.4}$$

Suppose f is real valued and $\lim_{y\to x} f(y) = L$. If $f(y) \leq a$ all y near x either to the right or to the left of x, then $L \leq a$ and if $f(y) \geq a$ then $L \geq a$.

Proof: All of these claims follow from Proposition 7.1.4 and the theorems on limits of sequences Theorem 4.4.8. For example, consider 7.4. Letting x be a limit point and $x_n \to x$, then by assumption $f(x_n) \to L$ and so, by continuity of h, it follows that $h(f(x_n)) \to h(L)$. If $f(y) \leq a$ for all y less than a and near to x, then if $x_n \to x$ from the left, eventually $f(x_n) \leq a$ and so $L \leq a$ also. The other case is similar.

A very useful theorem for finding limits is called the squeezing theorem.

Theorem 7.1.6 *Suppose f,g,h are real valued functions and that*

$$\lim_{x\to a} f(x) = L = \lim_{x\to a} g(x)$$

and for all x near a, $f(x) \leq h(x) \leq g(x)$.Then $\lim_{x\to a} h(x) = L$.

Proof: If $L \geq h(x)$, then $|h(x)-L| \leq |f(x)-L|$. If $L < h(x)$, then $|h(x)-L| \leq |g(x)-L|$. Therefore,

$$|h(x)-L| \leq |f(x)-L| + |g(x)-L|.$$

Now let $\varepsilon > 0$ be given. There exists δ_1 such that if $0 < |x-a| < \delta_1, |f(x) - L| < \varepsilon/2$ and there exists δ_2 such that if $0 < |x-a| < \delta_2$, then $|g(x) - L| < \varepsilon/2$. Letting $0 < \delta \leq \min(\delta_1, \delta_2)$, if $0 < |x-a| < \delta$, then

$$|h(x) - L| \leq |f(x) - L| + |g(x) - L| < \varepsilon/2 + \varepsilon/2 = \varepsilon. \blacksquare$$

Proposition 7.1.7 *If $x \in D(f)$ and x is a limit point of $D(f)$, then f is continuous at x if and only if* $\lim_{y \to x} f(y) = f(x)$.

Proof: $\Rightarrow$Let $y_n \to x$ where $y_n \in D(f)$ and the y_n are distinct. Then by continuity of f at x, it follows that $\lim_{n\to\infty} f(y_n) = f(x)$.

$\Leftarrow$ Let $\varepsilon > 0$ be given. Then there is $\delta > 0$ such that whenever $0 < \|y - x\| < \delta$, it will follow that $\|f(y) - f(x)\| < \varepsilon$. However, if $y = x$, then $\|f(y) - f(x)\| = 0 < \varepsilon$ and so the conditions for continuity are satisfied. $\blacksquare$

The case of this proposition which is of the most interest here in this chapter is the following theorem. Intervals are of the form $(a,b), [a,b), (a,b]$, or $[a,b]$, Endpoints of an interval are clearly limit points of the interval.

Theorem 7.1.8 *For $f : I \to \mathbb{R}$, and I is an interval. then f is continuous at $x \in I$ if and only if* $\lim_{y\to x} f(y) = f(x)$.

Example 7.1.9 *Find* $\lim_{x\to 3} \frac{x^2-9}{x-3}$.

Note that $\frac{x^2-9}{x-3} = x+3$ whenever $x \neq 3$. Therefore, if $0 < |x-3| < \varepsilon$,

$$\left| \frac{x^2-9}{x-3} - 6 \right| = |x+3-6| = |x-3| < \varepsilon.$$

It follows from the definition that this limit equals 6.

You should be careful to note that in the definition of limit, the variable **never equals the thing it is getting close to**. In this example, x is never equal to 3. This is very significant because, in interesting limits, the function whose limit is being taken will not be defined at the point of interest. The habit students acquire of plugging in the point to take the limit is only good on useless and uninteresting limits which are not good for anything other than to give a busy work exercise, deceiving people into thinking they know what is going on.

Example 7.1.10 *Let*

$$f(x) = \frac{x^2-9}{x-3} \text{ if } x \neq 3.$$

How should f be defined at $x = 3$ so that the resulting function will be continuous there?

The limit of this function equals 6 because for $x \neq 3$, $\frac{x^2-9}{x-3} = \frac{(x-3)(x+3)}{x-3} = x+3$. Therefore, by Theorem 7.1.8 it is necessary to define $f(3) \equiv 6$.

Example 7.1.11 *Find* $\lim_{x\to\infty} \frac{x}{1+x}$.

Write $\frac{x}{1+x} = \frac{1}{1+(1/x)}$. Now it seems clear that $\lim_{x\to\infty} 1 + (1/x) = 1 \neq 0$. Therefore, Theorem 7.1.5 implies $\lim_{x\to\infty} \frac{x}{1+x} = \lim_{x\to\infty} \frac{1}{1+(1/x)} = \frac{1}{1} = 1$.

Example 7.1.12 *Show* $\lim_{x\to a}\sqrt{x}=\sqrt{a}$ *whenever* $a\geq 0$. *In the case that* $a=0$, *take the limit from the right.*

There are two cases. First consider the case when $a>0$. Let $\varepsilon>0$ be given. Multiply and divide by $\sqrt{x}+\sqrt{a}$. This yields

$$\left|\sqrt{x}-\sqrt{a}\right|=\left|\frac{x-a}{\sqrt{x}+\sqrt{a}}\right|.$$

Now let $0<\delta_1<a/2$. Then if $|x-a|<\delta_1, x>a/2$ and so

$$\left|\sqrt{x}-\sqrt{a}\right|=\left|\frac{x-a}{\sqrt{x}+\sqrt{a}}\right|\leq\frac{|x-a|}{\left(\sqrt{a}/\sqrt{2}\right)+\sqrt{a}}\leq\frac{2\sqrt{2}}{\sqrt{a}}\left|x-a\right|.$$

Now let $0<\delta\leq\min\left(\delta_1,\frac{\varepsilon\sqrt{a}}{2\sqrt{2}}\right)$. Then for $0<|x-a|<\delta$,

$$\left|\sqrt{x}-\sqrt{a}\right|\leq\frac{2\sqrt{2}}{\sqrt{a}}\left|x-a\right|<\frac{2\sqrt{2}}{\sqrt{a}}\frac{\varepsilon\sqrt{a}}{2\sqrt{2}}=\varepsilon.$$

Next consider the case where $a=0$. In this case, let $\varepsilon>0$ and let $\delta=\varepsilon^2$. Then if $0<x-0<\delta=\varepsilon^2$, it follows that $0\leq\sqrt{x}<\left(\varepsilon^2\right)^{1/2}=\varepsilon$.

7.2 Exercises

1. Find the following limits if possible
   (a) $\lim_{x\to 0+}\frac{|x|}{x}$
   (b) $\lim_{x\to 0+}\frac{x}{|x|}$
   (c) $\lim_{x\to 0-}\frac{|x|}{x}$
   (d) $\lim_{x\to 4}\frac{x^2-16}{x+4}$
   (e) $\lim_{x\to 3}\frac{x^2-9}{x+3}$
   (f) $\lim_{x\to -2}\frac{x^2-4}{x-2}$
   (g) $\lim_{x\to\infty}\frac{x}{1+x^2}$
   (h) $\lim_{x\to\infty}-2\frac{x}{1+x^2}$
2. Find $\lim_{h\to 0}\frac{\frac{1}{(x+h)^3}-\frac{1}{x^3}}{h}$.
3. Find $\lim_{x\to 4}\frac{\sqrt[4]{x}-\sqrt{2}}{\sqrt{x}-2}$.
4. Find $\lim_{x\to\infty}\frac{\sqrt[5]{3x}+\sqrt[4]{x}+7\sqrt{x}}{\sqrt{3x+1}}$.
5. Find $\lim_{x\to\infty}\frac{(x-3)^{20}(2x+1)^{30}}{\left(2x^2+7\right)^{25}}$.
6. Find $\lim_{x\to 2}\frac{x^2-4}{x^3+3x^2-9x-2}$.
7. Find $\lim_{x\to\infty}\left(\sqrt{1-7x+x^2}-\sqrt{1+7x+x^2}\right)$.
8. Prove Theorem 7.1.3 for right, left and limits as $y\to\infty$.
9. Prove from the definition that $\lim_{x\to a}\sqrt[3]{x}=\sqrt[3]{a}$ for all $a\in\mathbb{R}$. **Hint:** You might want to use the formula for the difference of two cubes,
$$a^3-b^3=(a-b)\left(a^2+ab+b^2\right).$$
10. Prove Theorem 7.1.8 from the definitions of limit and continuity.

11. Find $\lim_{h\to 0} \frac{(x+h)^3 - x^3}{h}$

12. Find $\lim_{h\to 0} \frac{\frac{1}{x+h} - \frac{1}{x}}{h}$

13. Find $\lim_{x\to -3} \frac{x^3+27}{x+3}$

14. Find $\lim_{h\to 0} \frac{\sqrt{(3+h)^2} - 3}{h}$ if it exists.

15. Find the values of x for which $\lim_{h\to 0} \frac{\sqrt{(x+h)^2} - x}{h}$ exists and find the limit.

16. Find $\lim_{h\to 0} \frac{\sqrt[3]{(x+h)} - \sqrt[3]{x}}{h}$ if it exists. Here $x \neq 0$.

17. Suppose $\lim_{y\to x+} f(y) = L_1 \neq L_2 = \lim_{y\to x-} f(y)$. Show $\lim_{y\to x} f(x)$ does not exist. **Hint:** Roughly, the argument goes as follows: For $|y_1 - x|$ small and $y_1 > x$, $|f(y_1) - L_1|$ is small. Also, for $|y_2 - x|$ small and $y_2 < x$, $|f(y_2) - L_2|$ is small. However, if a limit existed, then $f(y_2)$ and $f(y_1)$ would both need to be close to some number and so both L_1 and L_2 would need to be close to some number. However, this is impossible because they are different.

18. Suppose f is an increasing function defined on $[a,b]$. Show f must be continuous at all but a countable set of points. **Hint:** Explain why every discontinuity of f is a jump discontinuity and

$$f(x-) \equiv \lim_{y\to x-} f(y) \leq f(x) \leq f(x+) \equiv \lim_{y\to x+} f(y)$$

with $f(x+) > f(x-)$. Now each of these intervals $(f(x-), f(x+))$ at a point, x where a discontinuity happens has positive length and they are disjoint. Furthermore, they have to all fit in $[f(a), f(b)]$. How many of them can there be which have length at least $1/n$?

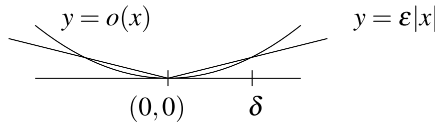
19. Let $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$. Find $\lim_{x\to 0}(\lim_{y\to 0} f(x,y))$, $\lim_{y\to 0}(\lim_{x\to 0} f(x,y))$. If you did it right you got -1 for one answer and 1 for the other. What does this tell you about interchanging limits?

20. The whole presentation of limits above is too specialized. Let D be the domain of a function f. A point x not necessarily in D, is said to be a limit point of D if $B(x,r)$ contains a point of D not equal to x for every $r > 0$. Now define the concept of limit in the same way as above and show that the limit is well defined if it exists. That is, if x is a limit point of D and $\lim_{y\to x} f(x) = L_1$ and $\lim_{y\to x} f(x) = L_2$, then $L_1 = L_2$. Is it possible to take a limit of a function at a point not a limit point of D? What would happen to the above property of the limit being well defined? Is it reasonable to define continuity at isolated points, those points which are not limit points, in terms of a limit as is often done in calculus books?

21. If f is an increasing function which is bounded above by a constant M, show that $\lim_{x\to\infty} f(x)$ exists. Give a similar theorem for decreasing functions.

7.3 The Definition of the Derivative

The following picture of a function $y = o(x)$ is an example of one which appears to be tangent to the line $y = 0$ at the point $(0,0)$.



You see in this picture, the graph of the function $y = \varepsilon|x|$ also where $\varepsilon > 0$ is just a positive number. Note there exists $\delta > 0$ such that if $|x| < \delta$, then $|o(x)| < \varepsilon|x|$ or in other words, $\frac{|o(x)|}{|x|} < \varepsilon$. You might draw a few other pictures of functions which would have the appearance of being tangent to the line $y = 0$ at the point $(0,0)$ and observe that in every case, it will follow that for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $0 < |x| < \delta$, then

$$\frac{|o(x)|}{|x|} < \varepsilon. \tag{7.5}$$

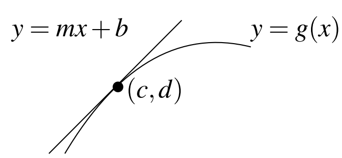
In other words, a reasonable way to say a function is tangent to the line $y = 0$ at $(0,0)$ is to say for all $\varepsilon > 0$ there exists $\delta > 0$ such that 7.5 holds. In other words, the function $y = o(x)$ is tangent at $(0,0)$ if and only if $\lim_{x\to 0} \frac{|o(x)|}{|x|} = 0$. More generally, even if the function has values in $\mathbb{F}$ or something even more general where it is not possible to draw pictures, the following is the definition of $o(x)$.

Definition 7.3.1 *A function $y = k(x)$ is said to be $o(x)$ if*

$$\lim_{x\to 0} \frac{|k(x)|}{|x|} = 0 \tag{7.6}$$

As was just discussed, in the case where $x \in \mathbb{R}$ and k is a function having values in $\mathbb{R}$ this is geometrically the same as saying the function is tangent to the line $y = 0$ at the point $(0,0)$. This terminology is used like an adjective. $k(x)$ is $o(x)$ means 7.6 holds. Thus $o(x) = 5o(x)$, $o(x) + o(x) = o(x)$, etc. The usage is very imprecise and sloppy, leaving out exactly the details which are of absolutely no significance in what is about to be discussed. It is this sloppiness which makes the notation so useful. It prevents you from fussing with things **which do not matter**. This takes some getting used to.

Now consider the case of the function $y = g(x)$ tangent to $y = b + mx$ at the point (c,d).



Thus, in particular, $g(c) = b + mc = d$. Then letting $x = c + h$, it follows x is close to c if and only if h is close to 0. Consider then the two functions

$$y = g(c+h),\ y = b + m(c+h).$$

If they are tangent as shown in the above picture, you should have the function

$$\begin{aligned} k(h) &\equiv g(c+h) - (b + m(c+h)) = g(c+h) - (b + mc) - mh \\ &= g(c+h) - g(c) - mh \end{aligned}$$

tangent to $y = 0$ at the point $(0,0)$. As explained above, the precise meaning of this function being tangent as described is to have $k(h) = o(h)$. This motivates (I hope) the following definition of the derivative which is the precise definition free of pictures and heuristics.

Definition 7.3.2 *Let g be a $\mathbb{F}$ (either $\mathbb{C}$ or $\mathbb{R}$) valued function defined on an open set in $\mathbb{F}$ containing c. Then $g'(c)$ is the number, if it exists, which satisfies*

$$(g(c+h) - g(c)) - g'(c)h = o(h)$$

where $o(h)$ is defined in Definition 7.3.1.

The above definition is more general than what will be extensively discussed here. I will usually consider the case where the function is defined on some interval contained in $\mathbb{R}$. In this context, where the function is defined on a subset of $\mathbb{R}$, the definition of derivative can also be extended to include right and left derivatives.

Definition 7.3.3 *Let g be a function defined on an interval, $[c,b)$. Then $g'_+(c)$ is the number, if it exists, which satisfies*

$$(g_+(c+h) - g_+(c)) - g'_+(c)h = o(h)$$

where $o(h)$ is defined in Definition 7.3.1 except you only consider positive h. Thus

$$\lim_{h \to 0+} \frac{|o(h)|}{|h|} = 0.$$

This is the derivative from the right. Let g be a function defined on an interval, $(a,c]$. Then $g'_-(c)$ is the number, if it exists, which satisfies

$$(g_-(c+h) - g_-(c)) - g'_-(c)h = o(h)$$

where $o(h)$ is defined in Definition 7.3.1 except you only consider negative h. Thus

$$\lim_{h \to 0-} \frac{|o(h)|}{|h|} = 0.$$

This is the derivative from the left.

I will not pay any attention to these distinctions from now on. In particular I will not write g'_- and g'_+ unless it is necessary. If the domain of a function defined on a subset of $\mathbb{R}$ is not open, it will be understood that at an endpoint, the derivative meant will be the appropriate derivative from the right or the left. First I need to show this is well defined because there cannot be two values for $g'(c)$.

Theorem 7.3.4 *The derivative is well defined because if*

$$\begin{aligned} (g(c+h) - g(c)) - m_1 h &= o(h) \\ (g(c+h) - g(c)) - m_2 h &= o(h) \end{aligned} \tag{7.7}$$

then $m_1 = m_2$.

Proof: Suppose 7.7. Then subtracting these, $(m_2 - m_1) h = o(h) - o(h) = o(h)$ and so dividing by $h \neq 0$ and then taking a limit as $h \to 0$ gives $m_2 - m_1 = \lim_{h\to 0} \frac{o(h)}{h} = 0$. Note the same argument holds for derivatives from the right or the left also. ■

Observation 7.3.5 *The familiar formula from calculus applies.*

$$g'(x) = \lim_{h\to 0} \frac{g(x+h) - g(x)}{h}$$

This is because the right side is

$$\lim_{h\to 0} \frac{g'(x) h + o(h)}{h} = g'(x).$$

So why am I fussing with little o notation? It is because functions of variables not in $\mathbb{R}$ must also be considered, although not so much in this book, but in this setting, you must do something like what I have just been discussing in this special case and in fact, the above definition continues to apply with no change in all situations while geometric notions involving "slope" don't.

Now the derivative has been defined, here are some properties.

Lemma 7.3.6 *Suppose $g'(c)$ exists. Then there exists $\delta > 0$ such that if $|h| < \delta$,*

$$|g(c+h) - g(c)| < \left(\left|g'(c)\right| + 1\right) |h| \tag{7.8}$$

$$o(|g(c+h) - g(c)|) = o(h) \tag{7.9}$$

g is continuous at c.

Proof: This follows from the definition of $g'(c)$.

$$(g(c+h) - g(c)) - g'(c) h = o(h)$$

and so there exists $\delta > 0$ such that if $0 < |h| < \delta$,

$$\frac{|(g(c+h) - g(c)) - g'(c) h|}{|h|} < 1$$

By the triangle inequality,

$$|g(c+h) - g(c)| - \left|g'(c) h\right| \leq \left|(g(c+h) - g(c)) - g'(c) h\right| < |h|$$

and so

$$|g(c+h) - g(c)| < \left(\left|g'(c)\right| + 1\right) |h|$$

Next consider the second claim. By definition of the little o notation, there exists a $\delta_1 > 0$ such that if

$$|g(c+h) - g(c)| < \delta_1,$$

then

$$o(|g(c+h) - g(c)|) < \frac{\varepsilon}{|g'(c)| + 1} |g(c+h) - g(c)|. \tag{7.10}$$

But from the first inequality, if $|h| < \delta$, then $|g(c+h) - g(c)| < (|g'(c)| + 1)|h|$ and so for $|h| < \min\left(\delta, \frac{\delta_1}{(|g'(c)|+1)}\right)$, it follows $|g(c+h) - g(c)| < (|g'(c)| + 1)|h| < \delta_1$ and so from 7.10,

$$\begin{aligned} o(|g(c+h) - g(c)|) &< \frac{\varepsilon}{|g'(c)| + 1} |g(c+h) - g(c)| \\ &< \frac{\varepsilon}{|g'(c)| + 1} \left(|g'(c)| + 1\right)|h| = \varepsilon |h| \end{aligned}$$

and this shows $\lim_{h \to 0} \frac{o(|g(c+h)-g(c)|)}{|h|} = 0$ because for nonzero h small enough,

$$\frac{o(|g(c+h) - g(c)|)}{|h|} < \varepsilon.$$

This proves 7.9.

The assertion about continuity follows from 7.8. Just let $h = x - c$ and the formula gives the following for $|x - c|$ small enough.

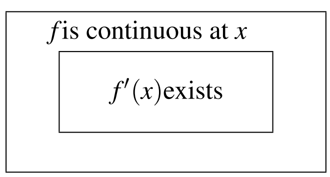
$$|g(x) - g(c)| < \left(|g'(c)| + 1\right)|x - c| \blacksquare$$

Of course some functions do not have derivatives at some points.

Example 7.3.7 *Let $f(x) = |x|$. Show $f'(0)$ does not exist.*

If $f'(0)$ did exist, then whenever $h_n \to 0$ where h_n are distinct, then $\lim_{n\to\infty} \frac{|0+h_n| - |h_n|}{h_n} = L$ for some L. However, if $h_n \to 0$ with each $h_n > 0$, you get 1 for the limit and if $h_n \to 0$ with each $h_n < 0$, then you get -1 for the limit. Thus the limit does not exist.

The following diagram shows how continuity at a point and differentiability there are related.



7.4 Continuous and Nowhere Differentiable

How bad can it get in terms of a continuous function not having a derivative at some points? It turns out it can be the case the function is nowhere differentiable but everywhere continuous. An example of such a pathological function different than the one I am about to present was discovered by Weierstrass in 1872. However, Bolzano was the first to produce a function in the 1830's which was continuous and nowhere differentiable although he did not show this completely.

Lemma 7.4.1 *Suppose $f'(x)$ exists and let c be a number. Then letting $g(x) \equiv f(cx)$,*

$$g'(x) = cf'(cx).$$

Here the derivative refers to either the derivative, the left derivative, or the right derivative. Also, if $f(x) = a + bx$, then $f'(x) = b$ where again, f' refers to either the left derivative, right derivative or derivative. Furthermore, in the case where $f(x) = a + bx$, $f(x+h) - f(x) = bh$.

Proof: It is known from the definition that $f(x+h) - f(x) - f'(x)h = o(h)$. Therefore,

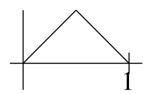
$$g(x+h) - g(x) = f(c(x+h)) - f(cx) = f'(cx)ch + o(ch)$$

and so $g(x+h) - g(x) - cf'(cx)h = o(ch) = o(h)$ and so this proves the first part of the lemma. Now consider the last claim.

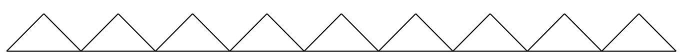
$$\begin{aligned} f(x+h) - f(x) &= a + b(x+h) - (a + bx) = bh \\ &= bh + 0 = bh + o(h). \end{aligned}$$

Thus $f'(x) = b$. ■

Now consider the following description of a function. The following is the graph of the function on $[0,1]$.



The height of the function is 1/2 and the slope of the rising line is 1 while the slope of the falling line is -1. Now extend this function to the whole real line to make it periodic of period 1. This means $f(x+n) = f(x)$ for all $x \in \mathbb{R}$ and $n \in \mathbb{Z}$, the integers. In other words to find the graph of f on $[1,2]$ you simply slide the graph of f on $[0,1]$ a distance of 1 to get the same tent shaped thing on $[1,2]$. Continue this way. The following picture illustrates what a piece of the graph of this function looks like. Some might call it an infinite sawtooth.



Now define

$$g(x) \equiv \sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k f\left(4^k x\right).$$

Letting $M_k = (3/4)^{-k}$, an application of the Weierstrass M test shows g is everywhere continuous. This is because each function in the sum is continuous and the series converges uniformly on $\mathbb{R}$.

Let $\delta_m = \pm \frac{1}{4}(4^{-m})$ where we assume $m > 2$. That of interest will be $m \to \infty$.

$$\frac{g(x+\delta_m) - g(x)}{\delta_m} = \frac{\sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k \left(f\left(4^k(x+\delta_m)\right) - f\left(4^k x\right)\right)}{\delta_m}$$

If you take $k > m$,

$$\begin{aligned} f\left(4^k(x+\delta_m)\right) - f\left(4^k x\right) &= f\left(4^k\left(x \pm \frac{1}{4}(4^{-m})\right)\right) - f\left(4^k x\right) \\ &= f\left(4^k x \pm \overbrace{\frac{1}{4}4^{k-m}}^{\text{integer}}\right) - f\left(4^k x\right) = 0 \end{aligned}$$

Therefore,

$$\frac{g(x+\delta_m)-g(x)}{\delta_m} = \frac{1}{\delta_m}\sum_{k=0}^{m}\left(\frac{3}{4}\right)^k \left(f\left(4^k(x+\delta_m)\right)-f\left(4^k x\right)\right)$$

The absolute value of the last term in the sum is

$$\left|\left(\frac{3}{4}\right)^m (f(4^m(x+\delta_m))-f(4^m x))\right|$$

and we choose the sign of δ_m such that both $4^m(x+\delta_m)$ and $4^m x$ are in some interval $[k/2,(k+1)/2)$ which is certainly possible because the distance between these two points is 1/4 and such half open intervals include all of $\mathbb{R}$. Thus, since f has slope ± 1 on the interval just mentioned,

$$\left|\left(\frac{3}{4}\right)^m (f(4^m(x+\delta_m))-f(4^m x))\right| = \left(\frac{3}{4}\right)^m 4^m|\delta_m| = 3^m|\delta_m|$$

As to the other terms, $0\leq f(x)\leq 1/2$ and so

$$\left|\sum_{k=0}^{m-1}\left(\frac{3}{4}\right)^k\left(f\left(4^k(x+\delta_m)\right)-f\left(4^k x\right)\right)\right| \leq \sum_{k=0}^{m-1}\left(\frac{3}{4}\right)^k = \frac{1-(3/4)^m}{1/4} = 4-4\left(\frac{3}{4}\right)^m$$

Thus

$$\left|\frac{g(x+\delta_m)-g(x)}{\delta_m}\right| \geq 3^m - \left(4-4\left(\frac{3}{4}\right)^m\right) \geq 3^m-4$$

Since $\delta_m\to 0$ as $m\to\infty$, $g'(x)$ does not exist because the difference quotients are not bounded. ■

This proves the following theorem.

Theorem 7.4.2 *There exists a function defined on $\mathbb{R}$ which is continuous and bounded but fails to have a derivative at any point.*

7.5 Finding the Derivative

Obviously there need to be simple ways of finding the derivative when it exists. There are rules of derivatives which make finding the derivative very easy. In the following theorem, the derivative could refer to right or left derivatives as well as regular derivatives.

Theorem 7.5.1 *Let a,b be numbers and suppose $f'(t)$ and $g'(t)$ exist. Then the following formulas are obtained.*

$$(af+bg)'(t) = af'(t)+bg'(t). \tag{7.11}$$

$$(fg)'(t) = f'(t)g(t)+f(t)g'(t). \tag{7.12}$$

The formula, 7.12 is referred to as the product rule.

If $f'(g(t))$ exists and $g'(t)$ exists, then $(f\circ g)'(t)$ also exists and

$$(f\circ g)'(t) = f'(g(t))g'(t).$$

This is called the chain rule. In this rule, for the sake of simiplicity, assume the derivatives are real derivatives, not derivatives from the right or the left. If $f(t) = t^n$ *where* n *is any integer, then*

$$f'(t) = nt^{n-1}. \tag{7.13}$$

Also, whenever $f'(t)$ *exists,* $f'(t) = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$ *where this definition can be adjusted in the case where the derivative is a right or left derivative by letting* $h > 0$ *or* $h < 0$ *only and considering a one sided limit. This is equivalent to* $f'(t) = \lim_{s\to t} \frac{f(s)-f(t)}{t-s}$ *with the limit being one sided in the case of a left or right derivative.*

Proof: 7.11 is left for you. Consider 7.12

$$fg(t+h) - fg(t) = f(t+h)g(t+h) - f(t)g(t+h) + f(t)g(t+h) - f(t)g(t)$$

$$= g(t+h)(f(t+h) - f(t)) + f(t)(g(t+h) - g(t))$$

$$\begin{aligned} &= g(t+h)\left(f'(t)h + o(h)\right) + f(t)\left(g'(t)h + o(h)\right) \\ &= g(t)f'(t)h + f(t)g'(t)h + f(t)o(h) \\ &\quad + (g(t+h) - g(t))f'(t)h + g(t+h)o(h) \end{aligned}$$

$$= g(t)f'(t)h + f(t)g'(t)h + o(h)$$

because by Lemma 7.3.6, g is continuous at t and so $(g(t+h) - g(t))f'(t)h = o(h)$. While $f(t)o(h)$ and $g(t+h)o(h)$ are both $o(h)$. This proves 7.12.

Next consider the chain rule. By Lemma 7.3.6 again,

$$f\circ g(t+h) - f\circ g(t) = f(g(t+h)) - f(g(t))$$

$$= f(g(t) + (g(t+h) - g(t))) - f(g(t))$$

$$= f'(g(t))(g(t+h) - g(t)) + o((g(t+h) - g(t)))$$

$$= f'(g(t))(g(t+h) - g(t)) + o(h)$$

$$\begin{aligned} &= f'(g(t))\left(g'(t)h + o(h)\right) + o(h) \\ &= f'(g(t))g'(t)h + o(h). \end{aligned}$$

This proves the chain rule.

Now consider the claim about $f(t) = t^n$ for n an integer. If $n = 0, 1$ the desired conclusion follows from Lemma 7.4.1. Suppose the claim is true for $n \geq 1$. Then let $f_{n+1}(t) = t^{n+1} = f_n(t)t$ where $f_n(t) \equiv t^n$. Then by the product rule, induction and the validity of the assertion for $n = 1$,

$$f'_{n+1}(t) = f'_n(t)t + f_n(t) = tnt^{n-1} + t^n = nt^{n+1}$$

and so the assertion is proved for all $n \geq 0$. Consider now $n = -1$.

$$\begin{aligned} (t+h)^{-1} - t^{-1} &= \frac{-1}{t(t+h)}h = \frac{-1}{t^2}h + \left(\frac{-1}{t(t+h)} + \frac{1}{t^2}\right)h \\ &= \frac{-1}{t^2}h + \frac{h^2}{t^2(t+h)} = -\frac{1}{t^2}h + o(h) = (-1)t^{-2}h + o(h) \end{aligned}$$

Therefore, the assertion is true for $n = -1$. Now consider $f(t) = t^{-n}$ where n is a positive integer. Then $f(t) = (t^n)^{-1}$ and so by the chain rule,

$$f'(t) = (-1)(t^n)^{-2} nt^{n-1} = -nt^{-n-1}.$$

This proves 7.13.

Finally, if $f'(t)$ exists,

$$f'(t)h + o(h) = f(t+h) - f(t).$$

Divide by h and take the limit as $h \to 0$, either a regular limit or a limit from one side or the other in the case of a right or left derivative.

$$f'(t) = \lim_{h \to 0} \left(\frac{f(t+h) - f(t)}{h} + \frac{o(h)}{h} \right) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}. \blacksquare$$

Note the last part is the usual definition of the derivative given in beginning calculus courses. There is nothing wrong with doing it this way from the beginning for a function of only one variable but it is not the right way to think of the derivative and does not generalize to the case of functions of many variables where the definition given in terms of $o(h)$ does.

Corollary 7.5.2 *Let $f'(t), g'(t)$ both exist and $g(t) \neq 0$, then the quotient rule holds.*

$$\left(\frac{f}{g} \right)' = \frac{f'(t)g(t) - f(t)g'(t)}{g(t)^2}$$

Proof: This is left to you. Use the chain rule and the product rule. $\blacksquare$

Higher order derivatives are defined in the usual way. $f'' \equiv (f')'$ etc. Also the Leibniz notation is defined by

$$\frac{dy}{dx} = f'(x) \text{ where } y = f(x)$$

and the second derivative is denoted as $\frac{d^2y}{dx^2}$ with various other higher order derivatives defined in the usual way.

The chain rule has a particularly attractive form in Leibniz's notation. Suppose $y = g(u)$ and $u = f(x)$. Thus $y = g \circ f(x)$. Then from the above theorem

$$(g \circ f)'(x) = g'(f(x)) f'(x) = g'(u) f'(x)$$

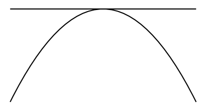
or in other words, $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$. Notice how the du cancels. This particular form is a very useful crutch and is used extensively in applications. Of course the problem is that we really don't know what du is. Nevertheless, it is great notation and in fact this can be made precise, but this book is on classical analysis.

7.6 Local Extreme Points

When you are on top of a hill, you are at a local maximum although there may be other hills higher than the one on which you are standing. Similarly, when you are at the bottom of a valley, you are at a local minimum even though there may be other valleys deeper than the one you are in. The word, "local" is applied to the situation because if you confine your attention only to points close to your location, you are indeed at either the top or the bottom.

Definition 7.6.1 *Let $f : D(f) \to \mathbb{R}$ where here $D(f)$ is only assumed to be some subset of $\mathbb{F}$. Then $x \in D(f)$ is a local minimum (maximum) if there exists $\delta > 0$ such that whenever $y \in B(x,\delta) \cap D(f)$, it follows $f(y) \geq (\leq) f(x)$. The plural of minimum is minima and the plural of maximum is maxima.*

Derivatives can be used to locate local maxima and local minima. The following picture suggests how to do this. This picture is of the graph of a function having a local maximum and the tangent line to it.



Note how the tangent line is horizontal. If you were not at a local maximum or local minimum, the function would be falling or climbing and the tangent line would not be horizontal.

Theorem 7.6.2 *Suppose $f : U \to \mathbb{R}$ where U is an open subset of $\mathbb{F}$ and suppose $x \in U$ is a local maximum or minimum. Then $f'(x) = 0$.*

Proof: Suppose x is a local maximum and let $\delta > 0$ is so small that $B(x,\delta) \subseteq U$. Then for $|h| < \delta$, both x and $x+h$ are contained in $B(x,\delta) \subseteq U$. Then letting h be real and positive,

$$f'(x)h + o(h) = f(x+h) - f(x) \leq 0.$$

Then dividing by h it follows from Theorem 7.1.5 on Page 130,

$$f'(x) = \lim_{h\to 0}\left(f'(x) + \frac{o(h)}{h}\right) = \lim_{h\to 0}\left(\frac{1}{h}(f(x+h) - f(x))\right) \leq 0$$

Next let $|h| < \delta$ and h is real and negative. Then

$$f'(x)h + o(h) = f(x+h) - f(x) \leq 0.$$

Then dividing by h,

$$f'(x) = \lim_{h\to 0} f'(x) + \frac{o(h)}{h} = \lim_{h\to 0}\left(\frac{1}{h}(f(x+h) - f(x))\right) \geq 0$$

Thus $f'(x) = 0$. The case where x is a local minimum is handled similarly. Alternatively, you could apply what was just shown to $-f(x)$. ■[1]

Points at which the derivative of a function equals 0 are sometimes called critical points. Included in the set of critical points are those points where f' fails to exist.

7.7 Exercises

1. If $f'(x) = 0$, is it necessary that x is either a local minimum or local maximum? **Hint:** Consider $f(x) = x^3$.

[1] Actually, the case where the function is defined on an open subset of $\mathbb{C}$ and yet has real values is not too interesting. However, this is information which depends on the theory of functions of a complex variable which is not considered yet.

2. A continuous function f defined on $[a,b]$ is to be maximized. It was shown above in Theorem 7.6.2 that if the maximum value of f occurs at $x \in (a,b)$, and if f is differentiable there, then $f'(x) = 0$. However, this theorem does not say anything about the case where the maximum of f occurs at either a or b. Describe how to find the point of $[a,b]$ where f achieves its maximum. Does f have a maximum? Explain.

3. Show that if the maximum value of a function f differentiable on $[a,b]$ occurs at the right endpoint, then for all $h > 0, f'(b)h \geq 0$. This is an example of a variational inequality. Describe what happens if the maximum occurs at the left end point and give a similar variational inequality. What is the situation for minima?

4. Find the maximum and minimum values and the values of x where these are achieved for the function $f(x) = x + \sqrt{25 - x^2}$.

5. A piece of wire of length L is to be cut in two pieces. One piece is bent into the shape of an equilateral triangle and the other piece is bent to form a square. How should the wire be cut to maximize the sum of the areas of the two shapes? How should the wire be bent to minimize the sum of the areas of the two shapes? **Hint:** Be sure to consider the case where all the wire is devoted to one of the shapes separately. This is a possible solution even though the derivative is not zero there.

6. Lets find the point on the graph of $y = \frac{x^2}{4}$ which is closest to $(0,1)$. One way to do it is to observe that a typical point on the graph is of the form $\left(x, \frac{x^2}{4}\right)$ and then to minimize the function $f(x) = x^2 + \left(\frac{x^2}{4} - 1\right)^2$. Taking the derivative of f yields $x + \frac{1}{4}x^3$ and setting this equal to 0 leads to the solution, $x = 0$. Therefore, the point closest to $(0,1)$ is $(0,0)$. Now lets do it another way. Lets use $y = \frac{x^2}{4}$ to write $x^2 = 4y$. Now for (x,y) on the graph, it follows it is of the form $\left(\sqrt{4y}, y\right)$. Therefore, minimize $f(y) = 4y + (y-1)^2$. Take the derivative to obtain $2 + 2y$ which requires $y = -1$. However, on this graph, y is never negative. What on earth is the problem?

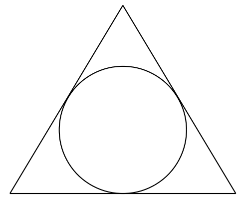
7. Find the dimensions of the largest rectangle that can be inscribed in the ellipse, $\frac{x^2}{9} + \frac{y^2}{4} = 1$.

8. A function f, is said to be odd if $f(-x) = -f(x)$ and a function is said to be even if $f(-x) = f(x)$. Show that if f is even, then f' is odd and if f is odd, then f' is even. Sketch the graph of a typical odd function and a typical even function.

9. Find the point on the curve, $y = \sqrt{25 - 2x}$ which is closest to $(0,0)$.

10. A street is 200 feet long and there are two lights located at the ends of the street. One of the lights is $\frac{1}{8}$ times as bright as the other. Assuming the brightness of light from one of these street lights is proportional to the brightness of the light and the reciprocal of the square of the distance from the light, locate the darkest point on the street.

11. Find the volume of the smallest right circular cone which can be circumscribed about a sphere of radius 4 inches.



12. Show that for r a rational number and $y = x^r$, it must be the case that if this function is differentiable, then $y' = rx^{r-1}$.

13. Let f be a continuous function defined on $[a,b]$. Let $\varepsilon > 0$ be given. Show there exists a polynomial p such that for all $x \in [a,b]$,

$$|f(x) - p(x)| < \varepsilon.$$

This follows from the Weierstrass approximation theorem, Theorem 6.10.3. Now here is the interesting part. Show there exists a function g which is also continuous on $[a,b]$ and for all $x \in [a,b]$,

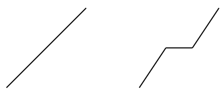
$$|f(x) - g(x)| < \varepsilon$$

but g has no derivative at any point. Thus there are enough nowhere differentiable functions that any continuous function is uniformly close to one. Explain why every continuous function is the uniform limit of nowhere differentiable functions. Also explain why every nowhere differentiable continuous function is the uniform limit of polynomials. **Hint:** You should look at the construction of the nowhere differentiable function which is everywhere continuous and bounded, given above.

14. Consider the following nested sequence of compact sets, $\{P_n\}$. Let $P_1 = [0,1]$, $P_2 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$, etc. To go from P_n to P_{n+1}, delete the open interval which is the middle third of each closed interval in P_n. Let $P = \cap_{n=1}^{\infty} P_n$. By Problem 16 on Page 74, $P \neq \emptyset$. If you have not worked this exercise, now is the time to do it. Show the total length of intervals removed from $[0,1]$ is equal to 1. If you feel ambitious also show there is a one to one onto mapping of $[0,1]$ to P. The set P is called the Cantor set. Thus P has the same number of points in it as $[0,1]$ in the sense that there is a one to one and onto mapping from one to the other even though the length of the intervals removed equals 1. **Hint:** There are various ways of doing this last part but the most enlightenment is obtained by exploiting the construction of the Cantor set rather than some silly representation in terms of sums of powers of two and three. All you need to do is use the theorems in the chapter on set theory related to the Schroder Bernstein theorem and show there is an onto map from the Cantor set to $[0,1]$. If you do this right it will provide a construction which is very useful to prove some even more surprising theorems which you may encounter later if you study compact metric spaces. The Cantor set is just a simple version of what is seen in some vegetables. Note in the following picture of Romanesco broccoli, the spirals of points each of which is a spiral of points each of which is a spiral of points...

15. $\uparrow$ Consider the sequence of functions defined in the following way. 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. The process of modifying a nonconstant section of the graph of this function is illustrated in the following picture.



Show $\{f_n\}$ converges uniformly on $[0,1]$. If $f(x)=\lim_{n\to\infty} f_n(x)$, show that $f(0)=0$, $f(1)=1$, f is continuous, and $f'(x)=0$ for all $x\notin P$ where P is the Cantor set of Problem 14. This function is called the Cantor function.It is a very important example to remember especially for those who like mathematical pathology. Note it has derivative equal to zero on all those intervals which were removed and whose total length was equal to 1 and yet it succeeds in climbing from 0 to 1. Isn't this amazing? **Hint:** This isn't too hard if you focus on getting a careful estimate on the difference between two successive functions in the list considering only a typical small interval in which the change takes place. The above picture should be helpful.

16. Let
$$f(x)=\begin{cases} 1 \text{ if } x\in\mathbb{Q} \\ 0 \text{ if } x\notin\mathbb{Q}\end{cases}$$
Now let $g(x)=x^2 f(x)$. Find where g is continuous and differentiable if anywhere.

7.8 Mean Value Theorem

The mean value theorem is possibly the most important theorem about the derivative of a function of one variable. It pertains only to a real valued function of a real variable. The best versions of many other theorems depend on this fundamental result. The mean value theorem is based on the following special case known as Rolle's theorem[2]. It is an existence

[2] Rolle is remembered for Rolle's theorem and not for anything else he did. Ironically, he did not like calculus. This may be because, until Bolzano and Cauchy and later Weierstrass, there were aspects of calculus which were fairly fuzzy. In particular, the notion of differential was not precise and yet it was being used.

theorem and like the other existence theorems in analysis, it depends on the completeness axiom.

Theorem 7.8.1 *Suppose $f:[a,b]\to\mathbb{R}$ is continuous, $f(a)=f(b)$, and $f:(a,b)\to\mathbb{R}$ has a derivative at every point of (a,b). Then there exists $x\in(a,b)$ such that $f'(x)=0$.*

Proof: Suppose first that $f(x)=f(a)$ for all $x\in[a,b]$. Then any $x\in(a,b)$ is a point such that $f'(x)=0$. If f is not constant, either there exists $y\in(a,b)$ such that $f(y)>f(a)$ or there exists $y\in(a,b)$ such that $f(y)<f(b)$. In the first case, the maximum of f is achieved at some $x\in(a,b)$ and in the second case, the minimum of f is achieved at some $x\in(a,b)$. Either way, Theorem 7.6.2 implies $f'(x)=0$. ■

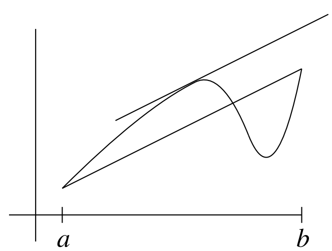
The next theorem is known as the Cauchy mean value theorem. It is the best version of this important theorem.

Theorem 7.8.2 *Suppose f,g are continuous on $[a,b]$ and differentiable on (a,b). Then there exists $x\in(a,b)$ such that $f'(x)(g(b)-g(a))=g'(x)(f(b)-f(a))$.*

Proof: Let $h(x)\equiv f(x)(g(b)-g(a))-g(x)(f(b)-f(a))$.Then letting $x=a$ and then letting $x=b$, a short computation shows $h(a)=h(b)$. Also, h is continuous on $[a,b]$ and differentiable on (a,b). Therefore Rolle's theorem applies and there exists $x\in(a,b)$ such that

$$h'(x)=f'(x)(g(b)-g(a))-g'(x)(f(b)-f(a))=0. \blacksquare$$

Letting $g(x)=x$, the usual version of the mean value theorem is obtained. Here is the usual picture which describes the theorem.



Corollary 7.8.3 *Let f be a continuous real valued function defined on $[a,b]$ and differentiable on (a,b). Then there exists $x\in(a,b)$ such that $f(b)-f(a)=f'(x)(b-a)$.*

Corollary 7.8.4 *Suppose $f'(x)=0$ for all $x\in(a,b)$ where $a\geq-\infty$ and $b\leq\infty$. Then $f(x)=f(y)$ for all $x,y\in(a,b)$. Thus f is a constant.*

Proof: If this is not true, there exists x_1 and x_2 such that $f(x_1)\neq f(x_2)$. Then by the mean value theorem,

$$0\neq\frac{f(x_1)-f(x_2)}{x_1-x_2}=f'(z)$$

for some z between x_1 and x_2. This contradicts the hypothesis that $f'(x)=0$ for all x. This proves the theorem in the case that f has real values. In the general case,

$$f(x+h)-f(x)-0h=o(h).$$

Then taking the real part of both sides,

$$\operatorname{Re} f(x+h) - \operatorname{Re} f(x) = \operatorname{Re} o(h) = o(h)$$

and so $\operatorname{Re} f'(x) = 0$ and by the first part, $\operatorname{Re} f$ must be a constant. The same reasoning applies to $\operatorname{Im} f$ and this proves the corollary. ■

Corollary 7.8.5 *Suppose $f : \mathbb{C} \to \mathbb{C}$ and $f'(x) = 0$ for all x. Then f is a constant.*

Proof: Let $t \in \mathbb{R}$ and consider $h(t) = f(x + t(y - x)) - f(x)$. Then by the chain rule,

$$h'(t) = f'(x + t(y - x))(y - x) = 0$$

and so by Corollary 7.8.4, h is a constant. In particular,

$$h(1) = f(y) - f(x) = h(0) = 0$$

which shows f is constant since x, y are arbitrary. ■

Corollary 7.8.6 *Suppose f has real values and $f'(x) > 0$ for all $x \in (a,b)$ where $a \geq -\infty$ and $b \leq \infty$. Then f is strictly increasing on (a,b). That is, if $x < y$, then $f(x) < f(y)$. If $f'(x) \geq 0$, then f is increasing in the sense that whenever $x < y$ it follows that $f(x) \leq f(y)$.*

Proof: Let $x < y$. Then by the mean value theorem, there exists $z \in (x, y)$ such that

$$0 < f'(z) = \frac{f(y) - f(x)}{y - x}.$$

Since $y > x$, it follows $f(y) > f(x)$ as claimed. Replacing $<$ by $\leq$ in the above equation and repeating the argument gives the second claim. ■

Corollary 7.8.7 *Suppose $f'(x) < 0$ for all $x \in (a,b)$ where $a \geq -\infty$ and $b \leq \infty$. Then f is strictly decreasing on (a,b). That is, if $x < y$, then $f(x) > f(y)$. If $f'(x) \leq 0$, then f is decreasing in the sense that for $x < y$, it follows that $f(x) \geq f(y)$*

Proof: Let $x < y$. Then by the mean value theorem, there exists $z \in (x, y)$ such that

$$0 > f'(z) = \frac{f(y) - f(x)}{y - x}.$$

Since $y > x$, it follows $f(y) < f(x)$ as claimed. The second claim is similar except instead of a strict inequality in the above formula, you put $\geq$. ■

7.9 Exercises

1. Sally drives her Saturn over the 110 mile toll road in exactly 1.3 hours. The speed limit on this toll road is 70 miles per hour and the fine for speeding is 10 dollars per mile per hour over the speed limit. How much should Sally pay?

2. Two cars are careening down a freeway in Utah weaving in and out of traffic. Car A passes car B and then car B passes car A as the driver makes obscene gestures. This infuriates the driver of car A who passes car B while firing his handgun at the driver of car B. Show there are at least two times when both cars have the same speed. Then show there exists at least one time when they have the same acceleration. The acceleration is the derivative of the velocity.

3. Show the cubic function $f(x) = 5x^3 + 7x - 18$ has only one real zero.

4. Suppose $f(x) = x^7 + |x| + x - 12$. How many solutions are there to the equation, $f(x) = 0$?

5. Let $f(x) = |x-7| + (x-7)^2 - 2$ on the interval $[6,8]$. Then $f(6) = 0 = f(8)$. Does it follow from Rolle's theorem that there exists $c \in (6,8)$ such that $f'(c) = 0$? Explain your answer.

6. Suppose f and g are differentiable functions defined on $\mathbb{R}$. Suppose also that it is known that $|f'(x)| > |g'(x)|$ for all x and that $|f'(t)| > 0$ for all t. Show that whenever $x \neq y$, it follows $|f(x) - f(y)| > |g(x) - g(y)|$. **Hint:** Use the Cauchy mean value theorem, Theorem 7.8.2.

7. Show that, like continuous functions, functions which are derivatives have the intermediate value property. This means that if $f'(a) < 0 < f'(b)$ then there exists $x \in (a,b)$ such that $f'(x) = 0$. **Hint:** Argue the minimum value of f occurs at an interior point of $[a,b]$.

8. Find an example of a function which has a derivative at every point but such that the derivative is not everywhere continuous.

9. Consider the function
$$f(x) \equiv \begin{cases} 1 \text{ if } x \geq 0 \\ -1 \text{ if } x < 0 \end{cases}.$$
Is it possible that this function could be the derivative of some function? Why?

10. Suppose $c \in I$, an open interval and that a function f, defined on I has $n+1$ derivatives. Then for each $m \leq n$ the following formula holds for $x \in I$.
$$f(x) = \sum_{k=0}^{m} f^{(k)}(c) \frac{(x-c)^k}{k!} + f^{(m+1)}(y) \frac{(x-c)^{m+1}}{(m+1)!} \tag{7.14}$$
where y is some point between x and c. Fix c, x in I. Let K be a number, depending on c, x such that
$$f(x) - \left(f(c) + \sum_{k=1}^{n} \frac{f^{(k)}(c)}{k!} (x-c)^k + K(x-c)^{n+1} \right) = 0$$
Now the idea is to find K. To do this, let
$$F(t) = f(x) - \left(f(t) + \sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!} (x-t)^k + K(x-t)^{n+1} \right)$$
Then $F(x) = F(c) = 0$. Therefore, by Rolle's theorem there exists y between c and x such that $F'(y) = 0$. Do the differentiation and solve for K. This is the main result on Taylor polynomials approximating a function f. The term $f^{(m+1)}(y) \frac{(x-c)^{m+1}}{(m+1)!}$ is called the Lagrange form of the remainder.

11. Let f be a real continuous function defined on the interval $[0,1]$. Also suppose $f(0)=0$ and $f(1)=1$ and $f'(t)$ exists for all $t\in(0,1)$. Show there exists n distinct points $\{s_i\}_{i=1}^n$ of the interval such that
$$\sum_{i=1}^n f'(s_i)=n.$$
Hint: Consider the mean value theorem applied to successive pairs in the following sum.
$$f\left(\frac{1}{3}\right)-f(0)+f\left(\frac{2}{3}\right)-f\left(\frac{1}{3}\right)+f(1)-f\left(\frac{2}{3}\right)$$

12. Now suppose $f:[0,1]\to\mathbb{R}$ is continuous and differentiable on $(0,1)$ and $f(0)=0$ while $f(1)=1$. Show there are distinct points $\{s_i\}_{i=1}^n\subseteq(0,1)$ such that
$$\sum_{i=1}^n \left(f'(s_i)\right)^{-1}=n.$$
Hint: Let $0=t_0<t_1<\cdots<t_n=1$ and pick $x_i\in f^{-1}(t_i)$ such that these x_i are increasing and $x_n=1,x_0=0$. Explain why you can do this. Then argue
$$t_{i+1}-t_i=f(x_{i+1})-f(x_i)=f'(s_i)(x_{i+1}-x_i)$$
and so
$$\frac{x_{i+1}-x_i}{t_{i+1}-t_i}=\frac{1}{f'(s_i)}$$
Now choose the t_i to be equally spaced.

13. Show that $(x+1)^{3/2}-x^{3/2}>2$ for all $x\geq 2$. Explain why for n a natural number larger than or equal to 1, there exists a natural number m such that $(n+1)^3>m^2>n^3$. **Hint:** Verify directly for $n=1$ and use the above inequality to take care of the case where $n\geq 2$. This shows that between the cubes of any two natural numbers there is the square of a natural number.

7.10 Derivatives of Inverse Functions

It happens that if f is a differentiable one to one function defined on an interval, $[a,b]$, and $f'(x)$ exists and is non zero then the inverse function f^{-1} has a derivative at the point $f(x)$. Recall that f^{-1} is defined according to the formula
$$f^{-1}(f(x))=x.$$

Let $f:[a,b]\to\mathbb{R}$ be a continuous function. Recall from Theorem 7.5.1
$$f'(a)\equiv\lim_{x\to a+}\frac{f(x)-f(a)}{x-a},\ f'(b)\equiv\lim_{x\to b-}\frac{f(x)-f(b)}{x-b}.$$

Recall the notation $x\to a+$ means that only $x>a$ are considered in the definition of limit, the notation $x\to b-$ defined similarly. Thus, this definition includes the derivative of f at the endpoints of the interval and to save notation,
$$f'(x_1)\equiv\lim_{x\to x_1}\frac{f(x)-f(x_1)}{x-x_1}$$
where it is understood that x is always in $[a,b]$.

Theorem 7.10.1 *Let $f : [a,b] \to \mathbb{R}$ be continuous and one to one. Suppose $f'(x_1)$ exists for some $x_1 \in [a,b]$ and $f'(x_1) \neq 0$. Then $\left(f^{-1}\right)'(f(x_1))$ exists and is given by the formula, $\left(f^{-1}\right)'(f(x_1)) = \frac{1}{f'(x_1)}$.*

Proof: As above, Lemma 6.4.3, f is either strictly increasing or strictly decreasing on $[a,b]$ and f^{-1} is continuous. Always y will be in the interval $f([a,b])$ if x_1 is at an end point. Then, from assumption that $f'(x_1)$ exists,

$$|y - f(x_1)| = \left|f\left(f^{-1}(y)\right) - f(x_1)\right| = \left|f'(x_1)\left(f^{-1}(y) - x_1\right) + o\left(f^{-1}(y) - x_1\right)\right|$$

by continuity, if $|y - f(x_1)|$ is small enough, then $\left|f^{-1}(y) - x_1\right|$ is small enough that

$$\left|o\left(f^{-1}(y) - x_1\right)\right| < \frac{|f'(x_1)|}{2}\left|f^{-1}(y) - x_1\right|.$$

Hence, if $|y - f(x_1)|$ is sufficiently small, then from the triangle inequality of the form $|p - q| \geq ||p| - |q||$,

$$\begin{aligned} |y - f(x_1)| &\geq \left|f'(x_1)\right|\left|f^{-1}(y) - x_1\right| - \frac{|f'(x_1)|}{2}\left|f^{-1}(y) - x_1\right| \\ &= \frac{|f'(x_1)|}{2}\left|f^{-1}(y) - x_1\right| \end{aligned}$$

It follows that for $|y - f(x_1)|$ small enough,

$$\left|\frac{o\left(f^{-1}(y) - x_1\right)}{y - f(x_1)}\right| \leq \left|\frac{o\left(f^{-1}(y) - x_1\right)}{f^{-1}(y) - x_1}\right| \frac{2}{|f'(x_1)|}$$

Then, using continuity of the inverse function again, it follows that if $|y - f(x_1)|$ is possibly still smaller, then $f^{-1}(y) - x_1$ is sufficiently small that the right side of the above inequality is no larger than ε. Since ε is arbitrary, it follows

$$o\left(f^{-1}(y) - x_1\right) = o(y - f(x_1))$$

Now from differentiability of f at x_1,

$$\begin{aligned} y - f(x_1) &= f\left(f^{-1}(y)\right) - f(x_1) = f'(x_1)\left(f^{-1}(y) - x_1\right) + o\left(f^{-1}(y) - x_1\right) \\ &= f'(x_1)\left(f^{-1}(y) - x_1\right) + o(y - f(x_1)) \\ &= f'(x_1)\left(f^{-1}(y) - f^{-1}(f(x_1))\right) + o(y - f(x_1)) \end{aligned}$$

Therefore,

$$f^{-1}(y) - f^{-1}(f(x_1)) = \frac{1}{f'(x_1)}(y - f(x_1)) + o(y - f(x_1))$$

From the definition of the derivative, this shows that $\left(f^{-1}\right)'(f(x_1)) = \frac{1}{f'(x_1)}$. ■

The following obvious corollary comes from the above by not bothering with end points. In this case, we can also consider the case where f is defined on an open set in $\mathbb{F}$ and has values in $\mathbb{F}$ where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. The new feature is that it might not make sense to consider one sided derivatives if $\mathbb{F} = \mathbb{C}$.

Corollary 7.10.2 *Let U be an open set in $\mathbb{F}$ and let $f : U \to \mathbb{F}$ be one to one and continuous such that $f'(x_1) \neq 0$ for $x_1 \in U$. Then, assuming $f(U)$ is an open set and f^{-1} is also continuous at $f(x_1)$,[3] it follows that $\left(f^{-1}\right)'(f(x_1)) = \frac{1}{f'(x_1)}$.*

Proof: The proof is exactly the same as the one given above except you don't consider the case of an endpoint and $|\cdot|$ refers to the absolute value in either $\mathbb{C}$ or $\mathbb{R}$.

Incidentally, the above argument works with virtually no change for n dimensional situations.

This is one of those theorems which is very easy to remember if you neglect the difficult questions and simply focus on formal manipulations. Consider the following.

$$f^{-1}(f(x)) = x.$$

Now use the chain rule on both sides to write

$$\left(f^{-1}\right)'(f(x)) f'(x) = 1,$$

and then divide both sides by $f'(x)$ to obtain

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)}.$$

Of course this gives the conclusion of the above theorem rather effortlessly and it is formal manipulations like this which aid in remembering formulas such as the one given in the theorem.

Example 7.10.3 *Let $f(x) = 1 + x^2 + x^3 + 7$. Show that f has an inverse and find $\left(f^{-1}\right)'(8)$.*

I am not able to find a formula for the inverse function. This is typical in useful applications so you need to get used to this idea. The methods of algebra are insufficient to solve hard problems in analysis. You need something more. The question is to determine whether f has an inverse. To do this,

$$f'(x) = 2x + 3x^2 + 7 > 0$$

By Corollary 7.8.6 on Page 147, this function is strictly increasing on $\mathbb{R}$ and so it has an inverse function although I have no idea how to find an explicit formula for this inverse function. However, I can see that $f(0) = 8$ and so by the formula for the derivative of an inverse function,

$$\left(f^{-1}\right)'(8) = \left(f^{-1}\right)'(f(0)) = \frac{1}{f'(0)} = \frac{1}{7}.$$

7.11 Derivatives and Limits of Sequences

When you have a function which is a limit of a sequence of functions, when can you say the derivative of the limit function is the limit of the derivatives of the functions in the sequence? The following theorem seems to be one of the best results available. It is based on the mean value theorem. Thus it is understood that the functions are real valued and defined on an interval of $\mathbb{R}$. First of all, recall Definition 6.9.6 on Page 117 listed here for convenience.

[3] In fact, this last assumption on the continuity of the inverse function is redundant but this is not a topic for this book. In case that U is an open interval this was proved earlier, but it also is true even if U is an open subset of $\mathbb{C}$ or $\mathbb{R}^p$ for $p > 1$.

Definition 7.11.1 *Let $\{f_n\}$ be a sequence of functions defined on D. Then $\{f_n\}$ is said to converge uniformly to f if it converges pointwise to f and for every $\varepsilon > 0$ there exists N such that for all $n \geq N$, $|f(x) - f_n(x)| < \varepsilon$ for all $x \in D$.*

To save on notation, denote by $\|k\| \equiv \sup\{|k(\xi)| : \xi \in D\}$. Then

$$\|k + l\| \leq \|k\| + \|l\| \tag{7.15}$$

because for each $\xi \in D, |k(\xi) + l(\xi)| \leq \|k\| + \|l\|$and taking sup yields 7.15. From the definition of uniform convergence, you see that f_n converges uniformly to f is the same as saying $\lim_{n\to\infty} \|f_n - f\| = 0$. Now here is the theorem. Note how the mean value theorem is one of the principal parts of the argument.

Theorem 7.11.2 *Let (a,b) be an open interval and let $f_k : (a,b) \to \mathbb{R}$ be differentiable and suppose there exists $x_0 \in (a,b)$ such that*

$$\{f_k(x_0)\} \text{ converges,}$$

$$\{f_k'\} \text{ converges uniformly to } g$$

Then there exists a function f defined on (a,b) such that

$$f_k \to f \text{ uniformly,}$$

and $f' = g$.

Proof: By the mean value theorem,

$$(f_k(x) - f_m(x)) - (f_k(x_0) - f_m(x_0)) = \left(f_k'(t_{km}) - f_m'(t_{km})\right)(x - x_0)$$

and so, if k, m is large enough,

$$|f_k(x) - f_m(x)| \leq |f_k(x_0) - f_m(x_0)| + \left\|f_k' - f_m'\right\| < \frac{\varepsilon}{2}$$

provided m, k are large enough. Therefore, $\|f_k - f_m\| < \varepsilon$ if k, m are large enough showing that $\{f_k\}$ converges uniformly to a continuous function f by Theorem 6.9.7. That is

$$\|f_k - f\| \to 0$$

I want to show that f' exists and equals g.

Let $c \in (a,b)$ and define

$$g_n(x,c) \equiv \begin{cases} \frac{f_n(x) - f_n(c)}{x - c} \text{ if } x \neq c \\ f_n'(c) \text{ if } x = c \end{cases}.$$

Thus $x \to g_n(x,c)$ is continuous.

Claim : For each c, $x \to g_n(x,c)$ converges uniformly to a continuous function h_c, on (a,b) and $h_c(c) = g(c)$.

Proof: Let $x \neq c$. Then by the mean value theorem applied to the function $x \to f_n(x) - f_m(x)$,

$$\begin{aligned} & \left|g_n(x,c) - g_m(x,c)\right| \\ = & \left|\frac{f_n(x) - f_m(x) - (f_n(c) - f_m(c))}{x-c}\right| \\ = & \left|f_n'(\xi) - f_m'(\xi)\right| \leq \left|f_n'(\xi) - g(\xi)\right| + \left|g(\xi) - f_m'(\xi)\right| \\ \leq & \left\|f_n' - g\right\| + \left\|f_m' - g\right\| \end{aligned}$$

By the assumption that $\{f_n'\}$ converges uniformly to g, it follows each of the last two terms converges to 0 as $n, m \to \infty$. If $x = c$, then

$$\left|g_n(c,c) - g_m(c,c)\right| = \left|f_n'(c) - f_m'(c)\right| \leq \left\|f_n' - g\right\| + \left\|f_m' - g\right\|$$

Thus $x \to g_n(x,c)$ is uniformly Cauchy and must converge uniformly to a continuous function h_c by Theorem 6.9.7 and Corollary 6.9.10. Also $h_c(c) = g(c)$ by the assumption that f_k' converges uniformly to g. This proves the claim.

Now to complete the proof of the theorem, for c given and $x \neq c$,

$$\frac{f(x) - f(c)}{x-c} = \lim_{n\to\infty} \frac{f_n(x) - f_n(c)}{x-c} = \lim_{n\to\infty} g_n(x,c) = h_c(x).$$

Since h_c is continuous,

$$f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x-c} = \lim_{x\to c} h_c(x) = h_c(c) = g(c). \blacksquare$$

7.12 Exercises

1. It was shown earlier that the n^{th} root of a positive number exists whenever n is a positive integer. Let $y = x^{1/n}$. Prove $y'(x) = \frac{1}{n}x^{(1/n)-1}$.

2. Now for positive x and p,q positive integers, $y = x^{p/q}$ is defined by $y = \sqrt[q]{x^p}$. Find and prove a formula for dy/dx.

3. For $1 \geq x \geq 0$, and $p \geq 1$, show that $(1-x)^p \geq 1 - px$. **Hint:** This can be done using the mean value theorem. Define $f(x) \equiv (1-x)^p - 1 + px$ and show that $f(0) = 0$ while $f'(x) \geq 0$ for all $x \in (0,1)$.

4. Using the result of Problem 3 establish Raabe's Test, an interesting variation on the ratio test. This test says the following. Suppose there exists a constant, C and a number p such that

$$\left|\frac{a_{k+1}}{a_k}\right| \leq 1 - \frac{p}{k+C} = \frac{1}{C+k}(C+k-p)$$

for all k large enough. Then if $p > 1$, it follows that $\sum_{k=1}^{\infty} a_k$ converges absolutely. **Hint:** Let $b_k \equiv k - 1 + C$ and note that for all k large enough, $b_k > 1$. Now conclude that there exists an integer, k_0 such that $b_{k_0} > 1$ and for all $k \geq k_0$ the given inequality above holds. Use Problem 3 to conclude that

$$\left|\frac{a_{k+1}}{a_k}\right| \leq 1 - \frac{p}{k+C} \leq \left(1 - \frac{1}{k+C}\right)^p = \left(\frac{b_k}{b_{k+1}}\right)^p$$

showing $|a_k| b_k^p$ is decreasing for $k \geq k_0$. Thus $|a_k| \leq M/b_k^p = M/(k-1+C)^p$. Now use comparison theorems and the p series to obtain the conclusion of the theorem.

5. The graph of a function $y = f(x)$ is said to be concave up or more simply "convex" if whenever (x_1, y_1) and (x_2, y_2) are two points such that $y_i \geq f(x_i)$, it follows that for each point, (x, y) on the straight line segment joining (x_1, y_1) and (x_2, y_2), $y \geq f(x)$. Show that if f is twice differentiable on an open interval, (a, b) and $f''(x) > 0$, then the graph of f is convex.

6. Show that if the graph of a function f defined on an interval (a, b) is convex, then if f' exists on (a, b), it must be the case that f' is a non decreasing function. Note you do not know the second derivative exists.

7. Convex functions defined in Problem 5 have a very interesting property. Suppose $\{a_i\}_{i=1}^n$ are all nonnegative, sum to 1, and suppose ϕ is a convex function defined on $\mathbb{R}$. Then
$$\phi\left(\sum_{k=1}^n a_k x_k\right) \leq \sum_{k=1}^n a_k \phi(x_k).$$
Verify this interesting inequality.

8. If ϕ is a convex function defined on $\mathbb{R}$, show that ϕ must be continuous at every point.

9. Prove the second derivative test. If $f'(x) = 0$ at $x \in (a, b)$, an interval on which f is defined and both f', f'' exist and are continuous on this interval, then if $f''(x) > 0$, it follows f has a local minimum at x and if $f''(x) < 0$, then f has a local maximum at x. Show that if $f''(x) = 0$ no conclusion about the nature of the critical point can be drawn. It might be a local minimum, local maximum or neither.

10. Recall the Bernstein polynomials which were used to prove the Weierstrass approximation theorem. For f a continuous function on $[0, 1]$,
$$p_n(x) = \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$$
It was shown these converge uniformly to f on $[0, 1]$. Now suppose f' exists and is continuous on $[0, 1]$. Show p_n' converges uniformly to f' on $[0, 1]$. **Hint:** Differentiate the above formula and massage to finally get
$$p_n'(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} \left(\frac{f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)}{1/n}\right) x^k (1-x)^{n-1-k}.$$
Then form the $(n-1)$ Bernstein polynomial for f' and show the two are uniformly close. You will need to estimate an expression of the form
$$f'\left(\frac{k}{n-1}\right) - \frac{f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)}{1/n}$$
which will be easy to do because of the mean value theorem and uniform continuity of f'.

11. In contrast to Problem 10, consider the sequence of functions
$$\{f_n(x)\}_{n=1}^{\infty} = \left\{\frac{x}{1+nx^2}\right\}_{n=1}^{\infty}.$$
Show it converges uniformly to $f(x) \equiv 0$. However, $f_n'(0)$ converges to 1, not $f'(0)$. **Hint:** To show the first part, find the value of x which maximizes the function $\left|\frac{x}{1+nx^2}\right|$. You know how to do this. Then plug it in and you will have an estimate sufficient to verify uniform convergence.

12. This next sequence of problems will give an independent treatment of the Riemann integral of piecewise continuous functions without the use of Riemann sums, being based instead on the mean value theorem and the Weierstrass theorem. For $p(x)$ a polynomial on $[a,b]$, let $P'(x) = p(x)$. Define $\int_a^b p(x)\,dx \equiv P(b) - P(a)$. Show, using the mean value theorem that this definition is well defined for polynomials and satisfies $\int_a^b (\alpha p + \beta q)\,dx = \alpha \int_a^b p\,dx + \beta \int_a^b q\,dx$. Also show $\int_a^b p\,dx + \int_b^c p\,dx = \int_a^c p\,dx$.

13. For f continuous on $[a,b]$, using the Weierstrass theorem, let $\|p_n - f\| \to 0$ where p_n is a polynomial and $\|g\| \equiv \max\{|g(x)| : x \in [a,b]\}$. Then define $\int_a^b f(x)\,dx \equiv \lim_{n\to\infty} \int_a^b p_n(x)\,dx$. Show this is well defined. **Hint:** If $\|p_n\|, \|\hat{p}_n\| \to 0$, then show $\|p_n - \hat{p}_m\|$ is small whenever n,m are large. Use the mean value theorem to verify that $\left|\int_a^b (p_n - \hat{p}_m)\,dx\right|$ is small. Thus $\left\{\int_a^b p_n\right\}$ is a Cauchy sequence and if another $\{\hat{p}_n\}$ is chosen, $\left\{\int_a^b \hat{p}_n\right\}$ is a Cauchy sequence which converges to the same thing.

14. For f continuous, show that if $F(x) \equiv \int_a^x f\,dt$, Then $F'(x) = f(x)$ so any continuous function has an antiderivative, and for any $a \neq b$, $\int_a^b f\,dx = G(b) - G(a)$ whenever $G' = f$ on the open interval determined by a,b and G continuous on the closed interval determined by a,b. Also verify that $\int_a^b (\alpha f + \beta g)\,dx = \alpha \int_a^b f\,dx + \beta \int_a^b g\,dx$, $\int_a^b f\,dx + \int_b^c f\,dx = \int_a^c f\,dx$, and $\int_a^b 1\,dx = b-a$, and $\int_a^b f\,dx = f(z)(b-a)$ for some $z \in (a,b)$. Also explain the change of variables formula from calculus, $\int_a^b f(g(x))\,g'(x)\,dx = F(g(b)) - F(g(a))$ and integration by parts.

15. If $f(x+) \equiv \lim_{y\to x+} f(y)$ and $f(x-) \equiv \lim_{y\to x-} f(y)$ both exist for all $x \in [a,b]$ and if f is continuous on $[\alpha_{k-1}, \alpha_k]$ where $a = \alpha_0 < \alpha_1 < \cdots < \alpha_n = b$, then define
$$\int_a^b f\,dx \equiv \sum_{k=1}^n \int_{\alpha_{k-1}}^{\alpha_k} f\,dx \equiv \sum_{k=1}^n G_k(f(\alpha_k-)) - G_k(f(\alpha_{k-1}+))$$
where $G_k' = f$ on (α_{k-1}, α_k) with G_k continuous on $[\alpha_{k-1}, \alpha_k]$. Show this coincides with the above definition on each sub interval and is a well defined way to define the integral of a function which is piecewise continuous, meaning: continuous on each sub interval and possessing right and left limits at every point.

16. Generalize differentiability to the case where f has values in $\mathbb{F}^p$. Simply replace $|\cdot|$ with $\|\cdot\|$ in the definition. Thus $o(u)$ means $\lim_{\|u\|\to 0} \frac{o(u)}{\|u\|} = 0$. Verify that if $f(t) = (f_1(t), \cdots, f_p(t))$ for $t \in [a,b]$, then f is differentiable (right or left differentiable at the end points) if and only if this is true for each of the component functions and that

$f'(t) = \left(f_1'(t), \cdots, f_p'(t)\right)$. Also show that it makes absolutely no difference whether we use the Euclidean norm $|\cdot|$ or the maximum norm $\|\cdot\|$ in defining $o(u)$.

Chapter 8

Power Series

8.1 Functions Defined in Terms of Series

It is time to consider functions other than polynomials. In particular it is time to give a mathematically acceptable definition of functions like $e^x, \sin(x)$ and $\cos(x)$. It has been assumed these functions are known from beginning calculus but this is a pretence. Most students who take calculus come through it without a complete understanding of the circular functions. This is because of the reliance on plane geometry in defining them. Fortunately, these functions can be completely understood in terms of power series rather than wretched plane geometry. The exponential function can also be defined in a simple manner using power series. It is tacitly assumed in this presentation that $x \in \mathbb{F}$, either $\mathbb{R}$ or $\mathbb{C}$.

Definition 8.1.1 *Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of numbers. The expression,*

$$\sum_{k=0}^{\infty} a_k (x-a)^k \tag{8.1}$$

is called a Taylor series centered at a. This is also called a power series centered at a. It is understood that x and $a \in \mathbb{F}$, that is, either $\mathbb{C}$ or $\mathbb{R}$.

In the above definition, x is a variable. Thus you can put in various values of x and ask whether the resulting series of numbers converges. Defining D to be the set of all values of x such that the resulting series does converge, define a new function f defined on D having values in $\mathbb{F}$ as $f(x) \equiv \sum_{k=0}^{\infty} a_k (x-a)^k$. This might be a totally new function, one which has no name. Nevertheless, much can be said about such functions. The following lemma is fundamental in considering the form of D which always turns out to be of the form $B(a,r)$ along with possibly some points z such that $|z-a| = r$. First here is a simple lemma which will be useful.

Lemma 8.1.2 $\lim_{n\to\infty} n^{1/n} = 1$.

Proof: It is clear $n^{1/n} \geq 1$. Let $n^{1/n} = 1 + e_n$ where $0 \leq e_n$. Then raising both sides to the n^{th} power for $n > 1$ and using the binomial theorem,

$$n = (1+e_n)^n = \sum_{k=0}^{n} \binom{n}{k} e_n^k \geq 1 + ne_n + (n(n-1)/2)\, e_n^2 \geq (n(n-1)/2)\, e_n^2$$

Thus $0 \leq e_n^2 \leq \frac{n}{n(n-1)} = \frac{1}{n-1}$. From this the desired result follows because $\left|n^{1/n} - 1\right| = e_n \leq \frac{1}{\sqrt{n-1}}$. ■

Theorem 8.1.3 *Let $\sum_{k=0}^{\infty} a_k (x-a)^k$ be a Taylor series. Then there exists $r \leq \infty$ such that the Taylor series converges absolutely if $|x-a| < r$. Furthermore, if $|x-a| > r$, the Taylor series diverges. If $\lambda < r$ then the Taylor series converges uniformly on the closed disk $|x-a| \leq \lambda$.*

Proof: Note $\limsup_{k\to\infty} \left|a_k (x-a)^k\right|^{1/k} = \limsup_{k\to\infty} |a_k|^{1/k} |x-a|$.Then by the root test, the series

converges absolutely if $|x-a| \limsup_{k\to\infty} |a_k|^{1/k} < 1$
diverges spectacularly if $|x-a| \limsup_{k\to\infty} |a_k|^{1/k} > 1$.

Thus define

$$r \equiv \begin{cases} 1/\limsup_{k\to\infty} |a_k|^{1/k} \text{ if } \infty > \limsup_{k\to\infty} |a_k|^{1/k} > 0 \\ \infty \text{ if } \limsup_{k\to\infty} |a_k|^{1/k} = 0 \\ 0 \text{ if } \limsup_{k\to\infty} |a_k|^{1/k} = \infty \end{cases}$$

Next let λ be as described. Then if $|x-a| \leq \lambda$, then

$$\limsup_{k\to\infty} \left|a_k (x-a)^k\right|^{1/k} = \limsup_{k\to\infty} |a_k|^{1/k} |x-a| \leq \lambda \limsup_{k\to\infty} |a_k|^{1/k} \leq \frac{\lambda}{r} < \alpha < 1$$

It follows that for all k large enough and such x, $\left|a_k (x-a)^k\right| < \alpha^k$. Then by the Weierstrass M test, convergence is uniform. ■

Note that the radius of convergence r is given by $\limsup_{k\to\infty} |a_k|^{1/k} r = 1$

Definition 8.1.4 *The number in the above theorem is called the radius of convergence and the set on which convergence takes place is called the disc of convergence.*

Now the theorem was proved using the root test but often you use the ratio test to find the radius of convergence. This kind of thing is typical in math and one must adjust to this fact. The proof of a theorem does not always yield a way to find the thing the theorem speaks about. The above is an existence theorem. There exists a disk of convergence from the above theorem. You find it in specific cases any way that is most convenient.

Example 8.1.5 *Find the disc of convergence of the Taylor series $\sum_{n=1}^{\infty} \frac{x^n}{n}$.*

Use Corollary 5.4.10. $\lim_{n\to\infty} \left(\frac{|x|^n}{n}\right)^{1/n} = \lim_{n\to\infty} \frac{|x|}{\sqrt[n]{n}} = |x|$ because, as shown earlier, $\lim_{n\to\infty} \sqrt[n]{n} = 1$ and so if $|x| < 1$ the series converges. The points satisfying $|z| = 1$ require special attention. When $x = 1$ the series diverges because it reduces to $\sum_{n=1}^{\infty} \frac{1}{n}$. At $x = -1$ the series converges because it reduces to $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ and the alternating series test applies and gives convergence. What of the other numbers z satisfying $|z| = 1$? It turns out this series will converge at all these numbers by the Dirichlet test.

Example 8.1.6 *Find the radius of convergence of $\sum_{n=1}^{\infty} \frac{n^n}{n!} x^n$.*

Apply the ratio test. Taking the ratio of the absolute values of the $(n+1)^{th}$ and the n^{th} terms

$$\frac{\frac{(n+1)^{(n+1)}}{(n+1)n!} |x|^{n+1}}{\frac{n^n}{n!} |x|^n} = (n+1)^n |x| n^{-n} = |x| \left(1+\frac{1}{n}\right)^n \to |x| e$$

Therefore the series converges absolutely if $|x| e < 1$ and diverges if $|x| e > 1$. Consequently, $r = 1/e$. This problem assumes that you remember from calculus the last limit. If not, this will be discussed later.

8.2 Operations on Power Series

It is desirable to be able to differentiate and multiply power series. Recall

$$f'(x) \equiv \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

Here h,x,f all can have values in $\mathbb{C}$. The definition is the same. The following theorem says you can differentiate power series in the most natural way on the disk of convergence, just as you would differentiate a polynomial. This theorem may seem obvious, but it is a serious mistake to think this. You usually cannot differentiate an infinite series whose terms are functions even if the functions are themselves polynomials. The following is special and pertains to power series. It is another example of the interchange of two limits, in this case, the limit involved in taking the derivative and the limit of the sequence of finite sums.

When you formally differentiate a series term by term, the result is called the derived series.

Theorem 8.2.1 *Let $\sum_{n=0}^{\infty} a_n (x-a)^n$ be a Taylor series having radius of convergence $R>0$ and let*

$$f(x) \equiv \sum_{n=0}^{\infty} a_n (x-a)^n \tag{8.2}$$

for $|x-a|<R$. Then

$$f'(x) = \sum_{n=0}^{\infty} a_n n (x-a)^{n-1} = \sum_{n=1}^{\infty} a_n n (x-a)^{n-1} \tag{8.3}$$

and this new differentiated power series, the derived series, has radius of convergence equal to R. Also, $f(x)$ given by the Taylor series is infinitely differentiable on the interior of its disk of convergence.

Proof: First consider the claim that the derived series has radius of convergence equal to R. Let $\hat{R}$ be the radius of convergence of the derived series. Then from Proposition 4.10.13 and Lemma 8.1.2,

$$\frac{1}{\hat{R}} \equiv \limsup_{n\to\infty} |a_n|^{1/n} n^{1/n} = \limsup_{n\to\infty} |a_n|^{1/n} \equiv \frac{1}{R},$$

and so $\hat{R}=R$. If $\limsup_{n\to\infty} |a_n|^{1/n} = 0$, then $\limsup_{n\to\infty} |a_n|^{1/n} n^{1/n}$ and in this case, the series and derived series both have radius of convergence equal to ∞.

Now let $r<R$, the radius of convergence of both series, and suppose $|x-a|<r$. Let δ be small enough that if $|h|<\delta$, then

$$|x+h-a| \leq |x-a|+|h| < r$$

also. Thus, $\limsup_{k\to\infty} |a_k|^{1/k} r < 1$.

Then for $|h|<\delta$, consider the difference quotient.

$$\frac{f(x+h)-f(x)}{h} = \frac{1}{h}\sum_{k=0}^{\infty} a_k \left((x+h-a)^k - (x-a)^k\right)$$

Using the binomial theorem,

$$\begin{aligned}
\frac{f(x+h)-f(x)}{h} &= \frac{1}{h}\sum_{k=0}^{\infty} a_k\left((x+h-a)^k-(x-a)^k\right) \\
&= \frac{1}{h}\sum_{k=1}^{\infty} a_k\left(\sum_{j=0}^{k}\binom{k}{j}(x-a)^j h^{k-j}-(x-a)^k\right) \\
&= \sum_{k=1}^{\infty} a_k\left(\sum_{j=0}^{k-1}\binom{k}{j}(x-a)^j h^{(k-1)-j}\right)
\end{aligned}$$

Then

$$\begin{aligned}
&\left|\frac{f(x+h)-f(x)}{h}-\sum_{k=1}^{\infty} a_k k(x-a)^{k-1}\right| \\
&= \left|\sum_{k=1}^{\infty} a_k\left(\sum_{j=0}^{k-1}\binom{k}{j}(x-a)^j h^{(k-1)-j}-k(x-a)^{k-1}\right)\right| \\
&= \left|\sum_{k=2}^{\infty} a_k\left(\sum_{j=0}^{k-1}\binom{k}{j}(x-a)^j h^{(k-1)-j}-k(x-a)^{k-1}\right)\right| \\
&= \left|\sum_{k=2}^{\infty} a_k\left(\sum_{j=0}^{k-2}\binom{k}{j}(x-a)^j h^{(k-1)-j}\right)\right|
\end{aligned}$$

Therefore,

$$\left|\frac{f(x+h)-f(x)}{h}-\sum_{k=1}^{\infty} a_k k(x-a)^{k-1}\right| \leq \sum_{k=2}^{\infty}|a_k|\left(\sum_{j=0}^{k-2}\binom{k}{j}|x-a|^j|h|^{(k-1)-j}\right)$$

Now it is clear that $k(k-1)\binom{k-2}{j} \geq \binom{k}{j}$ and so

$$\begin{aligned}
&= |h|\sum_{k=2}^{\infty}|a_k|\left(\sum_{j=0}^{k-2}\binom{k}{j}|x-a|^j|h|^{(k-2)-j}\right) \\
&\leq |h|\sum_{k=2}^{\infty}|a_k|k(k-1)\sum_{j=0}^{k-2}\binom{k-2}{j}|x-a|^j|h|^{(k-2)-j}
\end{aligned}$$

$$= |h|\sum_{k=2}^{\infty}|a_k|k(k-1)(|x-a|+|h|)^{k-2} < |h|\sum_{k=2}^{\infty}|a_k|k(k-1)r^{k-2} \tag{8.4}$$

By assumption and what was just observed about $\lim_{k\to\infty} k^{1/k}$,

$$\limsup_{k\to\infty}\left(|a_k|k(k-1)r^{k-2}\right)^{1/k} < 1$$

and so the series on the right in 8.4 converges. Therefore, assuming $|h|$ is small enough,

$$\left|\frac{f(x+h)-f(x)}{h}-\sum_{k=1}^{\infty} a_k k(x-a)^{k-1}\right| < C|h|$$

which shows that

$$\lim_{h\to 0}\left|\frac{f(x+h)-f(x)}{h}-\sum_{k=1}^{\infty}a_k k(x-a)^{k-1}\right|=0$$

Obviously, you can differentiate a power series infinitely often on the interior of the disk of convergence by simply repeating this theorem. ■

As an immediate corollary, it is possible to characterize the coefficients of a Taylor series.

Corollary 8.2.2 *Let $\sum_{n=0}^{\infty}a_n(x-a)^n$ be a Taylor series with radius of convergence $r>0$ and let*

$$f(x)\equiv\sum_{n=0}^{\infty}a_n(x-a)^n \text{ for } |x-a|<r \tag{8.5}$$

Then

$$a_n=\frac{f^{(n)}(a)}{n!}. \tag{8.6}$$

Also a Taylor series is infinitely differentiable on the interior of its disk of convergence.

Proof: From 8.5, $f(a)=a_0\equiv f^{(0)}(a)/0!$. From Theorem 8.2.1,

$$f'(x)=\sum_{n=1}^{\infty}a_n n(x-a)^{n-1}=a_1+\sum_{n=2}^{\infty}a_n n(x-a)^{n-1}.$$

Now let $x=a$ and obtain that $f'(a)=a_1=f'(a)/1!$. Next use Theorem 8.2.1 again to take the second derivative and obtain

$$f''(x)=2a_2+\sum_{n=3}^{\infty}a_n n(n-1)(x-a)^{n-2}$$

let $x=a$ in this equation and obtain $a_2=f''(a)/2=f''(a)/2!$. Continuing this way proves the corollary. ■

This also shows the coefficients of a Taylor series are unique. If

$$\sum_{k=0}^{\infty}a_k(x-a)^k=\sum_{k=0}^{\infty}b_k(x-a)^k$$

for all x in some open set containing a, then $a_k=b_k$ for all k.

It is possible to begin the study of complex analysis by defining the analytic functions to be those which are correctly given by a power series and this is sometimes done. For now, this will be the meaning of the word "analytic". The above theorem and corollary are the fundamental ideas in doing this.

Example 8.2.3 *Find the sum $\sum_{k=1}^{\infty}k2^{-k}$.*

It may not be obvious what this sum equals but with the above theorem it is easy to find. From the formula for the sum of a geometric series, $\frac{1}{1-t}=\sum_{k=0}^{\infty}t^k$ if $|t|<1$. Differentiate both sides to obtain $(1-t)^{-2}=\sum_{k=1}^{\infty}kt^{k-1}$ whenever $|t|<1$. Let $t=1/2$. Then $4=\frac{1}{(1-(1/2))^2}=\sum_{k=1}^{\infty}k2^{-(k-1)}$ and so if you multiply both sides by 2^{-1}, $2=\sum_{k=1}^{\infty}k2^{-k}$.

The above theorem shows that a power series is infinitely differentiable. Does it go the other way? That is, if the function has infinitely many continuous derivatives, is it correctly represented as a power series? The answer is no. See Problem 6 on Page 178 for an example. In fact, this is an important example and distinction. The modern theory of partial differential equations is built on just such functions which have many derivatives but no correct power series.

8.3 The Special Functions of Elementary Calculus

With this material on power series, it becomes possible to give an understandable treatment of the exponential function exp and the circular functions, sin and cos. A definition could be given directly for $x \in \mathbb{C}$ but in this section, it is assumed $x \in \mathbb{R}$.

8.3.1 Sines and Cosines

To begin with here is a definition of $\sin, \cos$, and exp.

Definition 8.3.1 *Define for all $x \in \mathbb{R}$*

$$\sin(x) \equiv \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \ \cos(x) \equiv \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \exp(x) \equiv \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Observation 8.3.2 *The above series converge for all $x \in \mathbb{F}$. This is most easily seen using the ratio test. Consider the series for $\sin(x)$ first. By the ratio test the series converges whenever*

$$\lim_{k\to\infty} \frac{\frac{|x|^{2k+3}}{(2k+3)!}}{\frac{|x|^{2k+1}}{(2k+1)!}} = \lim_{k\to\infty} \frac{1}{(2k+3)(2k+1)} |x|^2$$

is less than 1. However, this limit equals 0 for any x and so the series converges for all x. The verification of convergence for the other two series is left for you to do and is no harder. In what follows, I will emphasize x real. To do it for arbitrary $x \in \mathbb{C}$ is really a topic for complex analysis but you use the same series to define these functions.

Now that $\sin(x)$ and $\cos(x)$ have been defined, the properties of these functions must be considered. First, here is a fundamental lemma.

Lemma 8.3.3 *Suppose y is an $\mathbb{R}$ valued differentiable function and it solves the initial value problem,$y'' + y = 0,\ y(0) = 0, y'(0) = 0$. Then $y(x) = 0$.*

Proof: Multiply the equation by y' and use the chain rule to write

$$\frac{d}{dt}\left(\frac{1}{2}(y')^2 + \frac{1}{2}y^2\right) = 0.$$

Then by Corollary 7.8.5 $\frac{1}{2}(y')^2 + \frac{1}{2}y^2$ equals a constant. From the initial conditions, $y(0) = y'(0) = 0$, the constant can only be 0. ■

Theorem 8.3.4 $\sin'(x) = \cos(x)$ *and* $\cos'(x) = -\sin(x)$. *Also* $\cos(0) = 1, \sin(0) = 0$ *and*

$$\cos^2(x) + \sin^2(x) = 1 \tag{8.7}$$

for all x. *Also* $\sin(-x) = -\sin(x)$ *while* $\cos(-x) = \cos(x)$ *and the usual trig. identities hold,*

$$\sin(x+y) = \sin(x)\cos(y) + \sin(y)\cos(x) \tag{8.8}$$

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y) \tag{8.9}$$

Proof: That $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$ follows right away from differentiating the power series term by term using Theorem 8.2.1. It follows from the series that $\cos(0) = 1$ and $\sin(0) = 0$ and $\sin(-x) = -\sin(x)$ while $\cos(-x) = \cos(x)$ because the series for $\sin(x)$ only involves odd powers of x while the series for $\cos(x)$ only involves even powers.

For $x \in \mathbb{R}$, let $f(x) = \cos^2(x) + \sin^2(x)$, it follows from what was just discussed that $f(0) = 1$. Also from the chain rule, $f'(x) = 2\cos(x)(-\sin(x)) + 2\sin(x)\cos(x) = 0$ and so by Corollary 7.8.5, $f(x)$ is constant for all $x \in \mathbb{R}$. But $f(0) = 1$ so the constant can only be 1. Thus

$$\cos^2(x) + \sin^2(x) = 1 \tag{8.10}$$

as claimed.

It only remains to verify the identities. Consider 8.8 first. Fixing y and considering both sides as a function of x, it follows from the above that both sides of the identity satisfy the initial value problem

$$y'' + y = 0,\ y(0) = \sin(y), y'(0) = \cos(y)$$

Therefore, the difference satisfies the initial value problem of Lemma 8.3.3. Therefore, by this lemma, the difference equals 0. The next identity is handled similarly. ■

Note that 8.10 shows that the ordered pair $(\cos x, \sin x)$ lies on the unit circle with center at $(0,0)$.

Proposition 8.3.5 *The following important limits hold for* $a, b \neq 0$.

$$\lim_{x\to 0} \frac{\sin(ax)}{bx} = \frac{a}{b},\ \lim_{x\to 0} \frac{1-\cos(x)}{x} = 0.$$

Proof: From the definition of $\sin(x)$ given above, $\frac{\sin(ax)}{bx} =$

$$\frac{\sum_{k=0}^{\infty}(-1)^k \frac{(ax)^{2k+1}}{(2k+1)!}}{bx} = \frac{ax + \sum_{k=1}^{\infty}(-1)^k \frac{(ax)^{2k+1}}{(2k+1)!}}{bx} = \frac{a + \sum_{k=1}^{\infty}(-1)^k \frac{(ax)^{2k}}{(2k+1)!}}{b}$$

Now $\left|\sum_{k=1}^{\infty}(-1)^k \frac{(ax)^{2k}}{(2k+1)!}\right| \leq \sum_{k=1}^{\infty}|ax|^{2k} = \sum_{k=1}^{\infty}\left(|ax|^2\right)^k = \left(\frac{|ax|^2}{1-|ax|}\right)$ whenever $|ax| < 1$. Thus $\lim_{x\to 0}\sum_{k=1}^{\infty}(-1)^k \frac{(ax)^{2k}}{(2k+1)!} = 0$ and so $\lim_{x\to 0}\frac{\sin(ax)}{bx} = \frac{a}{b}$.The other limit can be handled similarly. ■

It is possible to verify the functions are periodic.

Lemma 8.3.6 *There exists a positive number* a, *such that* $\cos(a) = 0$.

Proof: To prove this, note that $\cos(0) = 1$ and so if it is false, it must be the case that $\cos(x) > 0$ for all positive x since otherwise, it would follow from the intermediate value theorem there would exist a point, x where $\cos x = 0$. Assume $\cos(x) > 0$ for all x. Then by Corollary 7.8.6 it would follow that $t \to \sin t$ is a strictly increasing function on $(0,\infty)$. Also note that $\sin(0) = 0$ and so $\sin(x) > 0$ for all $x > 0$. This is because, by the mean value theorem there exists $t \in (0,x)$ such that

$$\sin(x) = \sin(x) - \sin(0) = (\cos(t))(x-0) > 0.$$

By 8.7, $|f(x)| \leq 1$ for $f = \cos$ and $\sin$. Let $0 < x < y$. Then from the mean value theorem, $-\cos(y) - (-\cos(x)) = \sin(t)(y-x)$ for some $t \in (x,y)$. Since $t \to \sin(t)$ is increasing, it follows

$$-\cos(y) - (-\cos(x)) = \sin(t)(y-x) \geq \sin(x)(y-x).$$

This contradicts the inequality $|\cos(y)| \leq 1$ for all y because the right side is unbounded as $y \to \infty$. ■

Theorem 8.3.7 *Both* cos *and* sin *are periodic.*

Proof: Define a number π such that $\frac{\pi}{2} \equiv \inf\{x : x > 0 \text{ and } \cos(x) = 0\}$. Then $\frac{\pi}{2} > 0$ because $\cos(0) = 1$ and cos is continuous. On $\left[0, \frac{\pi}{2}\right]$ cos is positive and so it follows sin is increasing on this interval. Therefore, from 8.7, $\sin\left(\frac{\pi}{2}\right) = 1$. Now from Theorem 8.3.4,

$$\cos(\pi) = \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = -\sin^2\left(\frac{\pi}{2}\right) = -1, \ \sin(\pi) = 0$$

Using Theorem 8.3.4 again,

$$\cos(2\pi) = \cos^2(\pi) = 1 = \cos(0),$$

and so $\sin(2\pi) = 0$. From Theorem 8.3.4,

$$\cos(x + 2\pi) = \cos(x)\cos(2\pi) - \sin(x)\sin(2\pi) = \cos(x)$$

Thus cos is periodic of period 2π. By Theorem 8.3.4,

$$\sin(x + 2\pi) = \sin(x)\cos(2\pi) + \cos(x)\sin(2\pi) = \sin(x)$$

Using 8.7, it follows sin is also periodic of period 2π. ■

Note that 2π is the smallest period for these functions. This can be seen by observing that the above theorem and proof imply that cos is positive on $\left(0, \frac{\pi}{2}\right), \left(\frac{3\pi}{2}, 2\pi\right)$ and negative on $\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$ and that similar observations on sin are valid. Also, by considering where these functions are equal to 0, 1, and -1 along with where they are positive and negative, it follows that whenever $a^2 + b^2 = 1$, there exists a unique $t \in [0, 2\pi)$ such that $\cos(t) = a$ and $\sin(t) = b$. For example, if a and b are both positive, then since cos is continuous and strictly decreases from 1 to 0 on $\left[0, \frac{\pi}{2}\right]$, it follows there exists a unique $t \in (0, \pi/2)$ such that $\cos(t) = a$. Since $b > 0$ and sin is positive on $(0, \pi/2)$, it follows $\sin(t) = b$. No other value of t in $[0, 2\pi)$ will work since only on $(0, \pi/2)$ are both cos and sin positive. If $a > 0$ and $b < 0$ similar reasoning will show there exists a unique $t \in [0, 2\pi)$ with $\cos(t) = a$ and $\sin(t) = b$ and in this case, $t \in (3\pi/2, 2\pi)$. Other cases are similar and are left to the reader. Thus, every point on the unit circle is of the form $(\cos t, \sin t)$ for a unique $t \in [0, 2\pi)$.

This shows the unit circle is a smooth curve, however this notion will not be considered here.

Corollary 8.3.8 *For all* $x \in \mathbb{R}, \sin(x+2\pi) = \sin(x)$, $\cos(x+2\pi) = \cos(x)$

Proof: Let $y(x) \equiv \sin(x+2\pi) - \sin(x)$. Then from the above, $y'(0) = y(0) = 0$. It is also clear from the above that $y'' + y = 0$. Therefore, from Lemma 8.3.3 $y = 0$. Differentiating the identity just obtained yields the second identity. ■

Are these the same as the circular functions you studied very sloppily in calculus and trigonometry? They are.

If $\sin(x)$ defined above and $\sin(x)$ studied in a beginning calculus class both satisfy the initial value problem $y'' + y = 0$, $y(0) = 0, y'(0) = 1$ then they must be the same. However, if you remember anything from calculus you will realize $\sin(x)$ used there does satisfy the above initial value problem. If you don't remember anything from calculus, then it does not matter about harmonizing the functions. Just use the definition given above in terms of a power series. Similar considerations apply to $\cos$.

Of course all the other trig. functions are defined as earlier. Thus

$$\tan x = \frac{\sin x}{\cos x}, \cot x \equiv \frac{\cos x}{\sin x}, \sec x \equiv \frac{1}{\cos x}, \csc x \equiv \frac{1}{\sin x}.$$

Using the techniques of differentiation, you can find the derivatives of all these.

8.3.2 The Exponential Function

Now it is time to consider the exponential function $\exp(x)$ defined above. To do this, it is convenient to have the following uniqueness theorem.

Lemma 8.3.9 *Suppose*

$$y' - y = 0, \; y(0) = 0$$

Then $y = 0$. *Also for all* $x \in \mathbb{R}$, $\exp(-x)(\exp(x)) = 1$

Proof: The function exp has been defined above in terms of a power series. From this power series and Theorem 8.2.1 it follows that exp solves the above initial value problem. Thus $\exp' = \exp$. Multiply both sides of the differential equation by $\exp(-x)$. Then using the chain rule and product rule,

$$\frac{d}{dx}(\exp(-x)\, y(x)) = 0$$

and so $\exp(-x)\, y(x) = C$, a constant. The constant can only be 0 because of the initial condition. Therefore,

$$\exp(-x)\, y(x) = 0 \tag{$*$}$$

for all x.

Now I claim $\exp(-x)$ and $\exp(x)$ are never equal to 0. This is because by the chain rule, abusing notation slightly,

$$(\exp(-x)\exp(x))' = -\exp(-x)\exp(x) + \exp(-x)\exp(x) = 0$$

and so $\exp(-x)\exp(x) = C$ a constant. However, this constant can only be 1 because this is what it is when $x = 0$, a fact which follows right away from the definition in terms of power series. Thus from $*, y(x) = 0$. ■

Theorem 8.3.10 *The function* exp *satisfies the following properties.*

1. $\exp(x) > 0$ *for all* $x \in \mathbb{R}$, $\lim_{x\to\infty} \exp(x) = \infty, \lim_{x\to-\infty} \exp(x) = 0$.
2. exp *is the unique solution to the initial value problem*
$$y' - y = 0,\ y(0) = 1 \tag{8.11}$$
3. *For all* $x, y \in \mathbb{F}$
$$\exp(x+y) = \exp(x)\exp(y) \tag{8.12}$$
4. exp *is one to one mapping* $\mathbb{R}$ *onto* $(0,\infty)$.

Proof: To begin with consider 8.12. Fixing y it follows from the chain rule and the definition using power series that

$$x \to \exp(x+y) - \exp(x)\exp(y)$$

satisfies the initial value problem of Lemma 8.3.9 and so it is 0. This shows 8.12.

8.11 has already been noted. It comes directly from the definition and was proved in Lemma 8.3.9. The claim that $\exp(x) > 0$ was also established in the proof of this lemma.

Now from the power series, it is obvious that $\exp(x) > 0$ if $x > 0$ and by Lemma 8.3.9, $\exp(x)^{-1} = \exp(-x)$, so it follows $\exp(-x)$ is also positive. Since $\exp(x) > \sum_{k=0}^{2} \frac{x^k}{k!}$, it is clear $\lim_{x\to\infty} \exp(x) = \infty$ and it follows from this that $\lim_{x\to-\infty} \exp(x) = 0$.

It only remains to verify 4. Let $y \in (0,\infty)$. From the earlier properties, there exist x_1 such that $\exp(x_1) < y$ and x_2 such that $\exp(x_2) > y$. Then by the intermediate value theorem, there exists $x \in (x_1, x_2)$ such that $\exp(x) = y$. Thus exp maps onto $(0,\infty)$. It only remains to verify exp is one to one. Suppose then that $x_1 < x_2$. By the mean value theorem, there exists $x \in (x_1, x_2)$ such that

$$\exp(x)(x_2 - x_1) = \exp'(x)(x_2 - x_1) = \exp(x_2) - \exp(x_1).$$

Since $\exp(x) > 0$, it follows $\exp(x_2) \neq \exp(x_1)$. ■

8.4 ln and $\log_b$

In this section, the inverse function of $x \to \exp(x)$ is considered.

Definition 8.4.1 ln *is the inverse function of* exp. *It follows from the definition of inverse functions that* $\ln : (0,\infty) \to \mathbb{R}$ *,* $\ln(\exp(x)) = x$, *and* $\exp(\ln(x)) = x$. *The number* e *is that number such that* $\ln(e) = 1$.

By Corollary 7.10.2, it follows ln is differentiable. This makes possible the following simple theorem.

Theorem 8.4.2 *The following basic properties are available for* ln.

$$\ln'(x) = \frac{1}{x}. \tag{8.13}$$

Also for all $x, y > 0$,

$$\ln(xy) = \ln(x) + \ln(y), \tag{8.14}$$
$$\ln(1) = 0, \ln(x^m) = m\ln(x) \tag{8.15}$$

for all m *an integer.*

Proof: Since $\exp(\ln(x)) = x$ and $\ln'$ exists, it follows

$$\begin{aligned} x\ln'(x) &= \exp(\ln(x))\ln'(x) = \exp'(\ln(x))\ln'(x) \\ &= \exp(\ln(x))\ln'(x) = x\ln'(x) = 1 \end{aligned}$$

and this proves 8.13. Next consider 8.14.

$$xy = \exp(\ln(xy)),\ \exp(\ln(x) + \ln(y)) = \exp(\ln(x))\exp(\ln(y)) = xy.$$

Since exp was shown to be 1-1, it follows $\ln(xy) = \ln(x) + \ln(y)$. Next $\exp(0) = 1$ and $\exp(\ln(1)) = 1$ so $\ln(1) = 0$ again because exp is 1-1. Let $f(x) = \ln(x^m) - m\ln(x)$. Then $f(1) = \ln(1) - m\ln(1) = 0$. Also, by the chain rule, $f'(x) = \frac{1}{x^m}mx^{m-1} - m\frac{1}{x} = 0$ and so $f(x)$ equals a constant. The constant can only be 0 because $f(1) = 0$. This proves the last formula of 8.15 and completes the proof of the theorem. ■

The last formula tells how to define x^α for any $x > 0$ and $\alpha \in \mathbb{R}$. I want to stress that this is something new. Students are often deceived into thinking they know what x^α means for α a real number because they have a meaning for α an integer and with a little stretch for α a rational number. Such deception should never be tolerated in mathematics.

Definition 8.4.3 *Define x^α for $x > 0$ and $\alpha \in \mathbb{R}$ by $\ln(x^\alpha) = \alpha\ln(x)$. In other words, $x^\alpha \equiv \exp(\alpha\ln(x))$.*

From Theorem 8.4.2 this new definition does not contradict the usual definition in the case where α is an integer.

From this definition, the following properties are obtained.

Proposition 8.4.4 *For $x > 0$ let $f(x) = x^\alpha$ where $\alpha \in \mathbb{R}$. Then $f'(x) = \alpha x^{\alpha-1}$. Also $x^{\alpha+\beta} = x^\alpha x^\beta$ and $(x^\alpha)^\beta = x^{\alpha\beta}$.*

Proof: First consider the claim about the sum of the exponents.

$$\begin{aligned} x^{\alpha+\beta} &\equiv \exp((\alpha+\beta)\ln(x)) = \exp(\alpha\ln(x) + \beta\ln(x)) \\ &= \exp(\alpha\ln(x))\exp(\beta\ln(x)) \equiv x^\alpha x^\beta. \end{aligned}$$

$$\ln\left((x^\alpha)^\beta\right) = \beta\ln(x^\alpha) = \alpha\beta\ln(x),\ \ln\left(x^{\alpha\beta}\right) = \alpha\beta\ln(x).$$

The claim about the derivative follows from the chain rule. $f(x) = \exp(\alpha\ln(x))$ and so

$$f'(x) = \exp(\alpha\ln(x))\frac{\alpha}{x} \equiv \frac{\alpha}{x}x^\alpha = \alpha\left(x^{-1}\right)x^\alpha = \alpha x^{\alpha-1}. \blacksquare$$

Definition 8.4.5 *Define $\log_b$ for any $b > 0, b \neq 1$ by $\log_b(x) \equiv \frac{\ln(x)}{\ln(b)}$.*

Proposition 8.4.6 *The following hold for $\log_b(x)$.*

1. $b^{\log_b(x)} = x$, $\log_b(b^x) = x$.
2. $\log_b(xy) = \log_b(x) + \log_b(y)$
3. $\log_b(x^\alpha) = \alpha\log_b(x)$

Proof: $b^{\log_b(x)} \equiv \exp(\ln(b)\log_b(x)) = \exp\left(\ln(b)\frac{\ln(x)}{\ln(b)}\right) = \exp(\ln(x)) = x$ and also $\log_b(b^x) = \frac{\ln(b^x)}{\ln(b)} = \frac{x\ln(b)}{\ln(b)} = x$. This proves 1.

Now consider 2.

$$\log_b(xy) = \frac{\ln(xy)}{\ln(b)} = \frac{\ln(x)}{\ln(b)} + \frac{\ln(y)}{\ln(b)} = \log_b(x) + \log_b(y).$$

Finally, $\log_b(x^\alpha) = \frac{\ln(x^\alpha)}{\ln(b)} = \alpha\frac{\ln(x)}{\ln(b)} = \alpha\log_b(x)$. ■

8.5 The Complex Exponential

What does e^{ix} mean? Here $i^2 = -1$. Recall the complex numbers are of the form $a + ib$ and are identified as points in the plane. For $f(x) = e^{ix}$, you would want

$$f''(x) = i^2 f(x) = -f(x)$$

so $f''(x) + f(x) = 0$. Also, you would want $f(0) = e^0 = 1$, $f'(0) = ie^0 = i$. One solution to these conditions is $f(x) = \cos(x) + i\sin(x)$. Is it the only solution? Suppose $g(x)$ is another solution. Consider $u(x) = f(x) - g(x)$. Then it follows

$$u''(x) + u(x) = 0,\ u(0) = 0 = u'(0).$$

Thus both $\operatorname{Re} u$ and $\operatorname{Im} u$ solve the differential equation and 0 initial condition. By Lemma 8.3.3 both $\operatorname{Re} u$ and $\operatorname{Im} u$ are equal to 0. Thus the above is the only solution. Recall by De'Moivre's theorem

$$(\cos x + i\sin x)^n = \cos(nx) + i\sin(nx)$$

for any integer n and so $\left(e^{ix}\right)^n = e^{inx}$.

If you have a complex number $x + iy$, you can write it as

$$\sqrt{x^2+y^2}\left(\frac{x}{\sqrt{x^2+y^2}} + i\frac{y}{\sqrt{x^2+y^2}}\right)$$

and then note that $\left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right)$ is a point on the unit circle. Thus there is θ such that this ordered pair is $(\cos\theta, \sin\theta)$ and so if you let $r = \sqrt{x^2+y^2}$, the distance to the origin, the complex number can be written in the form $r(\cos\theta + i\sin\theta)$. From the above, this is of the form $re^{i\theta}$. This is called the polar form of a complex number. You should verify that with this convention, $re^{i\theta}\hat{r}e^{i\hat{\theta}} = r\hat{r}e^{i(\theta+\hat{\theta})}$. This reduces to using the trig identities for the cosine and sine of the sum of two angles.

In particular, this shows how to parametrize a circle in $\mathbb{C}$ centered at 0 which has radius r. It is just $\gamma(t) = re^{it}$ where $t \in [0, 2\pi]$. By this is meant that as t moves from 0 to 2π, the point $\gamma(t)$ is on the circle of radius r and moves in the counter clockwise direction over the circle.

8.6 The Binomial Theorem

The following is a very important example known as the binomial series. It was discovered by Newton.

Example 8.6.1 *Find a Taylor series for the function* $(1+x)^\alpha$ *centered at 0 valid for* $|x|<1$.

Use Theorem 8.2.1 to do this. First note that if $y(x) \equiv (1+x)^\alpha$, then y is a solution of the following initial value problem.

$$y' - \frac{\alpha}{(1+x)} y = 0, \ y(0) = 1. \tag{8.16}$$

Next it is necessary to observe there is only one solution to this initial value problem. To see this, multiply both sides of the differential equation in 8.16 by $(1+x)^{-\alpha}$. When this is done, one obtains

$$\frac{d}{dx}\left((1+x)^{-\alpha} y\right) = (1+x)^{-\alpha}\left(y' - \frac{\alpha}{(1+x)} y\right) = 0. \tag{8.17}$$

Therefore, from 8.17, there must exist a constant, C, such that $(1+x)^{-\alpha} y = C$. However, $y(0) = 1$ and so it must be that $C = 1$. Therefore, there is exactly one solution to the initial value problem in 8.16 and it is $y(x) = (1+x)^\alpha$.

The strategy for finding the Taylor series of this function consists of finding a series which solves the initial value problem above. Let $y(x) \equiv \sum_{n=0}^\infty a_n x^n$ be a solution to 8.16. Of course it is not known at this time whether such a series exists. However, the process of finding it will demonstrate its existence. From Theorem 8.2.1 and the initial value problem, $(1+x)\sum_{n=0}^\infty a_n n x^{n-1} - \sum_{n=0}^\infty \alpha a_n x^n = 0$ and so

$$\sum_{n=1}^\infty a_n n x^{n-1} + \sum_{n=0}^\infty a_n (n-\alpha) x^n = 0$$

Changing the variable of summation in the first sum,

$$\sum_{n=0}^\infty a_{n+1}(n+1)x^n + \sum_{n=0}^\infty a_n (n-\alpha) x^n = 0$$

and from Corollary 8.2.2 and the initial condition for 8.16 this requires

$$a_{n+1} = \frac{a_n(\alpha - n)}{n+1}, a_0 = 1. \tag{8.18}$$

Therefore, from 8.18 and letting $n = 0$, $a_1 = \alpha$, then using 8.18 again along with this information, $a_2 = \frac{\alpha(\alpha-1)}{2}$. Using the same process, $a_3 = \frac{\left(\frac{\alpha(\alpha-1)}{2}\right)(\alpha-2)}{3} = \frac{\alpha(\alpha-1)(\alpha-2)}{3!}$. By now you can spot the pattern. In general,

$$a_n = \frac{\overbrace{\alpha(\alpha-1)\cdots(\alpha-n+1)}^{n \text{ of these factors}}}{n!}.$$

Therefore, the candidate for the Taylor series is

$$y(x) = \sum_{n=0}^\infty \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n.$$

Furthermore, the above discussion shows this series solves the initial value problem on its interval of convergence. It only remains to show the radius of convergence of this series equals 1. It will then follow that this series equals $(1+x)^{\alpha}$ because of uniqueness of the initial value problem. To find the radius of convergence, use the ratio test. Thus the ratio of the absolute values of $(n+1)^{st}$ term to the absolute value of the n^{th} term is

$$\frac{\left|\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)(\alpha-n)}{(n+1)n!}\right||x|^{n+1}}{\left|\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}\right||x|^{n}} = |x|\,\frac{|\alpha-n|}{n+1} \to |x|$$

showing that the radius of convergence is 1 since the series converges if $|x|<1$ and diverges if $|x|>1$.

The expression, $\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$ is often denoted as $\binom{\alpha}{n}$. With this notation, the following theorem has been established.

Theorem 8.6.2 *Let α be a real number and let $|x|<1$. Then $(1+x)^{\alpha} = \sum_{n=0}^{\infty}\binom{\alpha}{n}x^n$.*

There is a very interesting issue related to the above theorem which illustrates the limitation of power series. The function $f(x) = (1+x)^{\alpha}$ makes sense for all $x>-1$ but one is only able to describe it with a power series on the interval $(-1,1)$. Think about this. The above technique is a standard one for obtaining solutions of differential equations and this example illustrates a deficiency in the method. An even more troubling example is $\frac{1}{1+x^2}$ which makes sense and is differentiable for all $x\in\mathbb{R}$ but its power series diverges for $|x|>1$. This is because the function fails to be differentiable at $x=i$ in the complex plane and so, if it had a valid power series, this could not happen thanks to the above theorems. However, we may only care about real values of x and if so, this reliance on power series is pretty useless.

To completely understand power series, it is necessary to take a course in complex analysis because this is where they make sense. It turns out that the right way to consider Taylor series is through the use of geometric series and something called the Cauchy integral formula of complex analysis. An introduction is given later.

8.7 Exercises

1. In each of the following, assume the relation defines y as a function of x for values of x and y of interest and find $y'(x)$.

   (a) $xy^2+\sin(y) = x^3+1$
   (b) $y^3+x\cos\left(y^2\right) = x^4$
   (c) $y\cos(x) = \tan(y)\cos\left(x^2\right)+2$
   (d) $\left(x^2+y^2\right)^6 = x^3y+3$
   (e) $\frac{xy^2+y}{y^5+x}+\cos(y) = 7$
   (f) $\sqrt{x^2+y^4}\sin(y) = 3x$
   (g) $y^3\sin(x)+y^2x^2 = 2^{x^2}y+\ln|y|$
   (h) $y^2\sin(y)x+\log_3(xy) = y^2+11$
   (i) $\sin\left(x^2+y^2\right)+\sec(xy) = e^{x+y}+y2^y+2$
   (j) $\sin\left(\tan\left(xy^2\right)\right)+y^3 = 16$
   (k) $\cos(\sec(\tan(y)))+\ln(5+\sin(xy)) = x^2y+3$

2. In each of the following, assume the relation defines y as a function of x for values of x and y of interest. Use the chain rule to show y satisfies the given differential equation.

(a) $x^2y+\sin y=7$, $\left(x^2+\cos y\right)y'+2xy=0$.

(b) $x^2y^3+\sin\left(y^2\right)=5$, $2xy^3+\left(3x^2y^2+2\left(\cos\left(y^2\right)\right)y\right)y'=0$.

(c) $y^2\sin(y)+xy=6$, $\left(2y(\sin(y))+y^2(\cos(y))+x\right)y'+y=0$.

3. Show that if $D(g)\subseteq U\subseteq D(f)$, and if f and g are both one to one, then $f\circ g$ is also one to one.

4. The number e is that number such that $\ln e=1$. Prove $e^x=\exp(x)$.

5. Find a formula for $\frac{dy}{dx}$ for $y=b^x$. Prove your formula.

6. Let $y=x^x$ for $x\in(0,\infty)$. Find $y'(x)$.

7. The logarithm test states the following. Suppose $a_k\neq 0$ for large k and that $p=\lim_{k\to\infty}\frac{\ln\left(\frac{1}{|a_k|}\right)}{\ln k}$ exists. If $p>1$, then $\sum_{k=1}^{\infty}a_k$ converges absolutely. If $p<1$, then the series, $\sum_{k=1}^{\infty}a_k$ does not converge absolutely. Prove this theorem.

8. Suppose $f(x+y)=f(x)+f(y)$ and f is continuous at 0. Find all solutions to this functional equation which are continuous at $x=0$. Now find all solutions which are bounded near 0. Next if you want an even more interesting version of this, find all solutions whose graphs are not dense in the plane. (A set S is dense in the plane if for every $(a,b)\in\mathbb{R}\times\mathbb{R}$ and $r>0$, there exists $(x,y)\in S$ such that $\sqrt{(x-a)^2+(y-b)^2}<r$. This is called the Cauchy equation.

9. Suppose $f(x+y)=f(x)f(y)$ and f is continuous and not identically zero. Find all solutions to this functional equation. **Hint:** First show the functional equation requires $f>0$.

10. Suppose $f(xy)=f(x)+f(y)$ for $x,y>0$. Suppose also f is continuous. Find all solutions to this functional equation.

11. Using the Cauchy condensation test, determine the convergence of $\sum_{k=2}^{\infty}\frac{1}{k\ln k}$. Now determine the convergence of $\sum_{k=2}^{\infty}\frac{1}{k(\ln k)^{1.001}}$.

12. Find the values of p for which the following series converges and the values of p for which it diverges.$\sum_{k=4}^{\infty}\frac{1}{\ln^p(\ln(k))\ln(k)k}$

13. For p a positive number, determine the convergence of $\sum_{n=2}^{\infty}\frac{\ln n}{n^p}$ for various values of p.

14. Determine whether the following series converge absolutely, conditionally, or not at all and give reasons for your answers.

(a) $\sum_{n=1}^{\infty}(-1)^n\frac{\ln\left(k^5\right)}{k}$

(b) $\sum_{n=1}^{\infty}(-1)^n\frac{\ln\left(k^5\right)}{k^{1.01}}$

(c) $\sum_{n=1}^{\infty}(-1)^n\frac{10^n}{(1.01)^n}$

(d) $\sum_{n=1}^{\infty}(-1)^n\sin\left(\frac{1}{n}\right)$

(e) $\sum_{n=1}^{\infty}(-1)^n\tan\left(\frac{1}{n^2}\right)$

(f) $\sum_{n=1}^{\infty}(-1)^n\cos\left(\frac{1}{n^2}\right)$

(g) $\sum_{n=1}^{\infty}(-1)^n\sin\left(\frac{\sqrt{n}}{n^2+1}\right)$

15. De Moivre's theorem says $[r(\cos t + i\sin t)]^n = r^n(\cos nt + i\sin nt)$ for n a positive integer. Prove this formula by induction. Does this formula continue to hold for all integers n, even negative integers? Explain.

16. Using De Moivre's theorem, show that if $z \in \mathbb{C}$ then z has n distinct n^{th} roots. **Hint:** Letting $z = x + iy$, $z = |z|\left(\frac{x}{|z|} + i\frac{y}{|z|}\right)$ and argue $\left(\frac{x}{|z|}, \frac{y}{|z|}\right)$ is a point on the unit circle. Hence $z = |z|(\cos(\theta) + i\sin(\theta))$. Then $w = |w|(\cos(\alpha) + i\sin(\alpha))$ is an n^{th} root if and only if $(|w|(\cos(\alpha) + i\sin(\alpha)))^n = z$. Show this happens exactly when $|w| = \sqrt[n]{|z|}$ and $\alpha = \frac{\theta + 2k\pi}{n}$ for $k = 0, 1, \cdots, n$.

17. Using De Moivre's theorem from Problem 15, derive a formula for $\sin(5x)$ and one for $\cos(5x)$.

18. Suppose $\sum_{n=0}^{\infty} a_n (x-c)^n$ is a power series with radius of convergence r. Show the series converge uniformly on any interval $[a,b]$ where $[a,b] \subseteq (c-r, c+r)$.

19. Find the disc of convergence of the series $\sum \frac{x^n}{n^p}$ for various values of p. **Hint:** Use Dirichlet's test.

20. Show $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ for all $x \in \mathbb{R}$ where e is the number such that $\ln e = 1$. Thus $e = \sum_{k=0}^{\infty} \frac{1}{k!}$. Show e is irrational. **Hint:** If $e = p/q$ for p, q positive integers, then argue $q!\left(\frac{p}{q} - \sum_{k=0}^{q} \frac{1}{k!}\right)$ is an integer. However, you can also show $q!\left(\sum_{k=0}^{\infty} \frac{1}{k!} - \sum_{k=0}^{q} \frac{1}{k!}\right) < 1$

21. Let $a \geq 1$. Show that for all $x > 0$, you have the inequality $ax > \ln(1 + x^a)$.

8.8 L'Hôpital's Rule

There is an interesting rule which is often useful for evaluating difficult limits. It is called L'Hôpital's [1] rule. The best versions of this rule are based on the Cauchy Mean value theorem, Theorem 7.8.2 on Page 146.

Theorem 8.8.1 *Let $[a,b] \subseteq [-\infty, \infty]$ and suppose f, g are functions which satisfy,*

$$\lim_{x \to b-} f(x) = \lim_{x \to b-} g(x) = 0, \tag{8.19}$$

and f' and g' exist on (a,b) with $g'(x) \neq 0$ on (a,b). Suppose also that

$$\lim_{x \to b-} \frac{f'(x)}{g'(x)} = L. \tag{8.20}$$

Then

$$\lim_{x \to b-} \frac{f(x)}{g(x)} = L. \tag{8.21}$$

[1] L'Hôpital published the first calculus book in 1696. This rule, named after him, appeared in this book. The rule was actually due to Bernoulli who had been L'Hôpital's teacher. L'Hôpital did not claim the rule as his own but Bernoulli accused him of plagarism. Nevertheless, this rule has become known as L'Hôpital's rule ever since. The version of the rule presented here is superior to what was discovered by Bernoulli and depends on the Cauchy mean value theorem which was found over 100 years after the time of L'Hôpital.

Proof: By the definition of limit and 8.20 there exists $c < b$ such that if $t > c$, then

$$\left|\frac{f'(t)}{g'(t)} - L\right| < \frac{\varepsilon}{2}.$$

Now pick x, y such that $c < x < y < b$. By the Cauchy mean value theorem, there exists $t \in (x, y)$ such that

$$g'(t)(f(x) - f(y)) = f'(t)(g(x) - g(y)).$$

Since $g'(s) \neq 0$ for all $s \in (a, b)$ it follows from the mean value theorem $g(x) - g(y) \neq 0$. Therefore,

$$\frac{f'(t)}{g'(t)} = \frac{f(x) - f(y)}{g(x) - g(y)}$$

and so, since $t > c$,

$$\left|\frac{f(x) - f(y)}{g(x) - g(y)} - L\right| < \frac{\varepsilon}{2}.$$

Now letting $y \to b-$,

$$\left|\frac{f(x)}{g(x)} - L\right| \leq \frac{\varepsilon}{2} < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows 8.21. ■

The following corollary is proved in the same way.

Corollary 8.8.2 *Let $[a, b] \subseteq [-\infty, \infty]$ and suppose f, g are functions which satisfy,*

$$\lim_{x \to a+} f(x) = \lim_{x \to a+} g(x) = 0, \tag{8.22}$$

and f' and g' exist on (a, b) with $g'(x) \neq 0$ on (a, b). Suppose also that

$$\lim_{x \to a+} \frac{f'(x)}{g'(x)} = L. \tag{8.23}$$

Then

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = L. \tag{8.24}$$

Here is a simple example which illustrates the use of this rule.

Example 8.8.3 *Find* $\lim_{x \to 0} \frac{5x + \sin 3x}{\tan 7x}$.

The conditions of L'Hôpital's rule are satisfied because the numerator and denominator both converge to 0 and the derivative of the denominator is nonzero for x close to 0. Therefore, if the limit of the quotient of the derivatives exists, it will equal the limit of the original function. Thus,

$$\lim_{x \to 0} \frac{5x + \sin 3x}{\tan 7x} = \lim_{x \to 0} \frac{5 + 3\cos 3x}{7\sec^2(7x)} = \frac{8}{7}.$$

Sometimes you have to use L'Hôpital's rule more than once.

Example 8.8.4 *Find* $\lim_{x \to 0} \frac{\sin x - x}{x^3}$.

Note that $\lim_{x\to 0}(\sin x - x) = 0$ and $\lim_{x\to 0} x^3 = 0$. Also, the derivative of the denominator is nonzero for x close to 0. Therefore, if $\lim_{x\to 0}\frac{\cos x-1}{3x^2}$ exists and equals L, it will follow from L'Hôpital's rule that the original limit exists and equals L. However, $\lim_{x\to 0}(\cos x - 1) = 0$ and $\lim_{x\to 0} 3x^2 = 0$ so L'Hôpital's rule can be applied again to consider $\lim_{x\to 0}\frac{-\sin x}{6x}$. From L'Hôpital's rule, if this limit exists and equals L, it will follow that $\lim_{x\to 0}\frac{\cos x-1}{3x^2} = L$ and consequently $\lim_{x\to 0}\frac{\sin x-x}{x^3} = L$. But from Proposition 8.3.5, $\lim_{x\to 0}\frac{-\sin x}{6x} = \frac{-1}{6}$. Therefore, by L'Hôpital's rule, $\lim_{x\to 0}\frac{\sin x-x}{x^3} = \frac{-1}{6}$.

Warning 8.8.5 ***Be sure to check the assumptions of L'Hôpital's rule before using it.***

Example 8.8.6 *Find* $\lim_{x\to 0+}\frac{\cos 2x}{x}$.

The numerator becomes close to 1 and the denominator gets close to 0. Therefore, the assumptions of L'Hôpital's rule do not hold and so it does not apply. In fact there is no limit unless you define the limit to equal $+\infty$. Now lets try to use the conclusion of L'Hôpital's rule even though the conditions for using this rule are not verified. Take the derivative of the numerator and the denominator which yields $\frac{-2\sin 2x}{1}$, an expression whose limit as $x \to 0+$ equals 0. This is a good illustration of the above warning.

Some people get the unfortunate idea that one can find limits by doing experiments with a calculator. If the limit is taken as x gets close to 0, these people think one can find the limit by evaluating the function at values of x which are closer and closer to 0. Theoretically, this should work although you have no way of knowing how small you need to take x to get a good estimate of the limit. In practice, the procedure may fail miserably.

Example 8.8.7 *Find* $\lim_{x\to 0}\frac{\ln\left|1+x^{10}\right|}{x^{10}}$.

This limit equals $\lim_{y\to 0}\frac{\ln|1+y|}{y} = \lim_{y\to 0}\frac{\left(\frac{1}{1+y}\right)}{1} = 1$ where L'Hôpital's rule has been used. This is an amusing example. You should plug .001 in to the function $\frac{\ln\left|1+x^{10}\right|}{x^{10}}$ and see what your calculator or computer gives you. If it is like mine, it will give 0 and will keep on returning the answer of 0 for smaller numbers than .001. This illustrates the folly of trying to compute limits through calculator or computer experiments. Indeed, you could say that a calculator is as useful for understanding limits as a bicycle is for swimming. Those who pretend otherwise are either guilty of ignorance or dishonesty.

There is another form of L'Hôpital's rule in which $\lim_{x\to b-} f(x) = \pm\infty$ and $\lim_{x\to b-} g(x) = \pm\infty$.

Theorem 8.8.8 *Let $[a,b] \subseteq [-\infty,\infty]$ and suppose f,g are functions which satisfy,*

$$\lim_{x\to b-} f(x) = \pm\infty \text{ and } \lim_{x\to b-} g(x) = \pm\infty, \tag{8.25}$$

and f' and g' exist on (a,b) with $g'(x) \neq 0$ on (a,b). Suppose also

$$\lim_{x\to b-}\frac{f'(x)}{g'(x)} = L. \tag{8.26}$$

Then

$$\lim_{x\to b-}\frac{f(x)}{g(x)} = L. \tag{8.27}$$

Proof: By the definition of limit and 8.26 there exists $c < b$ such that if $t > c$, then

$$\left|\frac{f'(t)}{g'(t)} - L\right| < \frac{\varepsilon}{2}.$$

Now pick x, y such that $c < x < y < b$. By the Cauchy mean value theorem, there exists $t \in (x, y)$ such that

$$g'(t)(f(x) - f(y)) = f'(t)(g(x) - g(y)).$$

Since $g'(s) \neq 0$ on (a, b), it follows from mean value theorem $g(x) - g(y) \neq 0$. Therefore,

$$\frac{f'(t)}{g'(t)} = \frac{f(x) - f(y)}{g(x) - g(y)}$$

and so, since $t > c$,

$$\left|\frac{f(x) - f(y)}{g(x) - g(y)} - L\right| < \frac{\varepsilon}{2}.$$

Now this implies

$$\left|\frac{f(y)}{g(y)} \frac{\left(\frac{f(x)}{f(y)} - 1\right)}{\left(\frac{g(x)}{g(y)} - 1\right)} - L\right| < \frac{\varepsilon}{2}$$

where for all y large enough, both $\frac{f(x)}{f(y)} - 1$ and $\frac{g(x)}{g(y)} - 1$ are not equal to zero. Continuing to rewrite the above inequality yields

$$\left|\frac{f(y)}{g(y)} - L\frac{\left(\frac{g(x)}{g(y)} - 1\right)}{\left(\frac{f(x)}{f(y)} - 1\right)}\right| < \frac{\varepsilon}{2}\left|\frac{\left(\frac{g(x)}{g(y)} - 1\right)}{\left(\frac{f(x)}{f(y)} - 1\right)}\right|.$$

Therefore, for y large enough,

$$\left|\frac{f(y)}{g(y)} - L\right| \leq \left|L - L\frac{\left(\frac{g(x)}{g(y)} - 1\right)}{\left(\frac{f(x)}{f(y)} - 1\right)}\right| + \frac{\varepsilon}{2}\left|\frac{\left(\frac{g(x)}{g(y)} - 1\right)}{\left(\frac{f(x)}{f(y)} - 1\right)}\right| < \varepsilon$$

due to the assumption 8.25 which implies

$$\lim_{y \to b-} \frac{\left(\frac{g(x)}{g(y)} - 1\right)}{\left(\frac{f(x)}{f(y)} - 1\right)} = 1.$$

Therefore, whenever y is large enough, $\left|\frac{f(y)}{g(y)} - L\right| < \varepsilon$ and this is what is meant by 8.27. ■

As before, there is no essential difference between the proof in the case where $x \to b-$ and the proof when $x \to a+$. This observation is stated as the next corollary.

Corollary 8.8.9 *Let $[a, b] \subseteq [-\infty, \infty]$ and suppose f, g are functions which satisfy,*

$$\lim_{x \to a+} f(x) = \pm\infty \text{ and } \lim_{x \to a+} g(x) = \pm\infty, \tag{8.28}$$

and f' and g' exist on (a,b) with $g'(x) \neq 0$ on (a,b). Suppose also that

$$\lim_{x\to a+} \frac{f'(x)}{g'(x)} = L. \tag{8.29}$$

Then

$$\lim_{x\to a+} \frac{f(x)}{g(x)} = L. \tag{8.30}$$

Theorems 8.8.1 8.8.8 and Corollaries 8.8.2 and 8.8.9 will be referred to as L'Hôpital's rule from now on. Theorem 8.8.1 and Corollary 8.8.2 involve the notion of indeterminate forms of the form $\frac{0}{0}$. Please do not think any meaning is being assigned to the nonsense expression $\frac{0}{0}$. It is just a symbol to help remember the sort of thing described by Theorem 8.8.1 and Corollary 8.8.2. Theorem 8.8.8 and Corollary 8.8.9 deal with indeterminate forms which are of the form $\frac{\pm\infty}{\infty}$. Again, this is just a symbol which is helpful in remembering the sort of thing being considered. There are other indeterminate forms which can be reduced to these forms just discussed. Don't ever try to assign meaning to such symbols.

Example 8.8.10 *Find* $\lim_{y\to\infty}\left(1+\frac{x}{y}\right)^y$.

It is good to first see why this is called an indeterminate form. One might think that as $y\to\infty$, it follows $x/y\to 0$ and so $1+\frac{x}{y}\to 1$. Now 1 raised to anything is 1 and so it would seem this limit should equal 1. On the other hand, if $x>0$, $1+\frac{x}{y}>1$ and a number raised to higher and higher powers should approach ∞. It really isn't clear what this limit should be. It is an indeterminate form which can be described as 1^∞. By definition,

$$\left(1+\frac{x}{y}\right)^y = \exp\left(y\ln\left(1+\frac{x}{y}\right)\right).$$

Now using L'Hôpital's rule,

$$\begin{aligned}\lim_{y\to\infty} y\ln\left(1+\frac{x}{y}\right) &= \lim_{y\to\infty}\frac{\ln\left(1+\frac{x}{y}\right)}{1/y} = \lim_{y\to\infty}\frac{\frac{1}{1+(x/y)}\left(-x/y^2\right)}{\left(-1/y^2\right)} \\ &= \lim_{y\to\infty}\frac{x}{1+(x/y)} = x\end{aligned}$$

Therefore, $\lim_{y\to\infty} y\ln\left(1+\frac{x}{y}\right)=x$. Since exp is continuous, it follows

$$\lim_{y\to\infty}\left(1+\frac{x}{y}\right)^y = \lim_{y\to\infty}\exp\left(y\ln\left(1+\frac{x}{y}\right)\right) = e^x.$$

8.8.1 Interest Compounded Continuously

Suppose you put money in the bank and it accrues interest at the rate of r per payment period. These terms need a little explanation. If the payment period is one month, and you started with \$100 then the amount at the end of one month would equal $100(1+r) = 100+100r$. In this the second term is the interest and the first is called the principal. Now you have $100(1+r)$ in the bank. This becomes the new principal. How much will you have at the end of the second month? By analogy to what was just done it would equal

$$100(1+r)+100(1+r)r = 100(1+r)^2.$$

In general, the amount you would have at the end of n months is $100(1+r)^n$.

When a bank says they offer 6% compounded monthly, this means r, the rate per payment period equals $.06/12$. Consider the problem of a rate of r per year and compounding the interest n times a year and letting n increase without bound. This is what is meant by compounding continuously. The interest rate per payment period is then r/n and the number of payment periods after time t years is approximately tn. From the above, the amount in the account after t years is

$$P\left(1+\frac{r}{n}\right)^{nt} \tag{8.31}$$

Recall from Example 8.8.10 that $\lim_{y\to\infty}\left(1+\frac{x}{y}\right)^y = e^x$. The expression in 8.31 can be written as

$$P\left[\left(1+\frac{r}{n}\right)^n\right]^t$$

and so, taking the limit as $n\to\infty$, you get $Pe^{rt}=A$. This shows how to compound interest continuously.

Example 8.8.11 *Suppose you have \$100 and you put it in a savings account which pays 6% compounded continuously. How much will you have at the end of* 4 *years?*

From the above discussion, this would be $100e^{(.06)4}=127.12$. Thus, in 4 years, you would gain interest of about \$27.

8.9 Exercises

1. Find the limits.

   (a) $\lim_{x\to 0}\frac{3x-4\sin 3x}{\tan 3x}$

   (b) $\lim_{x\to\frac{\pi}{2}-}(\tan x)^{x-(\pi/2)}$

   (c) $\lim_{x\to 1}\frac{\arctan(4x-4)}{\arcsin(4x-4)}$

   (d) $\lim_{x\to 0}\frac{\arctan 3x-3x}{x^3}$

   (e) $\lim_{x\to 0+}\frac{9^{\sec x-1}-1}{3^{\sec x-1}-1}$

   (f) $\lim_{x\to 0}\frac{3x+\sin 4x}{\tan 2x}$

   (g) $\lim_{x\to\pi/2}\frac{\ln(\sin x)}{x-(\pi/2)}$

   (h) $\lim_{x\to 0}\frac{\cosh 2x-1}{x^2}$

   (i) $\lim_{x\to 0}\frac{-\arctan x+x}{x^3}$

   (j) $\lim_{x\to 0}\frac{x^8\sin\frac{1}{x}}{\sin 3x}$

   (k) $\lim_{x\to\infty}(1+5^x)^{\frac{2}{x}}$

   (l) $\lim_{x\to 0}\frac{-2x+3\sin x}{x}$

   (m) $\lim_{x\to 1}\frac{\ln(\cos(x-1))}{(x-1)^2}$

   (n) $\lim_{x\to 0+}\sin^{\frac{1}{x}}x$

   (o) $\lim_{x\to 0}(\csc 5x-\cot 5x)$

   (p) $\lim_{x\to 0+}\frac{3^{\sin x}-1}{2^{\sin x}-1}$

   (q) $\lim_{x\to 0+}(4x)^{x^2}$

   (r) $\lim_{x\to\infty}\frac{x^{10}}{(1.01)^x}$

   (s) $\lim_{x\to 0}(\cos 4x)^{(1/x^2)}$

2. Find the following limits.

   (a) $\lim_{x\to 0+}\frac{1-\sqrt{\cos 2x}}{\sin^4(4\sqrt{x})}$.

   (b) $\lim_{x\to 0}\frac{2^{x^2}-2^{5x}}{\sin\left(\frac{x^2}{5}\right)-\sin(3x)}$.

   (c) $\lim_{n\to\infty}n\left(\sqrt[n]{7}-1\right)$.

   (d) $\lim_{x\to\infty}\left(\frac{3x+2}{5x-9}\right)^{x^2}$.

(e) $\lim_{x\to\infty}\left(\frac{3x+2}{5x-9}\right)^{1/x}$.

(f) $\lim_{n\to\infty}\left(\cos\frac{2x}{\sqrt{n}}\right)^n$.

(g) $\lim_{n\to\infty}\left(\cos\frac{2x}{\sqrt{5n}}\right)^n$.

(h) $\lim_{x\to3}\frac{x^x-27}{x-3}$.

(i) $\lim_{n\to\infty}\cos\left(\pi\frac{\sqrt{4n^2+13n}}{n}\right)$.

(j) $\lim_{x\to\infty}\left(\begin{array}{c}\sqrt[3]{x^3+7x^2}\\ -\sqrt{x^2-11x}\end{array}\right)$.

(k) $\lim_{x\to\infty}\left(\begin{array}{c}\sqrt[5]{x^5+7x^4}\\ -\sqrt[3]{x^3-11x^2}\end{array}\right)$.

(l) $\lim_{x\to\infty}\left(\frac{5x^2+7}{2x^2-11}\right)^{\frac{x}{1-x}}$.

(m) $\lim_{x\to\infty}\left(\frac{5x^2+7}{2x^2-11}\right)^{\frac{x\ln x}{1-x}}$.

(n) $\lim_{x\to0+}\frac{\ln\left(e^{2x^2}+7\sqrt{x}\right)}{\sinh(\sqrt{x})}$.

(o) $\lim_{x\to0+}\frac{\sqrt[7]{x}-\sqrt[5]{x}}{\sqrt[9]{x}-\sqrt[11]{x}}$.

3. Find the following limits.

(a) $\lim_{x\to0+}(1+3x)^{\cot 2x}$

(b) $\lim_{x\to0}\frac{\sin x-x}{x^2}=0$

(c) $\lim_{x\to0}\frac{\sin x-x}{x^3}$

(d) $\lim_{x\to0}\frac{\tan(\sin x)-\sin(\tan x)}{x^7}$

(e) $\lim_{x\to0}\frac{\tan(\sin 2x)-\sin(\tan 2x)}{x^7}$

(f) $\lim_{x\to0}\frac{\sin\left(x^2\right)-\sin^2(x)}{x^4}$

(g) $\lim_{x\to0}\frac{e^{-\left(1/x^2\right)}}{x}$

(h) $\lim_{x\to0}\left(\frac{1}{x}-\cot(x)\right)$

(i) $\lim_{x\to0}\frac{\cos(\sin x)-1}{x^2}$

(j) $\lim_{x\to\infty}\left(x^2\left(4x^4+7\right)^{1/2}-2x^4\right)$

(k) $\lim_{x\to0}\frac{\cos(x)-\cos(4x)}{\tan\left(x^2\right)}$

(l) $\lim_{x\to0}\frac{\arctan(3x)}{x}$

(m) $\lim_{x\to\infty}\left[\left(x^9+5x^6\right)^{1/3}-x^3\right]$

(n) $\lim_{x\to0}\frac{\ln(\sin(x)/x)}{x^2}$

4. Suppose you want to have \$2000 saved at the end of 5 years. How much money should you place into an account which pays 7% per year compounded continuously?

5. Using a good calculator, find $e^{.06}-\left(1+\frac{.06}{360}\right)^{360}$. Explain why this gives a measure of the difference between compounding continuously and compounding daily.

6. Consider the following function [2]

$$f(x)=\begin{cases} e^{-1/x^2} \text{ for } x\neq 0\\ 0 \text{ for } x=0\end{cases}$$

Show that $f^{(k)}(0)=0$ for all k so the power series of this function is of the form $\sum_{k=0}^{\infty}0x^k$ but the function is not identically equal to 0 on any interval containing 0. Thus this function has all derivatives at 0 and at every other point, yet fails to be correctly represented by its power series. This is an example of a smooth function which is not analytic. It is smooth because all derivatives exist and are continuous.

[2] Surprisingly, this function is very important to those who use modern techniques to study differential equations. One needs to consider test functions which have the property they have infinitely many derivatives but vanish outside of some interval. The theory of complex variables can be used to show there are no examples of such functions if they have a valid power series expansion. It even becomes a little questionable whether such strange functions even exist at all. Nevertheless, they do, there are enough of them, and it is this very example which is used to show this.

It fails to be analytic because it is not correctly given by its power series in any open set.

7. Euler seems to have done the following to find e^{ix}. He knew the power series for e^x and so substituted ix for x and this gave

$$\begin{aligned} & 1+ix+\left(i^2x^2\right)/2!+(ix)^3/3!+\cdots \\ = \; & 1+ix-x^2/2!-ix^3/3!+\cdots \end{aligned}$$

then he grouped the terms and got

$$1-\frac{x^2}{2!}+\frac{x^4}{4!}+\cdots+i\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}+\cdots\right)$$

and then recognized this as $\cos x+i\sin x$. What is wrong with this kind of thing? Why is it that in this case there is absolutely no problem and this is a legitimate explanation of Euler's formula.

8. Find $\lim_{x\to+\infty}\frac{x}{x+\sin(3x)}$. **Hint:** It might be good to not use L'Hospital's rule.

8.10 Multiplication of Power Series

Next consider the problem of multiplying two power series.

Theorem 8.10.1 *Let $\sum_{n=0}^{\infty}a_n(x-a)^n$, $\sum_{n=0}^{\infty}b_n(x-a)^n$ be two power series which have radius of convergence r_1 and r_2, both positive. Then*

$$\left(\sum_{n=0}^{\infty}a_n(x-a)^n\right)\left(\sum_{n=0}^{\infty}b_n(x-a)^n\right)=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a_kb_{n-k}\right)(x-a)^n$$

whenever $|x-a|<r\equiv\min(r_1,r_2)$.

Proof: By Theorem 8.1.3 both series converge absolutely if $|x-a|<r$. Therefore, by Theorem 5.5.6

$$\left(\sum_{n=0}^{\infty}a_n(x-a)^n\right)\left(\sum_{n=0}^{\infty}b_n(x-a)^n\right)=$$

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n}a_k(x-a)^kb_{n-k}(x-a)^{n-k}=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a_kb_{n-k}\right)(x-a)^n . \blacksquare$$

The significance of this theorem in terms of applications is that it states you can multiply power series just as you would multiply polynomials and everything will be all right on the common interval of convergence. This is called the Cauchy product.

This theorem can be used to find Taylor series which would perhaps be hard to find without it. Here is an example.

Example 8.10.2 *Find the Taylor series for $e^x\sin x$ centered at $x=0$.*

All that is required is to multiply

$$\left(\overbrace{1+x+\frac{x^2}{2!}+\frac{x^3}{3!}\cdots}^{e^x}\right)\left(\overbrace{x-\frac{x^3}{3!}+\frac{x^5}{5!}+\cdots}^{\sin x}\right)$$

From the above theorem the result should be

$$x+x^2+\left(-\frac{1}{3!}+\frac{1}{2!}\right)x^3+\cdots$$

$$=x+x^2+\frac{1}{3}x^3+\cdots$$

You can continue this way and get the following to a few more terms.

$$x+x^2+\frac{1}{3}x^3-\frac{1}{30}x^5-\frac{1}{90}x^6-\frac{1}{630}x^7+\cdots$$

I don't see a pattern in these coefficients but I can go on generating them as long as I want. (In practice this tends to not be very long.) I also know the resulting power series will converge for all x because both the series for e^x and the one for $\sin x$ converge for all x.

Example 8.10.3 *Find the Taylor series for* $\tan x$ *centered at* $x=0$.

Lets suppose it has a Taylor series $a_0+a_1x+a_2x^2+\cdots$. Then

$$\left(a_0+a_1x+a_2x^2+\cdots\right)\left(\overbrace{1-\frac{x^2}{2}+\frac{x^4}{4!}+\cdots}^{\cos x}\right)=\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}+\cdots\right).$$

Using the above, $a_0=0, a_1x=x$ so

$$a_1=1,\left(0\left(\frac{-1}{2}\right)+a_2\right)x^2=0$$

so $a_2=0$.

$$\left(a_3-\frac{a_1}{2}\right)x^3=\frac{-1}{3!}x^3$$

so $a_3-\frac{1}{2}=-\frac{1}{6}$ so $a_3=\frac{1}{3}$. Clearly one can continue in this manner. Thus the first several terms of the power series for tan are

$$\tan x=x+\frac{1}{3}x^3+\cdots.$$

You can go on calculating these terms and find the next two yielding

$$\tan x=x+\frac{1}{3}x^3+\frac{2}{15}x^5+\frac{17}{315}x^7+\cdots$$

This is a very significant technique because, as you see, there does not appear to be a very simple pattern for the coefficients of the power series for $\tan x$. Of course there are some issues here about whether $\tan x$ even has a power series, but if it does, the above must be it. In fact, $\tan(x)$ will have a power series valid on some interval centered at 0 and this becomes completely obvious when one uses methods from complex analysis but it isn't too obvious at this point. If you are interested in this issue, read the last section of the chapter. Note also that what has been accomplished is to divide the power series for $\sin x$ by the power series for $\cos x$ just like they were polynomials.

8.11 Exercises

1. Find the radius of convergence of the following.
   (a) $\sum_{k=1}^{\infty}\left(\frac{x}{2}\right)^n$
   (b) $\sum_{k=1}^{\infty}\sin\left(\frac{1}{n}\right)3^n x^n$
   (c) $\sum_{k=0}^{\infty}k!x^k$
   (d) $\sum_{n=0}^{\infty}\frac{(3n)^n}{(3n)!}x^n$
   (e) $\sum_{n=0}^{\infty}\frac{(2n)^n}{(2n)!}x^n$
2. Find $\sum_{k=1}^{\infty}k2^{-k}$.
3. Find $\sum_{k=1}^{\infty}k^2 3^{-k}$.
4. Find $\sum_{k=1}^{\infty}\frac{2^{-k}}{k}$.
5. Find $\sum_{k=1}^{\infty}\frac{3^{-k}}{k}$.
6. Show there exists a function f which is continuous on $[0,1]$ but nowhere differentiable and an infinite series of the form $\sum_{k=1}^{\infty}p_k(x)$ where each p_k is a polynomial which converges uniformly to $f(x)$ on $[0,1]$. Thus it makes absolutely no sense to write something like $f'(x)=\sum_{k=1}^{\infty}p_k'(x)$. **Hint:** Use the Weierstrass approximation theorem.
7. Find the power series centered at 0 for the function $1/\left(1+x^2\right)$ and give the radius of convergence. Where does the function make sense? Where does the power series equal the function?
8. Find a power series for the function $f(x)\equiv\frac{\sin(\sqrt{x})}{\sqrt{x}}$ for $x>0$. Where does $f(x)$ make sense? Where does the power series you found converge?
9. Use the power series technique which was applied in Example 8.6.1 to consider the initial value problem $y'=y, y(0)=1$. This yields another way to obtain the power series for e^x.
10. Use the power series technique on the initial value problem $y'+y=0,\ y(0)=1$. What is the solution to this initial value problem?
11. Use the power series technique to find solutions in terms of power series to the initial value problem
$$y''+xy=0,\ y(0)=0, y'(0)=1.$$
Tell where your solution gives a valid description of a solution for the initial value problem. **Hint:** This is a little different but you proceed the same way as in Example 8.6.1. The main difference is you have to do two differentiations of the power series instead of one.
12. Find several terms of a power series solution to the nonlinear initial value problem
$$y''+a\sin(y)=0,\ y(0)=1, y'(0)=0.$$

This is the equation which governs the vibration of a pendulum. Explain why there exists a power series which gives the solution to the above initial value problem. Multiply the equation by y' and identify what you have obtained as the derivative of an interesting quantity which must be constant.

13. Suppose the function e^x is defined in terms of a power series, $e^x \equiv \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Use Theorem 5.5.6 on Page 98 to show directly the usual law of exponents,
$$e^{x+y} = e^x e^y.$$
Be sure to check all the hypotheses.

14. Let $f_n(x) \equiv \left(\frac{1}{n} + x^2\right)^{1/2}$. Show that for all x,
$$||x| - f_n(x)| \leq \frac{1}{\sqrt{n}}.$$
Thus these approximate functions converge uniformly to the function $f(x) = |x|$. Now show $f_n'(0) = 0$ for all n and so $f_n'(0) \to 0$. However, the function $f(x) \equiv |x|$ has no derivative at $x = 0$. Thus even though $f_n(x) \to f(x)$ for all x, you cannot say that $f_n'(0) \to f'(0)$.

15. Let the functions, $f_n(x)$ be given in Problem 14 and consider
$$g_1(x) = f_1(x),\ g_n(x) = f_n(x) - f_{n-1}(x) \text{ if } n > 1.$$
Show that for all x,
$$\sum_{k=0}^{\infty} g_k(x) = |x|$$
and that $g_k'(0) = 0$ for all k. Therefore, you can't differentiate the series term by term and get the right answer[3].

16. Use the theorem about the binomial series to give a proof of the binomial theorem
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$
whenever n is a positive integer.

17. Find the power series for $\sin\left(x^2\right)$ by plugging in x^2 where ever there is an x in the power series for $\sin x$. How do you know this is the power series for $\sin\left(x^2\right)$?

18. Find the first several terms of the power series for $\sin^2(x)$ by multiplying the power series for $\sin(x)$. Next use the trig. identity, $\sin^2(x) = \frac{1-\cos(2x)}{2}$ and the power series for $\cos(2x)$ to find the power series.

19. Find the power series for $f(x) = \frac{1}{\sqrt{1-x^2}}$.

[3]How bad can this get? It can be much worse than this. In fact, there are functions which are continuous everywhere and differentiable nowhere. We typically don't have names for them but they are there just the same. Every such function can be written as an infinite sum of polynomials which of course have derivatives at every point. Thus it is nonsense to differentiate an infinite sum term by term without a theorem of some sort.

20. Let a,b be two positive numbers and let $p>1$. Choose q such that
$$\frac{1}{p}+\frac{1}{q}=1.$$
Now verify the important inequality
$$ab\leq\frac{a^p}{p}+\frac{b^q}{q}.$$
Hint: You might try considering $f(a)=\frac{a^p}{p}+\frac{b^q}{q}-ab$ for fixed $b>0$ and examine its graph using the derivative.

21. Using Problem 20, show that if $\alpha>0,p>1$, it follows that for all $x>0$
$$\left(\frac{p-1}{p}x+\frac{\alpha}{p}x^{1-p}\right)^p\geq\alpha.$$

22. Using Problem 21, define for $p>1$ and $\alpha>0$ the following sequence
$$x_{n+1}\equiv\frac{p-1}{p}x_n+\frac{\alpha}{p}x_n^{1-p},\ x_1>0.$$
Show $\lim_{n\to\infty}x_n=x$ where $x=\alpha^{1/p}$. In fact show that after x_1 the sequence decreases to $\alpha^{1/p}$.

23. Consider the sequence $\left\{\left(1+\frac{x}{n}\right)^n\right\}_{n=1}^{\infty}$ where x is a positive number. Using the binomial theorem show this sequence is increasing. Next show the sequence converges.

24. Consider the sequence $\left\{\left(1+\frac{x}{n}\right)^{n+1}\right\}_{n=1}^{\infty}$ where x is a positive number. Show this sequence decreases when $x>2$. **Hint:** You might consider showing $(1+y)^{(x/y)+1}$ is increasing in y provided $x>2$. To do this, you might use the following observation repeatedly. If $f(0)=0$ and $f'(y)>0$, then $f(y)\geq 0$. There may also be other ways to do this.

25. Let $\frac{z}{e^z-1}=\sum_{n=0}^{\infty}\frac{b_n}{n!}z^n$. The b_n are called the Bernoulli numbers. Show that
$$\frac{b_0}{0!n!}+\frac{b_1}{1!(n-1)!}+\frac{b_2}{2!(n-2)!}+\cdots+\frac{b_{n-1}}{(n-1)!1!}=\begin{cases}1 \text{ if } n=1\\ 0 \text{ id } n>1\end{cases}$$
Hint: You might use the Cauchy product after multiplying both sides by e^z-1.

8.12 The Fundamental Theorem of Algebra

The fundamental theorem of algebra states that every non constant polynomial having coefficients in $\mathbb{C}$ has a zero in $\mathbb{C}$. If $\mathbb{C}$ is replaced by $\mathbb{R}$, this is not true because of the example, $x^2+1=0$. This theorem is a very remarkable result and notwithstanding its title, all the best proofs of it depend on either analysis or topology. It was proved by Gauss in 1797. The proof given here follows Rudin [24]. See also Hardy [14] for a similar proof, more discussion and references. You can also see the interesting article on Wikipedia. You google fundamental theorem of algebra and go to this site. There are many ways to prove it. This article claims the first completely correct proof was done by Argand in 1806. The shortest proof is found in the theory of complex analysis and is a simple application of Liouville's theorem or the formula for counting zeros.

Recall De Moivre's theorem, Problem 15 on Page 172 from trigonometry which is listed here for convenience.

Theorem 8.12.1 *Let $r > 0$ be given. Then if n is a positive integer,*

$$[r(\cos t + i\sin t)]^n = r^n(\cos nt + i\sin nt).$$

Recall that this theorem is the basis for proving the following corollary from trigonometry, also listed here for convenience, see Problem 16 on Page 172.

Corollary 8.12.2 *Let z be a non zero complex number and let k be a positive integer. Then there are always exactly k k^{th} roots of z in $\mathbb{C}$.*

Lemma 8.12.3 *Let $a_k \in \mathbb{C}$ for $k = 1, \cdots, n$ and let $p(z) \equiv \sum_{k=1}^{n} a_k z^k$. Then p is continuous.*

Proof:
$$|az^n - aw^n| \le |a|\,|z-w|\,\left|z^{n-1} + z^{n-2}w + \cdots + w^{n-1}\right|.$$
Then for $|z-w| < 1$, the triangle inequality implies $|w| < 1 + |z|$ and so if $|z-w| < 1$,

$$|az^n - aw^n| \le |a|\,|z-w|\,n(1+|z|)^n.$$

If $\varepsilon > 0$ is given, let
$$\delta < \min\left(1, \frac{\varepsilon}{|a|\,n(1+|z|)^n}\right).$$
It follows from the above inequality that for $|z-w| < \delta$, $|az^n - aw^n| < \varepsilon$. The function of the lemma is just the sum of functions of this sort and so it follows that it is also continuous. In particular, if $z_k \to z$, then $p(z_k) \to p(z)$. ■

Here are some observations about compactness. Recall Theorem 4.8.14 which says that closed and bounded sets in $\mathbb{C}$ are sequentially compact. A version of this is the following theorem called the Weierstrass Bolzano theorem.

Theorem 8.12.4 *Let $\{z_k\}$ be a sequence of complex numbers such that $|z_k|$ is a bounded sequence of real numbers. Then there exists a subsequence $\{z_{n_k}\}$ and a complex number z such that $\lim_{k\to\infty} z_{n_k} = z$. Sets of the form $K \equiv \{z \in \mathbb{C} : |z| \le r\}$ are sequentially compact.*

Proof: The existence of z follows from the observation that

$$\{z_k\} \subseteq \{z \in \mathbb{C} \text{ such that } |z| \le r\}$$

for some $r > 0$. This disk just described is sequentially compact by Theorem 4.8.14. ■

Theorem 8.12.5 *Let $p(z)$ be a polynomial of degree $n \ge 1$ having complex coefficients. Then there exists z_0 such that $p(z_0) = 0$, a zero of the polynomial.*

Proof: Suppose the nonconstant polynomial

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n, a_n \neq 0,$$

has no zero in $\mathbb{C}$. By the triangle inequality,

$$|p(z)| \geq |a_n| |z|^n - \left| a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} \right|$$

$$\geq |a_n| |z|^n - \left(|a_0| + |a_1| |z| + \cdots + |a_{n-1}| |z|^{n-1} \right)$$

Now the term $|a_n| |z|^n$ dominates all the other terms which have $|z|$ raised to a lower power and so $\lim_{|z| \to \infty} |p(z)| = \infty$. Now let

$$0 \leq \lambda \equiv \inf \{|p(z)| : z \in \mathbb{C}\}$$

Then since $\lim_{|z| \to \infty} |p(z)| = \infty$, it follows that there exists $r > 0$ such that if $|z| > r$, then $|p(z)| \geq 1 + \lambda$. It follows that

$$\lambda = \inf \{|p(z)| : |z| \leq r\}$$

Since $K \equiv \{z : |z| \leq r\}$ is sequentially compact, it follows that, letting $\{z_k\} \subseteq K$ with $|p(z_k)| \leq \lambda + 1/k$, there is a subsequence still denoted as $\{z_k\}$ such that $\lim_{k \to \infty} z_k = z_0 \in K$. Then $|p(z_0)| = \lambda$ and so $\lambda > 0$. Thus,

$$|p(z_0)| = \min_{z \in K} |p(z)| = \min_{z \in \mathbb{C}} |p(z)| > 0$$

Then let $q(z) = \frac{p(z+z_0)}{p(z_0)}$. This is also a polynomial which has no zeros and the minimum of $|q(z)|$ is 1 and occurs at $z = 0$. Since $q(0) = 1$, it follows $q(z) = 1 + a_k z^k + r(z)$ where $r(z)$ consists of higher order terms. Here a_k is the first coefficient of $q(z)$ which is nonzero. Choose a sequence, $z_n \to 0$, such that $a_k z_n^k < 0$. For example, let $-a_k z_n^k = (1/n)$. Then for $r(z) = a_m z^m + a_{m+1} z^{m+1} + \ldots + a_n z^n$ for $m > k$,

$$\begin{aligned} |q(z_n)| &= \left| 1 + a_k z^k + r(z) \right| \leq 1 - 1/n + |r(z_n)| \\ &\leq 1 - \frac{1}{n} + \frac{1}{n} \sum_{j=m}^{n} |a_j| |a_k|^{1/k} \left(\frac{1}{n} \right)^{(j-k)/k} < 1 \end{aligned}$$

for all n large enough because the sum is smaller than 1 for n large enough. This contradicts $|q(z)| \geq 1$. ■

8.13 Some Other Theorems

First recall Theorem 5.5.6 on Page 98. For convenience, the version of this theorem which is of interest here is listed below.

Theorem 8.13.1 *Suppose $\sum_{i=0}^{\infty} a_i$ and $\sum_{j=0}^{\infty} b_j$ both converge absolutely. Then*

$$\left(\sum_{i=0}^{\infty} a_i \right) \left(\sum_{j=0}^{\infty} b_j \right) = \sum_{n=0}^{\infty} c_n$$

where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Furthermore, $\sum_{n=0}^{\infty} c_n$ converges absolutely.

Proof: It only remains to verify the last series converges absolutely. Letting p_{nk} equal 1 if $k \leq n$ and 0 if $k > n$. Then by Theorem 5.5.3 on Page 97

$$\begin{aligned}
\sum_{n=0}^{\infty} |c_n| &= \sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} a_k b_{n-k} \right| \leq \sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_k| |b_{n-k}| = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} p_{nk} |a_k| |b_{n-k}| \\
&= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} p_{nk} |a_k| |b_{n-k}| = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} |a_k| |b_{n-k}| = \sum_{k=0}^{\infty} |a_k| \sum_{n=0}^{\infty} |b_n| < \infty. \blacksquare
\end{aligned}$$

The above theorem is about multiplying two series. What if you wanted to consider $\left(\sum_{n=0}^{\infty} a_n\right)^p$ where p is a positive integer maybe larger than 2? Is there a similar theorem to the above?

Definition 8.13.2 *Define $\sum_{k_1+\cdots+k_p=m} a_{k_1} a_{k_2} \cdots a_{k_p}$ as follows. Consider all ordered lists of nonnegative integers $k_1, \cdots, k_p$ which have the property that $\sum_{i=1}^{p} k_i = m$. For each such list of integers, form the product, $a_{k_1} a_{k_2} \cdots a_{k_p}$ and then add all these products.*

Note that $\sum_{k=0}^{n} a_k a_{n-k} = \sum_{k_1+k_2=n} a_{k_1} a_{k_2}$. Therefore, from the above theorem, if $\sum a_i$ converges absolutely, it follows $\left(\sum_{i=0}^{\infty} a_i\right)^2 = \sum_{n=0}^{\infty} \left(\sum_{k_1+k_2=n} a_{k_1} a_{k_2}\right)$. It turns out a similar theorem holds for replacing 2 with p.

Theorem 8.13.3 *Suppose $\sum_{n=0}^{\infty} a_n$ converges absolutely. Then if p is a positive integer,*

$$\left(\sum_{n=0}^{\infty} a_n\right)^p = \sum_{m=0}^{\infty} c_{mp}$$

where $c_{mp} \equiv \sum_{k_1+\cdots+k_p=m} a_{k_1} \cdots a_{k_p}$.

Proof: First note this is obviously true if $p = 1$ and is also true if $p = 2$ from the above theorem. Now suppose this is true for p and consider $\left(\sum_{n=0}^{\infty} a_n\right)^{p+1}$. By the induction hypothesis and the above theorem on the Cauchy product,

$$\begin{aligned}
\left(\sum_{n=0}^{\infty} a_n\right)^{p+1} &= \left(\sum_{n=0}^{\infty} a_n\right)^p \left(\sum_{n=0}^{\infty} a_n\right) = \left(\sum_{m=0}^{\infty} c_{mp}\right) \left(\sum_{n=0}^{\infty} a_n\right) \\
&= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} c_{kp} a_{n-k}\right) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \sum_{k_1+\cdots+k_p=k} a_{k_1} \cdots a_{k_p} a_{n-k} \\
&= \sum_{n=0}^{\infty} \sum_{k_1+\cdots+k_{p+1}=n} a_{k_1} \cdots a_{k_{p+1}} \blacksquare
\end{aligned}$$

This theorem implies the following corollary for power series.

Corollary 8.13.4 *Let*

$$\sum_{n=0}^{\infty} a_n (x-a)^n$$

be a power series having radius of convergence, $r > 0$. Then if $|x-a| < r$,

$$\left(\sum_{n=0}^{\infty} a_n (x-a)^n\right)^p = \sum_{n=0}^{\infty} b_{np} (x-a)^n$$

where $b_{np} \equiv \sum_{k_1+\cdots+k_p=n} a_{k_1} \cdots a_{k_p}$.

Proof: Since $|x-a|<r$, the series, $\sum_{n=0}^{\infty} a_n (x-a)^n$, converges absolutely. Therefore, the above theorem applies and

$$\begin{aligned}\left(\sum_{n=0}^{\infty} a_n (x-a)^n\right)^p &= \sum_{n=0}^{\infty}\left(\sum_{k_1+\cdots+k_p=n} a_{k_1}(x-a)^{k_1}\cdots a_{k_p}(x-a)^{k_p}\right) \\ &= \sum_{n=0}^{\infty}\left(\sum_{k_1+\cdots+k_p=n} a_{k_1}\cdots a_{k_p}\right)(x-a)^n. \blacksquare\end{aligned}$$

With this theorem it is possible to consider the question raised in Example 8.10.3 on Page 180 about the existence of the power series for $\tan x$. This question is clearly included in the more general question of when $\left(\sum_{n=0}^{\infty} a_n (x-a)^n\right)^{-1}$ has a power series.

Lemma 8.13.5 *Let $f(x)=\sum_{n=0}^{\infty} a_n (x-a)^n$, a power series having radius of convergence $r>0$. Suppose also that $f(a)=1$. Then there exists $r_1>0$ and $\{b_n\}$ such that for all $|x-a|<r_1$,*

$$\frac{1}{f(x)}=\sum_{n=0}^{\infty} b_n (x-a)^n.$$

Proof: By continuity, there exists $r_1>0$ such that if $|x-a|<r_1$, then

$$\sum_{n=1}^{\infty} |a_n|\,|x-a|^n<1.$$

Now pick such an x. Then

$$\frac{1}{f(x)}=\frac{1}{1+\sum_{n=1}^{\infty} a_n (x-a)^n}=\frac{1}{1+\sum_{n=0}^{\infty} c_n (x-a)^n}$$

where $c_n=a_n$ if $n>0$ and $c_0=0$. Then

$$\left|\sum_{n=1}^{\infty} a_n (x-a)^n\right| \leq \sum_{n=1}^{\infty} |a_n|\,|x-a|^n<1 \tag{8.32}$$

and so from the formula for the sum of a geometric series,

$$\frac{1}{f(x)}=\sum_{p=0}^{\infty}\left(-\sum_{n=0}^{\infty} c_n (x-a)^n\right)^p.$$

By Corollary 8.13.4, this equals

$$\sum_{p=0}^{\infty}\sum_{n=0}^{\infty} b_{np}(x-a)^n \tag{8.33}$$

where $b_{np}=\sum_{k_1+\cdots+k_p=n}(-1)^p c_{k_1}\cdots c_{k_p}$. Thus $\left|b_{np}\right| \leq \sum_{k_1+\cdots+k_p=n}\left|c_{k_1}\right|\cdots\left|c_{k_p}\right| \equiv B_{np}$ and so by Theorem 8.13.3,

$$\sum_{p=0}^{\infty}\sum_{n=0}^{\infty}\left|b_{np}\right|\,|x-a|^n \leq \sum_{p=0}^{\infty}\sum_{n=0}^{\infty} B_{np}\,|x-a|^n=\sum_{p=0}^{\infty}\left(\sum_{n=0}^{\infty}|c_n|\,|x-a|^n\right)^p<\infty$$

by 8.32 and the formula for the sum of a geometric series. Since the series of 8.33 converges absolutely, Theorem 5.5.3 on Page 97 implies the series in 8.33 equals

$$\sum_{n=0}^{\infty}\left(\sum_{p=0}^{\infty} b_{np}\right)(x-a)^n$$

and so, letting $\sum_{p=0}^{\infty} b_{np} \equiv b_n$, this proves the lemma. ■

With this lemma, the following theorem is easy to obtain.

Theorem 8.13.6 *Let $f(x)=\sum_{n=0}^{\infty} a_n (x-a)^n$, a power series having radius of convergence $r>0$. Suppose also that $f(a)\neq 0$. Then there exists $r_1>0$ and $\{b_n\}$ such that for all $|x-a|<r_1$,*

$$\frac{1}{f(x)}=\sum_{n=0}^{\infty} b_n (x-a)^n .$$

Proof: Let $g(x)\equiv f(x)/f(a)$ so that $g(x)$ satisfies the conditions of the above lemma. Then by that lemma, there exists $r_1>0$ and a sequence, $\{b_n\}$ such that

$$\frac{f(a)}{f(x)}=\sum_{n=0}^{\infty} b_n (x-a)^n$$

for all $|x-a|<r_1$. Then $\frac{1}{f(x)}=\sum_{n=0}^{\infty}\widetilde{b_n}(x-a)^n$ where $\widetilde{b_n}=b_n/f(a)$. ■

There is a very interesting question related to r_1 in this theorem. Consider $f(x)=1+x^2$. In this case $r=\infty$ but the power series for $1/f(x)$ converges only if $|x|<1$. What happens is this, $1/f(x)$ will have a power series that will converge for $|x-a|<r_1$ where r_1 is the distance between a and the nearest singularity or zero of $f(x)$ in the complex plane. In the case of $f(x)=1+x^2$ this function has a zero at $x=\pm i$. This is just another instance of why the natural setting for the study of power series is the complex plane. To read more on power series, you should see the book by Apostol [3] or any text on complex variable. An introduction is given later in this book.

Chapter 9

Integration

The integral originated in attempts to find areas of various shapes and the ideas involved in finding integrals are much older than the ideas related to finding derivatives. In fact, Archimedes[1] was finding areas of various curved shapes about 250 B.C. using the main ideas of the integral. Newton and Leibniz first observed the relation between the integral and the derivative. However, their observations were incomplete because they did not have a precise definition for the integral. This came much later in the early 1800's and the first such definition sufficient to include continuous functions was due to Cauchy around 1820 who gave the first complete proof of the fundamental theorem of Calculus. Not much later, Dirichlet proved convergence of Fourier series to the mid-point of the jump of a piecewise continuous function under suitable conditions. However, a general theory for the integral which would include piecewise continuous functions did not come about till around 1854 with the work of Riemann and completed by Darboux. Lebesgue solved the hard questions about this integral in the early 1900's.

Of course people used the fundamental theorem of calculus, which was based on finding antiderivatives, to compute integrals all through the eighteenth century, but the fundamental question whether there exists an antiderivative for continuous functions was not considered till the time of Cauchy. However, using the later nineteenth century ideas of the Weierstrass approximation theorem as developed by Bernstein, we can consider this question. It will be entirely adequate to deal with all functions typically encountered in elementary calculus and is very short. After this, are sections devoted to the more general Riemann Stieltjes integrals due to Stieltjes which date from the late nineteenth century, used in number theory, probability, and functional analysis. In this more general theory, one uses an integrator function to include in the notion of an integral things like sums and a mixture of sums and integrals.

9.1 The Integral of 1700's

Recall the following definition from beginning calculus.

Definition 9.1.1 *$\int f(x)\,dx$ denotes the set of functions F which have the property that $F'(x) = f(x)$. These are called antiderivatives. When f is continuous on $[a,b]$, it is also required that F is continuous on $[a,b]$ in addition to having $F'(x) = f(x)$ on (a,b).*

From the chapter on the derivative, $\int \sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} a_k \frac{x^{k+1}}{k+1} + C$ where C is an arbitrary constant. Thus it is easy to find an antiderivative for any polynomial.

The next lemma shows that also every continuous function defined on a closed interval $[a,b]$ has an antiderivative.

Lemma 9.1.2 *Let f be a continuous, real valued function defined on $[a,b]$. Then there exists F such that $F'(x) = f(x)$ for all $x \in (a,b)$. At the end points, $F'(x)$ will refer to a one sided derivative.*

[1] Archimedes 287-212 B.C. found areas of curved regions by stuffing them with simple shapes which he knew the area of and taking a limit. He also made fundamental contributions to physics. The story is told about how he determined that a gold smith had cheated the king by giving him a crown which was not solid gold as had been claimed. He did this by finding the amount of water displaced by the crown and comparing with the amount of water it should have displaced if it had been solid gold.

Proof: Assume that $a < b$ in what follows. If not, simply switch a and b in the argument. Let $\{p_n\}$ be a sequence of polynomials for which $\|p_n - f\| \to 0$ and let $P_n'(x) = p_n(x)$ for all $x \in (a,b)$. By the mean value theorem,

$$|P_n(x) - P_n(a) - (P_m(x) - P_m(a))| =$$

$$\begin{aligned}
& |P_n(x) - P_m(x) - (P_n(a) - P_m(a))| \\
= & \ |(p_n(t) - p_m(t))(x-a)| \leq \|p_n - p_m\| \, |b-a| \\
\leq & \ (\|p_n - f\| + \|f - p_m\|) \, |b-a|
\end{aligned}$$

The right side converges to 0 as $n,m \to \infty$ and so by completeness, there exists

$$F(x) = \lim_{n\to\infty} (P_n(x) - P_n(a)),$$

this for any choice of x. It remains to verify that $F'(x) = f(x)$. Say $x \in [a,b)$ and let $h > 0$. Then by the mean value theorem,

$$\frac{P_n(x+h) - P_n(x)}{h} = \frac{(P_n(x+h) - P_n(a)) - (P_n(x) - P_n(a))}{h} = p_n(t_{hn}) \qquad (*)$$

for some $t_{hn} \in (x, x+h)$. By compactness, there is a subsequence, still denoted as t_{hn} for which $\lim_{n\to\infty} t_{hn} = t_h \in [x, x+h]$. Now

$$\begin{aligned}
|p_n(t_{hn}) - f(t_h)| &\leq |p_n(t_{hn}) - f(t_{hn})| + |f(t_{hn}) - f(t_h)| \\
&\leq \|p_n - f\| + |f(t_{hn}) - f(t_h)|
\end{aligned}$$

and so, letting $n \to \infty$, this shows, from continuity of f that $|p_n(t_{hn}) - f(t_h)| \to 0$. Taking a limit in $*$,

$$\frac{F(x+h) - F(x)}{h} = f(t_h), \ t_h \in [x, x+h]$$

Now by continuity of f, we can take a limit of this as $h \to 0$ and obtain $F'(x) = f(x)$, where $F'(x)$ is a right derivative at $x = a$. For $x \in (a,b]$, the situation is exactly the same for when h is restrained to be negative.

$$\frac{F(x+h) - F(x)}{h} = -\frac{F(x-(-h)) - F(x)}{-h} = \frac{F(x) - F(x-k)}{k}$$

where $k \equiv -h$ and so for $F'(x)$ the left derivative, it exists at each point of $(a,b]$ and equals $f(x)$ by exactly similar arguments to the above. Thus at every point of (a,b) both the right and left derivatives exist for F and both are equal to f so F is differentiable on (a,b). Also, the appropriate one sided derivatives for F exist at $x \in \{a,b\}$ and are likewise $f(x)$. ■

Definition 9.1.3 *For the rest of this section, $[a,b]$ will denote the closed interval having end points a and b but a could be larger than b or smaller than b. It is written this way to indicate that there is a direction of motion from a to b which will be reflected by the definition of the integral given below. It is an "oriented interval". Then for f continuous on $[a,b]$,*

$$\int_a^b f(x)\,dx \equiv F(b) - F(a)$$

where F is an antiderivative for f on $[a,b]$.

Proposition 9.1.4 *The integral is well defined for f continuous on $[a,b]$.*

Proof: Suppose F,G are both antiderivatives. Then letting

$$H(x) \equiv F(x) - G(x), H'(x) = 0$$

it follows by the mean value theorem, $H(b) - H(a) = 0(b-a) = 0$ so $F(b) - G(b) = F(a) - G(a)$ which implies $F(b) - F(a) = G(b) - G(a)$. ■

Proposition 9.1.5 *The above integral is well defined for f continuous on $[a,b]$ and satisfies the following properties.*

1. *$\int_a^b f dx = f(\hat{x})(b-a)$ for some $\hat{x}$ between a and $b, \hat{x} \notin \{a,b\}$. Thus $\left|\int_a^b f dx\right| \leq \|f\| |b-a|$.*

2. *If f is continuous on an interval which contains all necessary intervals,*

$$\int_a^c f dx + \int_c^b f dx = \int_a^b f dx, \text{ so } \int_a^b f dx + \int_b^a f dx = \int_b^b f dx = 0$$

3. *If $F(x) \equiv \int_a^x f dt$, Then $F'(x) = f(x)$. Also,*

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a \beta g(x) dx$$

If $a < b$, and $f(x) \geq 0$, then $\int_a^b f dx \geq 0$. Also $\left|\int_a^b f dx\right| \leq \left|\int_a^b |f| dx\right|$.

4. *$\int_a^b 1 dx = b - a$.*

Proof: The integral is well defined by Proposition 9.1.4 and Lemma 9.1.2. Consider 1. Let $F'(x) = f(x), F$ as in Lemma 9.1.2 so

$$\int_a^b f(x) dx \equiv F(b) - F(a) = f(\hat{x})(b-a)$$

for some $\hat{x}$ in the open interval determined by a,b. This is by the mean value theorem. Hence $\left|\int_a^b f dx\right| \leq \|f\| |b-a|$.

Now consider 2. Let $F' = f$ on a closed interval which contains all necessary intervals. Then from the definition,

$$\int_a^c f dx + \int_c^b f dx = F(c) - F(a) + F(b) - F(c) = F(b) - F(a) \equiv \int_a^b f(x) dx$$

Next consider 3. For $F(x) \equiv \int_a^x f(x) dx$, the definition says that $F(x) = G(x) - G(a)$ where G is an antiderivative of f. Since $F(x) = G(x) - G(a), f = G' = F'$. It follows that $F'(x) = f(x)$ with an appropriate one sided derivative at the ends of the interval. Now let $F' = f, G' = g$. Then $\alpha f + \beta g = (\alpha F + \beta G)'$ and so

$$\begin{aligned}
\int_a^b (\alpha f(x) + \beta g(x)) dx &\equiv (\alpha F + \beta G)(b) - (\alpha F + \beta G)(a) \\
&= \alpha F(b) + \beta G(b) - (\alpha F(a) + \beta G(a)) \\
&= \alpha (F(b) - F(a)) + \beta (G(b) - G(a)) \\
&\equiv \alpha \int_a^b f(x) dx + \beta \int_a \beta g(x) dx
\end{aligned}$$

If $f \geq 0, a < b$, then the mean value theorem implies that for $F' = f$, and some

$$t \in (a,b), F(b) - F(a) = \int_a^b f dx = f(t)(b-a) \geq 0.$$

Thus

$$\begin{aligned} \int_a^b (|f| - f) dx &\geq 0, \int_a^b (|f| + f) dx \geq 0 \text{ so} \\ \int_a^b |f| dx &\geq \int_a^b f dx \text{ and } \int_a^b |f| dx \geq -\int_a^b f dx \end{aligned}$$

so this proves $\left|\int_a^b f dx\right| \leq \int_a^b |f| dx$. This, along with part 2 implies the other claim that $\left|\int_a^b f dx\right| \leq \left|\int_a^b |f| dx\right|$ even if $a > b$.

The last claim is obvious because an antiderivative of 1 is $F(x) = x$. ■

Note also that the usual change of variables theorem from elementary calculus is available because, by the chain rule, if $F' = f$, then $f(g(x)) g'(x) = \frac{d}{dx} F(g(x))$ so that, from the above proposition,

$$F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(y) dy = \int_a^b f(g(x)) g'(x) dx.$$

We usually let $y = g(x)$ and $dy = g'(x) dx$ and then change the limits as indicated above, or equivalently we massage the expression to look like the above. Integration by parts also follows from differentiation rules.

Also notice that, by considering real and imaginary parts, you can define the integral of a complex valued continuous function

$$\int_a^b f(t) dt \equiv \int_a^b \operatorname{Re} f(t) dt + i \int_a^b \operatorname{Im} f(t) dt$$

and that the change of variables formula just described would hold. Just apply the above to the real and imaginary parts. Similarly, you could consider continuous functions with values in $\mathbb{R}^p$ by considering the component functions.

Definition 9.1.6 *A function $f : [a,b] \to \mathbb{R}$ is piecewise continuous if there is an ordered list of intermediate points z_i having an order consistent with*

$$[a,b], \quad (z_i - z_{i-1})(b-a) > 0$$

$a = z_0, z_1, \cdots, z_n = b$, called a partition of $[a,b]$, and functions f_i continuous on $[z_{i-1}, z_i]$ such that $f = f_i$ on (z_{i-1}, z_i). For f piecewise continuous, define

$$\int_a^b f(t) dt \equiv \sum_{i=1}^n \int_{z_{i-1}}^{z_i} f_i(s) ds$$

Observation 9.1.7 *Note that this defines the integral when the function has finitely many discontinuities and that changing the value of the function at finitely many points does not affect the integral.*

Of course this gives what appears to be a new definition because if f is continuous on $[a,b]$, then it is piecewise continuous for any such partition. However, it gives the same answer because, from this new definition,

$$\int_a^b f(t)\,dt = \sum_{i=1}^n \left(F(z_i) - F(z_{i-1})\right) = F(b) - F(a)$$

Does this give the main properties of the integral? In particular, is the integral still linear? Suppose f, g are piecewise continuous. Then let $\{z_i\}_{i=1}^n$ include all the partition points of both of these functions. Then, since it was just shown that no harm is done by including more partition points, $\int_a^b \alpha f(t) + \beta g(t)\,dt \equiv$

$$\sum_{i=1}^n \int_{z_{i-1}}^{z_i} \left(\alpha f_i(s) + \beta g_i(s)\right) ds = \sum_{i=1}^n \alpha \int_{z_{i-1}}^{z_i} f_i(s)\,ds + \sum_{i=1}^n \beta \int_{z_{i-1}}^{z_i} g_i(s)\,ds$$

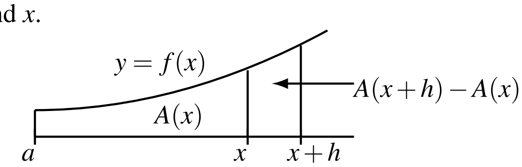
$$= \alpha \sum_{i=1}^n \int_{z_{i-1}}^{z_i} f_i(s)\,ds + \beta \sum_{i=1}^n \int_{z_{i-1}}^{z_i} g_i(s)\,ds = \alpha \int_a^b f(t)\,dt + \beta \int_a^b g(t)\,dt$$

Also, the claim that $\int_a^b f\,dt = \int_a^c f\,dt + \int_c^b f\,dt$ is obtained exactly as before by considering all partition points on each integral preserving the order of the limits in the small intervals determined by the partition points. That is, if $a > c$, you would have $z_{i-1} > z_i$ in computing $\int_a^c f\,dt$.

Definition 9.1.8 *Let I be an interval. Then $\mathscr{X}_I(t)$ is 1 if $t \in I$ and 0 if $t \notin I$. Then a step function will be of the form $\sum_{k=1}^n c_k \mathscr{X}_{I_k}(t)$ where $I_k = [a_{k-1}, a_k]$ is an interval and $\{I_k\}_{k=1}^n$ are non-overlapping intervals whose union is an interval $[a,b]$ so $b - a = \sum_{k=1}^n (a_k - a_{k-1})$. Then, as explained above,*

$$\int_a^b \sum_{k=1}^n c_k \mathscr{X}_{I_k}(t)\,dt = \sum_{k=1}^n c_k \int_{a_{k-1}}^{a_k} 1\,dt = \sum_{k=1}^n c_k (a_k - a_{k-1}).$$

The main assertion of the above Proposition 9.1.5 is that for any f continuous, there exists a unique solution to the initial value problem $F'(t) = f(t)$, along with $F(a) = 0$ and it is $F(t) = \int_a^t f(x)\,dx$. As an example of something which satisfies this initial value problem consider $A(x)$ the area under the graph of a curve $y = f(x)$ as shown in the following picture between a and x.



Thus $A(x+h) - A(x) \in [f(x)h, f(x+h)h]$ and so

$$\frac{A(x+h) - A(x)}{h} \in [f(x), f(x+h)].$$

Then taking a limit as $h \to 0$, one obtains $A'(x) = f(x), A(a) = 0$ and so one would have $A(x) = \int_a^x f(t)\,dt$. Other situations for the graph of $y = f(x)$ are similar. This suggests that we should **define** the area under the graph of the curve between a and $x > a$ as this integral.

Is this as general as a complete treatment of Riemann integration? No it is not. In particular, it does not include the well known example where $f(x) = \sin\left(\frac{1}{x}\right)$ for $x \in (0,1]$

and $f(0) \equiv 0$. However, it is sufficiently general to include all cases which are typically of interest starting with Dirichlet and his consideration of convergence of Fourier series. It is also enough to build a theory of ordinary differential equations and do all standard examples of beginning calculus. However, this integral, as well as the more general Riemann integral discussed below are woefully inadequate when it comes to a need to handle limits. You need the Lebesgue integral or something more sophisticated to obtain this. Such integrals are considered later.

9.2 The Riemann Stieltjes Integral

In this section are the principal theorems about Stieltjes integrals, including the classical Riemann integral, as a special case. A good source for more of these things is the book by Apostol, [2] and Hobson [18]. The difference here is that instead of dx you use dg where g is a function, usually of bounded variation. This is more general than the previous section. You simply take $g(x) = x$ in what follows and obtain it. I wanted to give a review of familiar material first before launching in to Stieltjes integrals.

In all which follows I will always tacitly assume that f is a bounded function defined on some finite interval.

9.3 Fundamental Definitions and Properties

First we need to define what is meant by finite total variation.

Definition 9.3.1 *Let g be a function defined on $[a,b]$. For*

$$P_{[a,x]} \equiv \{x_0, \cdots, x_n\}$$

a partition of $[a,x]$,define $V\left(P_{[a,x]}, g\right)$ by

$$\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| .$$

Define the total variation of g on $[a,x]$ by

$$V_{[a,x]}(g) \equiv \sup\left\{V\left(P_{[a,x]}, g\right) : P_{[a,x]} \text{ is a partition of } [a,x]\right\}.$$

Then g is said to be of bounded variation on $[a,b]$ if $V_{[a,b]}(g)$ is finite. Also, for $P = \{x_0, x_1, \cdots, x_n\}$ where $x_k \geq x_{k-1}$, $\|P\| \equiv \max\{|x_k - x_{k-1}| : k = 1, 2, \cdots, n\}$.

Then with this definition, one has an important proposition which pertains to the case of principal interest here in which the functions are all real valued. The above definition of finite total variation works for functions which have values in some normed linear space however.

Proposition 9.3.2 *Every real valued function g of bounded variation can be written as the difference of two increasing function, one of which is the function $x \to V_{[a,x]}(g)$. Furthermore, the functions of bounded variation are exactly those functions which are the difference of two increasing functions.*

Proof: Let g be of bounded variation. It is obvious from the definition that $x \to V_{[a,x]}(g)$ is an increasing function. Also $g(x) = V_{[a,x]}(g) - \left(V_{[a,x]}(g) - g(x)\right)$ The first part of the proposition is proved if I can show $x \to V_{[a,x]}(g) - g(x)$ is increasing. Let $x \leq y$. Is it true that $V_{[a,x]}(g) - g(x) \leq V_{[a,y]}(g) - g(y)$? This is true if and only if

$$g(y) - g(x) \leq V_{[a,y]}(g) - V_{[a,x]}(g) \tag{9.1}$$

To show this is so, first note that $V\left(P_{[a,x]}, g\right) \leq V\left(Q_{[a,x]}, g\right)$ whenever the partition $Q_{[a,x]} \supseteq P_{[a,x]}$. You demonstrate this by adding in one point at a time and using the triangle inequality. Now let P_y and $P_{[a,x]}$ be partitions of $[a,y]$ and $[a,x]$ respectively such that

$$V\left(P_{[a,x]}, g\right) + \varepsilon > V_{[a,x]}(g),\ V(P_y, g) + \varepsilon > V_{[a,y]}(g)$$

Without loss of generality P_y contains x because from what was just shown you could add in the point x and the approximation of $V(P_y, g)$ to $V_{[a,y]}(g)$ would only be better. Then from the definition,

$$\begin{aligned} V_{[a,y]}(g) - V_{[a,x]}(g) &\geq V(P_y, g) - \left(V\left(P_{[a,x]}, g\right) + \varepsilon\right) \\ &\geq |g(y) - g(x)| - \varepsilon \geq g(y) - g(x) - \varepsilon \end{aligned}$$

and since ε is arbitrary, this establishes 9.1. This proves the first part of the proposition.

Now suppose $g(x) = g_1(x) - g_2(x)$ where each g_i is an increasing function. Why is g of bounded variation? Letting $x < y$

$$\begin{aligned} |g(y) - g(x)| &= |g_1(y) - g_2(y) - (g_1(x) - g_2(x))| \\ &\leq (g_1(y) - g_1(x)) + (g_2(y) - g_2(x)) \end{aligned}$$

Therefore, if $P = \{x_0, \cdots, x_n\}$ is any partition of $[a,b]$

$$\begin{aligned} \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| &\leq \sum_{i=1}^{n} (g_1(x_i) - g_1(x_{i-1})) + (g_2(x_i) - g_2(x_{i-1})) \\ &= (g_1(b) - g_1(a)) + (g_2(b) - g_2(a)) \end{aligned}$$

and this shows $V_{[a,b]}(g) \leq (g_1(b) - g_1(a)) + (g_2(b) - g_2(a))$ so g is of bounded variation. ■

The following is the definition of the Riemann Stieltjes integral.

Definition 9.3.3 *A bounded function f defined on $[a,b]$ is said to be Riemann Stieltjes integrable if there exists a number I with the property that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if*

$$P \equiv \{x_0, x_1, \cdots, x_n\},\ a = x_0 < \cdots < x_n = b$$

is any partition having $\|P\| < \delta$, and $z_i \in [x_{i-1}, x_i]$,

$$\left| I - \sum_{i=1}^{n} f(z_i)(g(x_i) - g(x_{i-1})) \right| < \varepsilon.$$

The number $\int_a^b f(x)\, dg(x)$is defined as I. I will denote this Riemann Stieltjes sum approximating I as $\sum_P f(z_i)(g(x_i) - g(x_{i-1}))$. When f is Riemann Stieltjes integrable on $[a,b]$ with respect to g as just described, this is denoted as $f \in R([a,b], g)$ or simply as $R[a,b]$ if the definition is clear for g.

A special case is the following definition.

Definition 9.3.4 *The Riemann integral is a special case of the above in which the integrator function $g(x) = x$. We write $\int_a^b f(x)\,dg(x)$ in the form $\int_a^b f(x)\,dx$ to signify the Riemann integral.*

There is only one possible number I satisfying the above definition.

Lemma 9.3.5 *The integral $\int_a^b f(x)\,dg(x)$ is well defined in the sense that if there is such a number I, then there is only one.*

Proof: Suppose you have two of them $I, \hat{I}$ and that $P, \hat{P}$ are corresponding partitions such that $\|P\|, \|\hat{P}\|$ are both small enough that

$$\left| I - \sum_P f(z_i)(g(x_i) - g(x_{i-1})) \right| < \varepsilon, \left| \hat{I} - \sum_{\hat{P}} f(\hat{z}_i)(g(x_i) - g(x_{i-1})) \right| < \varepsilon$$

whenever z_i or $\hat{z}_i$ are in $[x_{i-1}, x_i]$. Let $Q \equiv P \cup \hat{P}$ and choose z_i and $\hat{z}_i$ to be the left endpoint of the sub intervals defined by the partition Q. Then $\|Q\| \leq \min\left(\|P\|, \|\hat{P}\|\right)$ and so $|I - S| < \varepsilon, \left|\hat{I} - S\right| < \varepsilon$ where

$$S = \sum_Q f(x_{i-1})(g(x_i) - g(x_{i-1})).$$

Then $\left|I - \hat{I}\right| \leq |I - S| + \left|S - \hat{I}\right| < 2\varepsilon$. Since ε is arbitrary, $I = \hat{I}$. ■

Next is a fairly obvious theorem which says essentially that things which hold for sums typically hold for integrals also, provided the integrals exist.

Theorem 9.3.6 *Assuming all integrals make sense, the following relation exists for f, g functions and a, b scalars.*

$$\int_a^b \left(af + b\hat{f}\right) dg = a \int_a^b f dg + b \int_a^b \hat{f} dg.$$

Assuming all integrals make sense and g is increasing, it follows that

$$\int_a^b |f|\, dg \geq \left| \int_a^b f dg \right|.$$

Also, if $a < c < b$ and all integrals make sense for $I = [a,c], [a,b], [c,b]$, then

$$\int_a^b f dg = \int_a^c f dg + \int_c^b f dg$$

Proof: Consider the first claim. Since all is assumed to make sense, (It is shown soon that continuity of the f functions and bounded variation of the g is sufficient.) there exists $\delta > 0$ such that if $||P|| < \delta$, then

$$\left| \int_a^b \left(af + b\hat{f}\right) dg - S_P\left(af + b\hat{f}\right) \right|, \left| aS_P(f) - a \int_a^b f dg \right|$$

$$\left| bS_P\left(\hat{f}\right) - b \int_a^b \hat{f} dg \right| < \varepsilon$$

where $S_P(h)$ denotes a suitable Riemann sum with respect to such a partition. Choose the same point in $[x_{i-1}, x_i]$ for each function in the list. Then from the above, $3\varepsilon >$

$$\left| bS_P\left(\hat{f}\right) + aS_P(f) - S_P\left(af + b\hat{f}\right) - \left(a\int_a^b f dg + b\int_a^b \hat{f} dg - \int_a^b \left(af + b\hat{f}\right) dg \right) \right|$$

However, $bS_P\left(\hat{f}\right) + aS_P(f) - S_P\left(af + b\hat{f}\right) = 0$ from properties of sums. Therefore,

$$\left| a\int_a^b f dg + b\int_a^b \hat{f} dg - \int_a^b \left(af + b\hat{f}\right) dg \right| < 3\varepsilon$$

and since ε is arbitrary, this shows that the expression inside $|\cdot|$ equals 0.

If g is increasing, then the Riemann sums are all nonnegative if $f \geq 0$. Thus

$$\int_a^b (|f| - f)\, dg, \ \int_a^b (|f| + f)\, dg \geq 0$$

and so

$$\int_a^b |f|\, dg \geq \max\left(\int_a^b f dg, - \int_a^b f dg \right), \ \int_a^b |f|\, dg \geq \left| \int_a^b f dg \right|$$

For the last claim, let $\delta > 0$ be such that when $\|P\| < \delta$, all integrals are approximated within ε be a Riemann sum based on P. Without loss of generality, letting P_I be such a partition for I each of the intervals needed, we can assume $P_{[a,b]}$ contains c since adding it in will not increase $\|P\|$. Also we can let each of the other two P_I be the restriction of $P_{[a,b]}$ to $[a,c]$ or $[c,b]$. Then

$$\left| \int_a^c f dg - S_{P_{[a,c]}}(f) \right| < \varepsilon, \left| \int_c^b f dg - S_{P_{[c,b]}} \right| < \varepsilon, \left| \int_a^b f dg - S_{P_{[a,b]}} \right| < \varepsilon$$

we can also pick the same intermediate point in each of these sums. Then $S_{P_{[a,b]}} = S_{P_{[a,c]}} + S_{P_{[c,b]}}$ and so, from the triangle inequality, $\left| \int_a^b f dg - \left(\int_a^c f dg + \int_c^b f dg \right) \right| < 3\varepsilon$ and since ε is arbitrary, the desired relation follows. ■

When does the integral make sense? The main result is the next theorem. We have in mind the case where f and g have real values but there is no change in the argument if they have complex values or even more general situations such as where g has values in a complete normed linear space and f has scalar values or when f has values in a normed linear space and g has scalar values or even more general situations. You simply change the meaning of the symbols used in the following argument. This is why I am being vague about where f and g have their values.

Theorem 9.3.7 *Let f be continuous on $[a,b]$ and let g be of finite total variation on $[a,b]$. Then f is Riemann Stieltjes integrable in the sense of Definition 9.3.3, $f \in R([a,b], g)$.*

Proof: Since f is continuous and $[a,b]$ is sequentially compact, it follows from Theorem 6.7.2 that f is uniformly continuous. Thus if $\varepsilon > 0$ is given, there exists $\delta > 0$ such that if $|x - y| < \delta$, then

$$|f(x) - f(y)| < \frac{\varepsilon}{2\left(V_{[a,b]}(g) + 1\right)}.$$

Let $P = \{x_0, \cdots, x_n\}$ be a partition such that $||P|| < \delta$. Now if you add in a point z on the interior of I_j and consider the new partition,

$$x_0 < \cdots < x_{j-1} < z < x_j < \cdots < x_n$$

denoting it by P',

$$S(P,f) - S(P',f) = \sum_{i=1}^{j-1} \left(f(t_i) - f(t_i')\right)(g(x_i) - g(x_{i-1}))$$

$$+ f(t_j)\left(g(x_j) - g(x_{j-1})\right) - f(t_j')\left(g(z) - g(x_{j-1})\right)$$

$$- f(t_{j+1}')(g(x_j) - g(z)) + \sum_{i=j+1}^{n} \left(f(t_i) - f(t_{i+1}')\right)(g(x_i) - g(x_{i-1}))$$

The term, $f(t_j)\left(g(x_j) - g(x_{j-1})\right)$ can be written as

$$f(t_j)\left(g(x_j) - g(x_{j-1})\right) = f(t_j)(g(x_j) - g(z)) + f(t_j)\left(g(z) - g(x_{j-1})\right)$$

and so, the middle terms can be written as

$$\begin{gathered} f(t_j)(g(x_j) - g(z)) + f(t_j)\left(g(z) - g(x_{j-1})\right) \\ - f(t_j')\left(g(z) - g(x_{j-1})\right) - f(t_{j+1}')(g(x_j) - g(z)) \end{gathered}$$

$$= \left(f(t_j) - f(t_{j+1}')\right)(g(x_j) - g(z)) + \left(f(t_j) - f(t_j')\right)\left(g(z) - g(x_{j-1})\right)$$

The absolute value of this is dominated by

$$< \frac{\varepsilon}{2\left(V_{[a,b]}(g) + 1\right)} \left(\left|g(x_j) - g(z)\right| + \left|g(z) - g(x_{j-1})\right|\right)$$

This is because the various pairs of values at which f is evaluated are closer than δ. Similarly,

$$\left|\sum_{i=1}^{j-1} \left(f(t_i) - f(t_i')\right)(g(x_i) - g(x_{i-1}))\right| \leq \sum_{i=1}^{j-1} \left|f(t_i) - f(t_i')\right| |g(x_i) - g(x_{i-1})|$$

$$\leq \sum_{i=1}^{j-1} \frac{\varepsilon}{2\left(V_{[a,b]}(g) + 1\right)} |g(x_i) - g(x_{i-1})|$$

and

$$\left|\sum_{i=j+1}^{n} \left(f(t_i) - f(t_{i+1}')\right)(g(x_i) - g(x_{i-1}))\right| \leq \sum_{i=j+1}^{n} \frac{\varepsilon}{2\left(V_{[a,b]}(g) + 1\right)} |g(x_i) - g(x_{i-1})|.$$

Thus renumbering the points to include z,

$$\left|S(P,f) - S(P',f)\right| \leq \sum_{i=1}^{n+1} \frac{\varepsilon}{2\left(V_{[a,b]}(g) + 1\right)} |g(x_i) - g(x_{i-1})| < \varepsilon/2.$$

Similar reasoning would apply if you added in two new points in the partition or more generally, any finite number of new points. You would just have to consider more exceptional terms. Therefore, if $||P|| < \delta$ and Q is any partition, then from what was just shown, you can pick the points on the intervals any way you like and

$$|S(P,f) - S(P \cup Q, f)| < \varepsilon/2.$$

Therefore, if $||P||, ||Q|| < \delta$,

$$\begin{aligned} |S(P,f) - S(Q,f)| &\leq |S(P,f) - S(P\cup Q,f)| + |S(P\cup Q,f) - S(Q,f)| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

Now consider a sequence $\varepsilon_n \to 0$. Then from what was just shown, there exist $\delta_n > 0$ such that for $\|P\|, \|Q\| < \delta_n$, $|S(P,f) - S(Q,f)| < \varepsilon_n$. Let K_n be defined by

$$K_n \equiv \overline{\cup \{S(P,f) : \|P\| < \delta_n\}}.$$

In other words, take the closure of the set of numbers consisting of all Riemann sums, $S(P,f)$ such that $\|P\| < \delta_n$. It follows from the definition, $K_n \supseteq K_{n+1}$ for all n and each K_n is closed with $\operatorname{diam}(K_n) = \varepsilon_n \to 0$. Then by Theorem 4.10.15 there exists a unique $I \in \cap_{n=1}^{\infty} K_n$. (In more general situations, you would use completeness and a corresponding theorem which says the intersection of nested closed sets having diameters converging to 0 contains a unique point. I mention this because this all holds just as well in case the continuous function has values in an infinite dimensional space in which closed and bounded sets are maybe not compact. We are not concerned with this case in this book.) Letting $\varepsilon > 0$, there exists n such that $\varepsilon_n < \varepsilon$. Then if $\|P\| < \delta_n$, it follows $|S(P,f) - I| \leq \varepsilon_n < \varepsilon$. Thus f is Riemann Stieltjes integrable in the sense of Definition 9.3.3 and $I = \int_a^b f dg$. ■

Are there easy to apply theorems which will let you conclude that something is or is not Riemann Stieltjes integrable? This involves the relation between integrability and upper and lower sums. It is specific to the case where f and g have real values and does not generalize readily like the above theorem does.

Definition 9.3.8 *Let f be real valued, bounded on $[a,b]$, and*

$$P \equiv \{x_0, \cdots, x_m\}$$

be a partition with g an increasing integrator function. Then the upper and lower sums are defined respectively as

$$U(f,P) \equiv \sum_{i=1}^{m} M_i (g(t_i) - g(t_{i-1})), \; L(f,P) \equiv \sum_{i=1}^{m} m_i (g(t_i) - g(t_{i-1}))$$

where

$$M_i \equiv \sup\{f(x) : x \in [x_{i-1}, x_i]\}, \; m_i \equiv \inf\{f(x) : x \in [x_{i-1}, x_i]\}$$

Now here is an interesting lemma about partitions.

Lemma 9.3.9 *Suppose f is a bounded function defined on $[a,b]$ and $|f(x)| < M$ for all $x \in [a,b]$ and let g be increasing. Let Q be a partition having n points, $\{x_0^*, \cdots, x_n^*\}$ and let P be any other partition. Then*

$$|U(f,P) - L(f,P)| \leq 2Mn \|P_g\| + |U(f,Q) - L(f,Q)|$$

where $\|P_g\|$ is defined by $\max\{g(x_i) - g(x_{i-1}) : P = \{x_0, \cdots, x_m\}\}$.

Proof: Let $P = \{x_0, \cdots, x_m\}$. Let I denote the set

$$I \equiv \{i : [x_{i-1}, x_i] \text{ contains some point of } Q\}$$

and $I^C \equiv \{0, \cdots, m\} \setminus I$. Then

$$U(f,P) - L(f,P) = \sum_{i \in I} (M_i - m_i)(g(x_i) - g(x_{i-1})) + \sum_{i \in I^C} (M_i - m_i)(g(x_i) - g(x_{i-1}))$$

In the second sum above, for each $i \in I^C$, $[x_{i-1}, x_i]$ must be contained in $[x_{k-1}^*, x_k^*]$ for some k because there are no points of Q in $[x_{i-1}, x_i]$. Therefore, the sum of the terms for which $[x_{i-1}, x_i]$ is contained in $[x_{k-1}^*, x_k^*]$ is no larger than the term in the sum which equals $U(f,Q) - L(f,Q)$. (Note how it is important that g be increasing.) Therefore, the second sum is no larger than $U(f,Q) - L(f,Q)$ and

$$U(f,P) - L(f,P) \leq \sum_{i \in I} (M_i - m_i)(g(x_i) - g(x_{i-1})) + U(f,Q) - L(f,Q) \tag{9.2}$$

Now consider the first sum. Since $|f(x)| \leq M$, $(M_i - m_i)(g(x_i) - g(x_{i-1})) \leq 2M \left|\left|P_g\right|\right|$ and so, since each of these intervals $[x_{i-1}, x_i]$ for $i \in I$ contains at least one point of Q, there can be no more than n of these. Hence the first sum is dominated by $2Mn \left\|P_g\right\|$. ■

The following theorem applies to the case where f, g are real valued and g is continuous. This includes the case of greatest interest which is $g(x) = x$ in which one is considering the Riemann integral. Historically, it shows that Riemann integrability is equivalent to Darboux integrability. The latter consists of defining the integral in terms of upper and lower sums, the integral being the unique number between all upper sums and all lower sums. Darboux showed that his integral was equivalent to the Riemann integral which has been described above by using the special case $g(x) = x$ which is a special case of the following theorem. Thus the following theorem is essentially due to Darboux.

Theorem 9.3.10 *Let g be an increasing continuous function and let f be real valued and $|f(x)| \leq M$ for all x. Then $f \in R([a,b], g)$ if and only if for each $\varepsilon > 0$, there exists a partition Q such that*

$$U(f,Q) - L(f,Q) < \varepsilon \tag{9.3}$$

Proof: $\Leftarrow$ Let $\varepsilon > 0$ be given and let Q be a partition such that $U(f,Q) - L(f,Q) < \varepsilon/3$. Say $Q = \{x_0^*, \cdots, x_n^*\}$. Now there exists δ such that if each $x_i - x_{i-1} < \delta$, then $|g(x_i) - g(x_{i-1})| < \frac{\varepsilon}{6Mn}$. Thus $\left\|P_g\right\|$ in the above lemma is no larger than $\frac{\varepsilon}{6Mn}$. Therefore, from Lemma 9.3.9 above, if $\|P\| < \delta$, then

$$|U(f,P) - L(f,P)| \leq 2Mn \left\|P_g\right\| + |U(f,Q) - L(f,Q)| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon.$$

It follows, since ε is arbitrary that there exists $\delta_m \to 0$ such that the diameter of S_m defined by

$$S_m \equiv \overline{\cup \left\{ \sum_P f(t_k)(g(x_k) - g(x_{k-1})) : \|P\| \leq \delta_m \right\}}$$

converges to 0. Thus there is a unique I in the intersection of these S_m and by definition, this is the integral.

$\Rightarrow$ Suppose $f \in R([a,b],g)$. Then given $\varepsilon > 0$, there is $\delta > 0$ and I such that if $\|P\| < \delta$

$$\left| \sum_P f(t_k)(g(x_k) - g(x_{k-1})) - I \right| < \frac{\varepsilon}{3}$$

Pick such a partition $P = a = x_0 < \cdots < x_n = b$. Here $t_k \in [x_{k-1}, x_k]$ is arbitrary. Pick $t_k, s_k \in [x_{k-1}, x_k]$ such that

$$f(t_k) + \frac{\varepsilon}{6(g(b) - g(a) + 1)n} > M_k,\ f(s_k) - \frac{\varepsilon}{6(g(b) - g(a) + 1)n} < m_k$$

Then letting $r(\varepsilon) \equiv \frac{\varepsilon}{6(g(b)-g(a)+1)n}$,

$$U(f,P) - L(f,P) < \sum_{i=1}^{n} (f(t_k) + r(\varepsilon) - (f(s_k) - r(\varepsilon)))(g(x_k) - g(x_{k-1}))$$

$$< \sum_{i=1}^{n} 2r(\varepsilon)(g(x_k) - g(x_{k-1})) + \left| \sum_P f(t_k)(g(x_k) - g(x_{k-1})) - I \right|$$

$$+ \left| \sum_P f(s_k)(g(x_k) - g(x_{k-1})) - I \right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \ \blacksquare$$

Example 9.3.11 *Let $g(x) = x$ be increasing and continuous and let f be decreasing. Then for $P = a = x_0 < \cdots < x_n = b$ a partition of equally spaced points,*

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{k=1}^{n} f(x_{k-1})(x_k - x_{k-1}) - \sum_{k=1}^{n} f(x_k)(x_k - x_{k-1}) \\ &= \frac{b-a}{n} \sum_{k=1}^{n} (f(x_{k-1}) - f(x_k)) = \frac{b-a}{n}(f(a) - f(b)) \end{aligned}$$

Thus for n large enough, $U(f,P) - L(f,P) < \varepsilon$. It follows $\int_a^b f dx$ exists. A similar argument shows that if f is increasing, then the integral exists.

From this important result, one can obtain fairly easily the fact that various functions of Riemann integrable functions are Riemann integrable.

Definition 9.3.12 *Let $k : D \times D \to \mathbb{R}$ satisfy*

$$\left| k(a,b) - k\left(\hat{a}, \hat{b}\right) \right| \le K\left(|a - \hat{a}| + \left| b - \hat{b} \right| \right).$$

Such a function is called Lipschitz and K is called the Lipschitz constant.

Theorem 9.3.13 *Let $f, h \in R([a,b],g)$ for g an increasing function and suppose $f([a,b]), g([a,b])$ are both contained in D. Let $k : D \times D \to \mathbb{R}$ be Lipschitz with constant K. Then $k(f,h) \in R([a,b],g)$.*

Proof: By assumption and Theorem 9.3.13, along with the observation that if $P \subseteq Q$, then $U(f,P) \ge U(f,Q)$ and $L(f,P) \le L(f,Q)$, there exists a partition P such that

$$U(f,P) - L(f,P) < \varepsilon,\ U(h,P) - L(h,P) < \varepsilon.$$

Now let

$$\begin{aligned} M_i^k &\equiv \sup \{k(f(x), h(x)) : x \in [x_{i-1}, x_i]\}, \\ m_i^k &\equiv \inf \{k(f(x), h(x)) : x \in [x_{i-1}, x_i]\} \end{aligned}$$

$$M_i^h \equiv \sup \{h(x) : x \in [x_{i-1}, x_i]\}, m_i^h \equiv \inf \{h(x) : x \in [x_{i-1}, x_i]\},$$

a similar convention holding for M_i^f, m_i^f. Here $P = \{x_0, x_1, \cdots, x_n\}$. Then

$$U(k(f,h)) - L(k(f,h)) = \sum_{k=1}^{n} \left(M_i^k - m_i^k\right)(g(x_i) - g(x_{i-1}))$$

$$\leq \sum_{k=1}^{n} K\left(\left(M_i^f - m_i^f\right) + \left(M_i^h - m_i^h\right)\right)(g(x_i) - g(x_{i-1})) \leq K(\varepsilon + \varepsilon)$$

Since ε is arbitrary, this shows $k(f,g) \in R([a,b], g)$. In case this went by too fast, see the following explanation. ■

In case there is a question about the assertion that

$$\left(M_i^k - m_i^k\right) \leq K\left(\left(M_i^f - m_i^f\right) + \left(M_i^h - m_i^h\right)\right)$$

Pick $x_M, x_m \in [x_{i-1}, x_i]$ such that $k(f(x_M), h(x_M)) > M_i^k - \varepsilon,\ k(f(x_m), h(x_m)) < m_i^k + \varepsilon$. Then

$$\left(M_i^k - m_i^k\right) \leq k(f(x_M), h(x_M)) - k(f(x_m), h(x_m)) + 2\varepsilon$$

$$\leq K(|f(x_M) - f(x_m)| + |h(x_M) - h(x_m)|) + 2\varepsilon \leq K\left(\left(M_i^f - m_i^f\right) + \left(M_i^h - m_i^h\right)\right) + 2\varepsilon$$

and since ε is arbitrary, this shows the assertion.

This theorem could be generalized by letting k be continuous, but if you want to do everything right, the context of Riemann integration is not the right place to look anyway. Much more satisfactory results are available in the theory of Lebesgue integration. I am trying to keep things simple without excluding the most important examples.

Here is an example of the kind of thing considered obtained from the above theorem.

Example 9.3.14 *Let f, h be Riemann Stieltjes integrable, real valued functions with respect to g an increasing integrator function. Then so is* $\max(f,h), \min(f,h), af + bh$ *for a, b real numbers, fh, and likely many other combinations of these functions.*

The only claims not obvious are the one about fh and $\max(f,h), \min(f,h)$. However, g, h are bounded by assumption. Therefore, h, f have all values in some interval $[-R, R]$. Let $k(a,b) \equiv ab$ for $(a,b) \in [-R,R] \times [-R,R]$.

$$\left|k(a,b) - k\left(\hat{a}, \hat{b}\right)\right| \leq |ab - \hat{a}b| + \left|\hat{a}b - \hat{a}\hat{b}\right| \leq R|a - \hat{a}| + R\left|b - \hat{b}\right|$$

As to $\max(f,h)$, it equals $\frac{|f-h|+f+h}{2}$ which clearly satisfies the necessary condition.

Note that this theorem includes the most important example in which $g(t) = t$, the Riemann integral.

The following corollary gives many examples of functions which are integrable. This corollary includes the case of Riemann integrability of a piecewise continuous function. This was first shown by Riemann. However, it is important that either f or g is continuous at the exceptional points for f.

Corollary 9.3.15 *Let g be an increasing function defined on $[a,b]$ and let $|f|$ be bounded by M and f is continuous on $[a,b]$ except for $\{c_1,\cdots,c_r\}$. At these points either f or g is continuous. Then $\int_a^b f dg$ exists.*

Proof: By Theorem 9.3.7, I just need to show that there is a partition P such that $U(f,P) - L(f,P) < \varepsilon$. Let $\|\tilde{P}\| < \delta, \tilde{P} = a = y_0 < \cdots < y_n = b$ where δ is so small that if some $c_i \in [y_{k-1}, y_k]$, then for M_k, m_k having the meaning described above as sup and inf of f on the k^{th} interval,

$$|(M_k - m_k)(g(y_k) - g(y_{k-1}))| < \frac{\varepsilon}{5r} \tag{9.4}$$

This is possible because either f or g is continuous at c_i. Note there are at most r of these intervals. The validity of the above inequality only depends on $\|\tilde{P}\|$ so to eliminate possible cases, assume that none of the y_i equal any of the finitely many points in the list except possibly y_n and y_0 if there is a discontinuity at an end point. Then there are $m \leq r+1$ closed intervals $\{I_j\}_{j=1}^m$ which remain, other than these special ones which contain an exceptional point and on each of these intervals, f is continuous. Hence denoting as $U(f,P_j)$ an upper sum corresponding to a partition P_j of I_j and $L(f,P_j)$ defined similarly, we can choose P_j on I_j such that $U(f,P_i) - L(f,P_j) < \frac{\varepsilon}{5(r+1)}$. Letting P be a partition of $[a,b]$ consisting of the y_k along with the points of each P_j, it follows that

$$\begin{aligned} U(f,P) - L(f,P) &< \sum_{j=1}^{m} (U(f,P_i) - L(f,P_j)) + r(M_k - m_k)(g(y_k) - g(y_{k-1})) \\ \leq \quad & \frac{\varepsilon}{5(r+1)} m + \frac{\varepsilon}{5} < \frac{2\varepsilon}{5} < \varepsilon \blacksquare \end{aligned}$$

Proposition 9.3.16 *Suppose $f \in R([a,b],g)$ where g is an increasing function. If $f = \hat{f}$ except at finitely many points $\{z_1,\cdots,z_n\}$ at which g is continuous, then $\hat{f} \in R([a,b],g)$ and $\int \hat{f} dg = \int f dg$.*

Proof: By assumption, there exists P such that $U(f,P) - L(f,P) < \frac{\varepsilon}{4}$. Then by adding in more points to P by including points on either side of the exceptional points and using the continuity of g at these exceptional points, we can pick these extra points close enough to the exceptional points such that if $\hat{P}$ consists of the new partition with the new points added in, $|U(\hat{f},\hat{P}) - U(f,P)| < \frac{\varepsilon}{4}$ and $|L(\hat{f},\hat{P}) - L(f,P)| < \frac{\varepsilon}{4}$. Therefore, $U(\hat{f},\hat{P}) - L(\hat{f},\hat{P}) < \frac{3\varepsilon}{4}$ showing that $\hat{f}$ is also in $R([a,b],g)$. Also, the two integrals are between all upper and lower sums. Thus these integrals are equal because if $\int \hat{f} dg \geq \int f dg$,

$$\left| \int \hat{f} dg - \int f dg \right| \leq U(\hat{f},\hat{P}) - L(f,P) \leq U(\hat{f},\hat{P}) - L(\hat{f},P) + \frac{\varepsilon}{4} < \varepsilon$$

Since ε is arbitrary, this shows the two are equal. It works the same if $\int f dg \geq \int \hat{f} dg$. ■

In case $g(x) = x$ so you are considering the Riemann integral, this theorem is more general than the one which says that piecewise continuous functions are Riemann integrable. It is more general because you could have a function which is continuous except at finitely many points but maybe the limit of the function from one side or another does not even exist. A piecewise continuous function is defined next.

Definition 9.3.17 *A bounded function $f : [a,b] \to \mathbb{R}$ is called piecewise continuous if there are points z_i such that $a = z_0 < z_1 < \cdots < z_n = b$ and continuous functions $g_i : [z_{i-1}, z_i] \to \mathbb{R}$ such that for $t \in (z_{i-1}, z_i)$, $g_i(t) = f(t)$.*

I think that the case of piecewise continuous functions is certainly the case of most interest, however. What happens is that the right and left limits of f exist at each of the exceptional points. Thus we can speak of

$$\lim_{x \to z_i +} f(x) \equiv f(z_i +), \text{ and } \lim_{x \to z_i -} f(x) \equiv f(z_i -)$$

where the first limit is taken from the right and the second limit from the left. On $[z_{i-1}, z_i]$, the function g_i equals f except at the endpoints it is the right or left limit of f.

The situation for piecewise continuous functions is stated as the following corollary. I will give a proof because this is, in my opinion, the most important case and it won't hurt to have a new and maybe more elementary proof. In this corollary, the integrator function is $g(t) = t$.

Corollary 9.3.18 *Let $f : [a,b] \to \mathbb{R}$ be piecewise continuous. Then f is Riemann integrable. Also*

$$\int_a^b f\,dt = \sum_{i=1}^n \int_{z_{i-1}}^{z_i} g_i\,dt \tag{9.5}$$

Where $g_i = f$ on (z_{i-1}, z_i) with g_i continuous on $[z_{i-1}, z_i]$.

Proof: Let P_i be a partition for $[z_{i-1}, z_i]$. Since there are only finitely many of these intervals, there exists $\delta > 0$ such that if $\|P_i\| < \delta$, then for each i,

$$\left| \sum_{P_i} g_i - \int_{z_{i-1}}^{z_i} g_i\,dt \right| < \varepsilon$$

Let M_f be an upper bound for $|f|$ on $[a,b]$, M_g an upper bound for all $|g_i|$. Now let $\|P\| < \delta < \varepsilon$ where P is a partition of $[a,b]$, these points denoted as x_j. Let $\hat{P}_i$ be those points of P which are in $(z_{i-1}, z_i]$ and let P_i consist of $\hat{P}_i$ along with z_{i-1} and z_i. Thus $\|P_i\| < \delta$. Then for $y_i \in [x_{i-1}, x_i]$,

$$\left| \sum_{i=1}^n \int_{z_{i-1}}^{z_i} g_i\,dt - \sum_P f \right| \leq \sum_{i=1}^n \left| \int_{z_{i-1}}^{z_i} g_i\,dt - \sum_{x_j \in \hat{P}_i} f(y_j)(x_j - x_{j-1}) \right|$$

Now for $x_j \in \hat{P}_i$, $f(y_j) = g_i(y_j)$ except maybe at end points where these differ by no more than $2(M_f + M_g) \equiv 2M$. Thus the above is no more than

$$\begin{aligned} &\leq \sum_{i=1}^n \left| \int_{z_{i-1}}^{z_i} g_i\,dt - \sum_{x_j \in P_i} g_i(y_j)(x_j - x_{j-1}) \right| + \sum_{i=1}^n 4(M_f + M_g)\delta \\ &< n\varepsilon + 4Mn\delta < \varepsilon(n + 4Mn) \end{aligned}$$

Since ε is arbitrary, this shows that f is indeed Riemann integrable and equals 9.5. ■

Recall that every bounded variation real function is the difference of two increasing functions. This is Proposition 9.3.2. Then one can easily generalize the above Corollary 9.3.15 to the case where g is only real and of bounded variation.

Corollary 9.3.19 *Let g be the difference of two increasing functions $g = g_1 - g_2$ defined on $[a,b]$ and let f be bounded by M and is continuous on $[a,b]$ except at finitely many points $\{c_1, \cdots, c_r\}$. At these points either f or both g_1, g_2 are continuous. Then $\int_a^b f dg$ exists.*

Proof: The details are left to the reader. However, you simply apply the above Corollary 9.3.15 to $\int_a^b f dg_1$ and $\int_a^b f dg_2$ and then note that from the definition, $\int_a^b f dg = \int_a^b f dg_1 - \int_a^b f dg_2$ ■

The above has given many examples of functions which are integrable. Now here is one which is not integrable with respect to any continuous bounded variation function.

Example 9.3.20 *Here is an example of a function which is not integrable. Let*

$$f(t) \equiv \begin{cases} 1 \text{ if } x \text{ is rational} \\ 0 \text{ if } x \text{ is not rational} \end{cases} \tag{9.6}$$

and let g be an increasing continuous function, $g(b) > g(a)$. Then in reference to $[a,b]$, $U(f,g) = g(b) - g(a)$ and $L(f,g) = 0$ so $\int_a^b f dg$ does not exist. This follows from Theorem 9.3.10.

There is a fundamental relationship between $f \in R([a,b], g)$ and $g \in R([a,b], f)$. It turns out that if you have one, then you have the other also and in addition to this, there is a fundamental integration by parts formula. This is a very remarkable formula.

9.4 Integration by Parts

Theorem 9.4.1 *Let f, g be two functions defined on $[a,b]$. Suppose $f \in R([a,b], g)$. Then $g \in R([a,b], f)$ and the following integration by parts formula holds.*

$$\int_a^b f dg + \int_a^b g df = fg(b) - fg(a).$$

Proof: By definition there exists $\delta > 0$ such that if $\|P\| < \delta$ then whenever $z_i \in [x_{i-1}, x_i]$,

$$\left| \sum_{i=1}^n f(z_i)(g(x_i) - g(x_{i-1})) - \int_a^b f dg \right| < \varepsilon$$

Pick such a partition. Notice $fg(b) - fg(a) = \sum_{i=1}^n fg(x_i) - fg(x_{i-1})$. Therefore, subtracting $\sum_{i=1}^n g(t_i)(f(x_i) - f(x_{i-1}))$ from both sides where $t_i \in [x_{i-1}, x_i]$,

$$fg(b) - fg(a) - \sum_{i=1}^n g(t_i)(f(x_i) - f(x_{i-1}))$$

$$= \sum_{i=1}^n (fg(x_i) - fg(x_{i-1})) - \sum_{i=1}^n g(t_i)(f(x_i) - f(x_{i-1}))$$

$$= \sum_{i=1}^n f(x_i)(g(x_i) - g(t_i)) + f(x_{i-1})(g(t_i) - g(x_{i-1}))$$

But this is just a Riemann Stieltjes sum for $\int_a^b f dg$ corresponding to the partition which consists of all the x_i along with all the t_i and if P' is this partition, $\|P'\| < \delta$ because it has at least as many points in it as P. Therefore,

$$\left| \overbrace{\left(fg(b) - fg(a) - \underbrace{\sum_{i=1}^{n} g(t_i)(f(x_i) - f(x_{i-1}))}_{\text{Riemann sum for } \int_a^b g df} \right)}^{\text{Riemann sum for } \int_a^b f dg} - \int_a^b f dg \right| < \varepsilon$$

and this has shown that from the definition, $g \in R([a,b], f)$ and

$$\int_a^b g df = fg(b) - fg(a) - \int_a^b f dg. \blacksquare$$

It is an easy theorem to remember. Think something sloppy like this: $d(fg) = fdg + gdf$ and so

$$fg(b) - fg(a) = \int_a^b d(fg) = \int_a^b fdg + \int_a^b gdf \tag{9.7}$$

and all you need is for at least one of these integrals on the right to make sense. Then the other automatically does and the formula follows.

Corollary 9.4.2 *$\int_a^b f dg$ exists if f is continuous and g of bounded variation or if g is continuous and f of bounded variation.*

Proof: This follows from the above integration by parts result and Theorem 9.3.7. ■

The following proposition shows $\int_a^b f(t) dg(t) = \int_a^b f(t) g'(t) dt$ under suitable assumptions. Actually, this will be true even if g has values in some vector space. To begin with is an argument which could be used to show this. This argument is technical and for the topics of interest in this book, the proposition is sufficient. Therefore, you might want to go directly to the proposition. The reason that it is more complicated is that the mean value theorem is not available unless the function is real valued.

Suppose now that g is differentiable on $[a,b]$ and has a continuous derivative $x \to g'(x)$. Consider the following function where $(x,y) \in [a,b] \times [a,b]$

$$k(x,y) \equiv \begin{cases} \frac{|g(y) - g(x) - g'(x)(y-x)|}{|y-x|} \text{ if } x \neq y \\ 0 \text{ if } x = y \end{cases}$$

Then k is continuous in the sense that if $x_n \to x, y_n \to y$, it follows that $k(x_n, y_n) \to k(x,y)$. Thus it is uniformly continuous by Theorem 6.7.2. Thus if $\|P\| < \delta$, for P a partition of $[a,b]$,

$$\begin{aligned} \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| &\leq \sum_{i=1}^{n} |g'(x_i)| |x_i - x_{i-1}| + \sum_{i=1}^{n} \varepsilon |x_i - x_{i-1}| \\ &\leq (\varepsilon + \max\{|g'(x)| : x \in [a,b]\})(b-a) \end{aligned}$$

Now note that whenever Q is a partition, the estimate for total variation obtained from $V(Q_b, g) \leq V((Q \cup P)_b, g)$. Therefore, in finding the total variation of g one can assume that all partitions have norm no more than δ. Then since ε is arbitrary, this shows that

$$V_{[a,b]}(g) \leq |b-a| \max\left\{\left|g'(x)\right| : x \in [a,b]\right\}$$

Then by Theorem 9.4.1, both $\int_a^b f dg$ and $\int_a^b f g' dt$ exist. Then one can use a similar argument to what is about to be presented in the next proposition to conclude that these two integrals are equal. The only difference is that in the general case, you would need to avoid the mean value theorem and instead use a left sum to approximate the difference in g and a similar argument to what was just presented to show that g is of finite total variation. The reason for noting this argument is to allow the possibility that g has vector values, possibly in an infinite dimensional space. However, for this book, we are mainly concerned with g having real values. For this case, there is a much easier argument based on the mean value theorem.

In the assumptions for the following Proposition, it is easier to simply assume that $g'(x)$ exists and g' is continuous on $[a,b]$ where $g'(a), g'(b)$ are defined as appropriate one sided derivatives, but in the interest of generality, it is only assumed in the following that $g'(t)$ exists on (a,b) and is continuous on $[a,b]$. However, giving the condition on $[a,b]$ is likely easier to remember and will suffice in all typical situations.

Proposition 9.4.3 *Let g be real valued, continuous on $[a,b]$, and have a derivative which is continuous and bounded on (a,b). Then g has finite total variation and for f continuous,*

$$\int_a^b f(t)\, dg(t) = \int_a^b f(t)\, g'(t)\, dt,$$

the integral on the right existing because $t \to f(t) g'(t)$ is continuous on (a,b), bounded, and $\gamma(t) \equiv t$ is continuous at the endpoints, Corollary 9.3.15.

Proof: First consider the claim that g has finite total variation. Let $a = x_0 < x_1 < \cdots < x_n = b$. By the mean value theorem,

$$\sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| = \sum_{i=1}^{n} \left|g'(s_i)\right| (t_i - t_{i-1}) \leq \max\left\{\left|g'(x)\right| : x \in [a,b]\right\}(b-a).$$

Thus $V_{[a,b]}(g) < \infty$.

Now, by Theorem 9.4.1, $\int_a^b f(t) g'(t)\, dt, \int_a^b f(t)\, dg(t)$ both exist. Therefore, there exists a δ such that if $||P|| < \delta$, then letting P be $\{x_0, \cdots, x_n\}$,

$$\begin{aligned}
\left|\sum_{i=1}^{n} f(u_i)(g(x_i) - g(x_{i-1})) - \int_a^b f(t)\, dg(t)\right| &< \varepsilon, \\
\left|\sum_{i=1}^{n} f(u_i) g'(u_i)(x_i - x_{i-1}) - \int_a^b f(t) g'(t)\, dt\right| &< \varepsilon
\end{aligned}$$

for any choice of $u_i \in [x_{i-1}, x_i]$. But in the first expression, the mean value theorem implies $g(x_i) - g(x_{i-1}) = g'(v_i)(x_i - x_{i-1})$ for some $v_i \in (x_{i-1}, x_i)$. Let $u_i = v_i$ in the top inequality to obtain

$$\begin{aligned} & \left| \sum_{i=1}^{n} f(v_i)(g(x_i) - g(x_{i-1})) - \int_a^b f(t)\, dg(t) \right| \\ = & \left| \sum_{i=1}^{n} f(v_i) g'(v_i)(x_i - x_{i-1}) - \int_a^b f(t)\, dg(t) \right| < \varepsilon \end{aligned}$$

Then let $u_i = v_i$ in the bottom inequality as well to obtain

$$\begin{aligned} \left| \int_a^b f(t)\, dg(t) - \int_a^b f(t) g'(t)\, dt \right| &\leq \left| \int_a^b f(t)\, dg(t) - \sum_{i=1}^{n} f(v_i) g'(v_i)(x_i - x_{i-1}) \right| \\ &+ \left| \sum_{i=1}^{n} f(v_i) g'(v_i)(x_i - x_{i-1}) - \sum_{i=1}^{n} f(v_i)(g(x_i) - g(x_{i-1})) \right| \\ &+ \left| \sum_{i=1}^{n} f(v_i)(g(x_i) - g(x_{i-1})) - \int_a^b f(t)\, dg(t) \right| < \varepsilon + 0 + \varepsilon = 2\varepsilon \end{aligned}$$

Since ε is arbitrary, this shows $\int_a^b f(t)\, dg(t) = \int_a^b f(t) g'(t)\, dt$. ■

What does the integration by parts formula 9.7 say in case g' exists and is continuous and f' exists and is continuous? By Proposition 9.4.3 above, both $\int_a^b f dg$ and $\int_a^b g df$ exist. Now the integration by parts formula says $fg(b) - fg(a) = \int_a^b f dg + \int_a^b g df$ and from what was just shown in Proposition 9.4.3, this reduces to

$$fg(b) - fg(a) = \int_a^b f(t) g'(t)\, dt + \int_a^b g(t) f'(t)\, dt \tag{9.8}$$

which is the usual integration by parts formula from calculus.

Proposition 9.4.4 *Let f, g both be continuous on $[a,b]$ with continuous bounded derivatives on (a,b). Then the usual calculus integration by parts formula 9.8 is valid.*

9.5 The Fundamental Theorem of Calculus

Note how as a special case, you get the usual fundamental theorem of calculus by letting $f(t) \equiv 1$. Indeed, from Theorem 9.4.1

$$\int_a^b 1 g'(t)\, dt = \int_a^b 1 dg(t) + \overbrace{\int_a^b g df}^{\text{obviously } 0} = 1g(b) - 1g(a) = g(b) - g(a)$$

This proves:

Theorem 9.5.1 *If g' is continuous on $[a,b]$, then $g(b) - g(a) = \int_a^b g'(t)\, dt$.*

A version of this presented more directly is the following.

Theorem 9.5.2 *Suppose $\int_a^b f(t)\, dt$ exists and $F'(t) = f(t)$ for each $t \in (a,b)$ for some F continuous on $[a,b]$. Then $\int_a^b f(t)\, dt = F(b) - F(a)$.*

Proof: There exists $\delta > 0$ such that if $\|P\| < \delta$, then for $P = x_0, \cdots, x_n$,

$$\left| \int_a^b f(t)\,dt - \sum_{k=1}^n f(t_k)(x_k - x_{k-1}) \right| < \varepsilon, \text{ for any } t_k \in [x_{k-1}, x_k].$$

Use the mean value theorem to pick $t_k \in (x_{k-1}, x_k)$ such that $f(t_k)(x_k - x_{k-1}) = F(x_k) - F(x_{k-1})$. Then

$$\begin{aligned} \left| \int_a^b f(t)\,dt - \sum_{k=1}^n f(t_k)(x_k - x_{k-1}) \right| &= \left| \int_a^b f(t)\,dt - \sum_{k=1}^n F(x_k) - F(x_{k-1}) \right| \\ &= \left| \int_a^b f(t)\,dt - (F(b) - F(a)) \right| < \varepsilon \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this proves the theorem. ■

Note that this did not explicitly require f to be continuous, just Riemann integrable. In case f is continuous or piecewise continuous, this shows the integral discussed here is the same as the integral of Definition 9.1.3 because, as shown in Proposition 9.1.5 both can be computed by using an antiderivative. If $\widehat{\int}$ is the earlier 1700's integral and if f is piecewise continuous with exceptional points between a and b, $\alpha_0, \alpha_1, ..., \alpha_n$, and $f = f_i$ on (α_{i-1}, α_i), then

$$\widehat{\int_a^b} f dx \equiv \sum_{i=1}^n \widehat{\int}_{\alpha_{i-1}}^{\alpha_i} f_i dx = \sum_{i=1}^n \int_{\alpha_{i-1}}^{\alpha_i} f_i dx = \int_a^b f dx.$$

Definition 9.5.3 *When everything makes sense,* $\int_a^b f(t)\,dg(t) \equiv -\int_b^a f(t)\,dg(t)$

This also shows the well known change of variables formula.

Theorem 9.5.4 *Suppose $F'(y) = f(y)$ for y between $\phi(a)$ and $\phi(b)$ and f, ϕ' are continuous. Then*

$$\int_{\phi(a)}^{\phi(b)} f(y)\,dy = \int_a^b f(\phi(t))\,\phi'(t)\,dt$$

Proof: This is true because both sides reduce to $F(\phi(b)) - F(\phi(a))$. ■

The other form of the fundamental theorem of calculus is also obtained. Note that in case of the Riemann integral if $f \geq 0$, and $a \leq b$, then $\int_a^b f(t)\,dt \geq 0$. Therefore, assuming f is continuous,

$$\int_a^b (|f(t)| - f(t))\,dt \geq 0, \quad \int_a^b (|f(t)| + f(t))\,dt \geq 0$$

and so $\int_a^b |f(t)|\,dt \geq \int_a^b f(t)\,dt$, $\int_a^b |f(t)|\,dt \geq -\int_a^b f(t)\,dt$ which implies $\int_a^b |f(t)|\,dt \geq \left| \int_a^b f(t)\,dt \right|$. This also follows from Theorem 9.3.6.

Theorem 9.5.5 *Let f be continuous on $[a,b]$ and suppose $V_{[a,b]}(g) < \infty$. Then $f \in R([a,x], g)$ for all $x \in [a,b]$. If $g(t) = t$, then for $F(x) = \int_a^x f(t)\,dt$, it follows that $F'(t) = f(t)$ for $t \in (a,b)$.*

Proof: The first part is obvious because $V_{[a,x]}(g) \leq V_{[a,b]}(g) < \infty$ and so $f \in R([a,x], g)$ by Theorem 9.3.7. It remains to verify the last part. Let $t \in (a,b)$ and let $|h|$ be small enough that everything of interest is in $[a,b]$. First suppose $h > 0$. Then

$$\left|\frac{F(t+h) - F(t)}{h} - f(t)\right| = \left|\frac{1}{h}\int_t^{t+h} f(s)\,ds - \frac{1}{h}\int_t^{t+h} f(t)\,ds\right|$$

$$\leq \frac{1}{h}\int_t^{t+h} |f(s) - f(t)|\,ds \leq \frac{1}{h}\int_t^{t+h} \varepsilon ds = \varepsilon$$

provided that ε is small enough due to continuity of f at t. A similar inequality is obtained if $h < 0$ except in the argument, you will have $t + h < t$ so you have to switch the order of integration in going to the second line and replace $1/h$ with $1/(-h)$. Thus $\lim_{h\to 0} \frac{F(t+h)-F(t)}{h} = f(t)$. ■

An examination of the proof along with Corollary 9.3.15 yields the following corollary.

Corollary 9.5.6 *Let f be continuous on $[a,b]$ except for finitely many points. Then $F(x) \equiv \int_a^x f dx$ exists for all $x \in [a,b]$. $F'(x) = f(x)$ for x any point in (a,b) at which f is continuous.*

9.6 Uniform Convergence and the Integral

It turns out that uniform convergence is very agreeable in terms of the integral. The following is the main result.

Theorem 9.6.1 *Let g be of bounded variation and let f_n be continuous and converging uniformly to f on $[a,b]$. Then f is also integrable and $\int_a^b f dg = \lim_{n\to\infty} \int_a^b f_n dg$.*

Proof: The uniform convergence implies f is also continuous. See Theorem 6.9.7. Therefore, $\int_a^b f dg$ exists. Now let n be given large enough that

$$||f - f_n|| \equiv \max_{x\in[a,b]} |f(x) - f_n(x)| < \varepsilon$$

Next pick $\delta > 0$ small enough that if $||P|| < \delta$, then

$$\left|\int_a^b f dg - \sum_{k=1}^n f(t_k)(g(x_k) - g(x_{k-1}))\right| < \varepsilon$$

$$\left|\int_a^b f_n dg - \sum_{k=1}^n f_n(t_k)(g(x_k) - g(x_{k-1}))\right| < \varepsilon$$

for any choice $t_k \in [x_{k-1}, x_k]$. Pick such a P and the same t_k for both sums. Then

$$\left|\int_a^b f dg - \int_a^b f_n dg\right| \leq \left|\int_a^b f dg - \sum_{k=1}^n f(t_k)(g(x_k) - g(x_{k-1}))\right|$$

$$+\left|\sum_{k=1}^n (f(t_k) - f_n(t_k))(g(x_k) - g(x_{k-1}))\right| + \left|\int_a^b f_n dg - \sum_{k=1}^n f_n(t_k)(g(x_k) - g(x_{k-1}))\right|$$

$$< \varepsilon + \sum_{k=1}^n \varepsilon |g(x_k) - g(x_{k-1})| + \varepsilon \leq 2\varepsilon + V_{[a,b]}(g)\varepsilon$$

Since ε is arbitrary, this shows that $\lim_{n\to\infty} \left|\int_a^b f dg - \int_a^b f_n dg\right| = 0$. ■

9.7 A Simple Procedure for Finding Integrals

Suppose f is a continuous function and F is an increasing integrator function. How do you find $\int_a^b f(x)\,dF$? Is there some sort of easy way to do it which will handle lots of simple cases? It turns out there is a way. It is based on Lemma 9.4.3. First of all

$$F(x+) \equiv \lim_{y\to x+} F(y),\; F(x-) \equiv \lim_{y\to x-} F(y)$$

For an increasing function F, the jump of the function at x equals $F(x+) - F(x-)$.

Procedure 9.7.1 *Suppose f is continuous on $[a,b]$ and F is an increasing function defined on $[a,b]$ such that there are finitely many intervals determined by the partition $a = x_0 < x_1 < \cdots < x_n = b$ which have the property that on $[x_i, x_{i+1}]$, the following function is differentiable and has a continuous derivative.*

$$G_i(x) \equiv \begin{cases} F(x) \text{ on } (x_i, x_{i+1}) \\ F(x_i+) \text{ when } x = x_i \\ F(x_{i+1}-) \text{ when } x = x_{i+1} \end{cases}$$

Also assume $F(a) = F(a+), F(b) = F(b-)$. Then

$$\int_a^b f(x)\,dF = \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f(x)\,G_j'(x)\,dx + \sum_{i=1}^{n-1} f(x_i)\left(F(x_i+) - F(x_i-)\right)$$

Here is why this procedure works. Let δ be very small and consider the partition

$$\begin{aligned} a &= x_0 < x_1 - \delta < x_1 < x_1 + \delta < x_2 - \delta < x_2 < x_2 + \delta < \\ \cdots x_{n-1} - \delta &< x_{n-1} < x_{n-1} + \delta < x_n - \delta < x_n = b \end{aligned}$$

where δ is also small enough that whenever $|x-y| < \delta$, it follows $|f(x) - f(y)| < \varepsilon$. Then from the properties of the integral presented above,

$$\int_a^{x_1-\delta} f dF + \int_{x_1+\delta}^{x_2-\delta} f dF + \cdots + \int_{x_{n-1}+\delta}^{b} f dF + \sum_{i=1}^{n-1} (f(x_i) - \varepsilon)(F(x_i+\delta) - F(x_i-\delta))$$

$$\begin{aligned} \leq \quad & \int_a^b f dF \leq \int_a^{x_1-\delta} f dF + \int_{x_1+\delta}^{x_2-\delta} f dF + \cdots + \int_{x_{n-1}+\delta}^{b} f dF \\ & + \sum_{i=1}^{n-1} (f(x_i) + \varepsilon)(F(x_i+\delta) - F(x_i-\delta)) \end{aligned}$$

By Lemma 9.4.3 this implies

$$\begin{aligned} & \int_a^{x_1-\delta} f G_0' dx + \int_{x_1+\delta}^{x_2-\delta} f G_1' dx + \cdots + \int_{x_{n-1}+\delta}^{b} f G_{n-1}' dx \\ & + \sum_{i=1}^{n-1} (f(x_i) - \varepsilon)(F(x_i+\delta) - F(x_i-\delta)) \leq \int_a^b f dF \leq \end{aligned}$$

$$\int_a^{x_1-\delta} fG_0' dx + \int_{x_1+\delta}^{x_2-\delta} fG_1' dx + \cdots + \int_{x_{n-1}+\delta}^{b} fG_{n-1}' dx$$
$$+ \sum_{i=1}^{n-1} \left(f\left(x_i\right)+\varepsilon\right)\left(F\left(x_i+\delta\right)-F\left(x_i-\delta\right)\right)$$

Now let $\delta \to 0$ to obtain the desired integral is between

$$\sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f(x) G_j'(x) dx + \sum_{i=1}^{n-1} \left(f\left(x_i\right)+\varepsilon\right)\left(F\left(x_i+\right)-F\left(x_i-\right)\right)$$

and

$$\sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f(x) G_j'(x) dx + \sum_{i=1}^{n-1} \left(f\left(x_i\right)-\varepsilon\right)\left(F\left(x_i+\right)-F\left(x_i-\right)\right)$$

Since ε is arbitrary, this shows the procedure is valid. This yields the following right away.

Lemma 9.7.2 *Let F be increasing. Then $\mathscr{X}_{[a,b]} \in R([a,b], F)$ and the following formula is valid.* $\int_a^b \mathscr{X}_{[a,b]} dF = F(b+) - F(a-)$.

9.8 Stirling's Formula

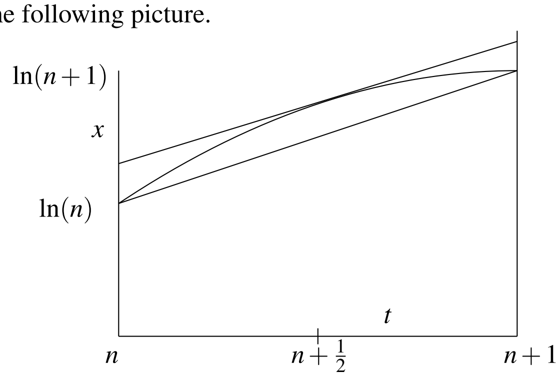
In this section is an elementary approach to Stirlings formula. This formula is an asymptotic approximation for $n!$. It is quite old, dating to about 1730. The approach followed here is like the one in the Calculus book of Courant found in the references. Later I will give a different one found in [24]. See also [9].

To begin with is a simple lemma which really depends on the shape of the graph of $t \to \ln t$.

Lemma 9.8.1 *For n a positive integer,*

$$\frac{1}{2}\left(\ln(n+1)+\ln(n)\right) \leq \int_n^{n+1} \ln(t)\, dt \leq \ln\left(n+\frac{1}{2}\right) \tag{9.9}$$

Proof: Consider the following picture.



There are two trapezoids, the area of the larger one is larger than $\int_n^{n+1} \ln(t)\, dt$ and the area of the smaller being smaller than this integral. The equation of the line which forms the top of the large trapezoid is

$$y - \ln\left(n+\frac{1}{2}\right) = \frac{1}{n+\frac{1}{2}}\left(x - \left(n+\frac{1}{2}\right)\right)$$

Thus the area of the large trapezoid is obtained by averaging the two vertical sides and multiplying by the length of the base which is 1. This is easily found to be $\ln\left(n+\frac{1}{2}\right)$. Then the area of the smaller trapezoid is obtained also as the average of the two vertical sides times the length of the base which is $\frac{1}{2}\left(\ln(n+1)+\ln(n)\right)$. ■

Now observe the following:

$$\exp\left(\sum_{k=1}^{n-1}\frac{1}{2}\left(\ln(k)+\ln(k+1)\right)\right)=\prod_{k=1}^{n-1}\left(k(k+1)\right)^{1/2}$$

$$=(1\cdot 2)^{1/2}(2\cdot 3)^{1/2}\cdots((n-1)\cdot n)^{1/2}=(n-1)!\sqrt{n}=n!n^{-1/2}$$

Letting $T_n\equiv\sum_{k=1}^{n-1}\frac{1}{2}\left(\ln(k)+\ln(k+1)\right), \exp(T_n)=n!n^{-1/2}$. Then

$$\begin{aligned}
&\int_1^n \ln(t)\,dt-T_n \\
\leq\ &\sum_{k=1}^{n-1}\ln\left(k+\frac{1}{2}\right)-\sum_{k=1}^{n-1}\frac{1}{2}\left(\ln(k)+\ln(k+1)\right)
\end{aligned}$$

$$\begin{aligned}
=\ &\frac{1}{2}\sum_{k=1}^{n-1}\left(\ln\left(k+\frac{1}{2}\right)-\ln(k)\right)-\frac{1}{2}\sum_{k=1}^{n-1}\left(\ln(k+1)-\ln\left(k+\frac{1}{2}\right)\right) \\
\leq\ &\frac{1}{2}\sum_{k=1}^{n-1}\left(\ln(k)-\ln\left(k-\frac{1}{2}\right)\right)-\frac{1}{2}\sum_{k=1}^{n-1}\left(\ln(k+1)-\ln\left(k+\frac{1}{2}\right)\right)
\end{aligned}$$

$$\begin{aligned}
=\ &\frac{1}{2}\sum_{k=1}^{n-1}\left(\ln(k)-\ln\left(k-\frac{1}{2}\right)\right)-\frac{1}{2}\sum_{k=2}^{n}\left(\ln(k)-\ln\left(k-\frac{1}{2}\right)\right) \\
=\ &\frac{1}{2}\ln 2-\left(\frac{1}{2}\ln(n)-\ln\left(n-\frac{1}{2}\right)\right)\leq\frac{\ln 2}{2}
\end{aligned}$$

$$\begin{aligned}
&\left(\int_1^{n+1}\ln(t)\,dt-T_{n+1}\right)-\left(\int_1^n \ln(t)\,dt-T_n\right) \\
=\ &\int_n^{n+1}\ln(t)\,dt-\left(\frac{1}{2}\left(\ln(n)+\ln(n+1)\right)\right)\geq 0
\end{aligned}$$

Thus $\left\{\int_1^n \ln(t)\,dt-T_n\right\}$ increases to some $\alpha\leq\frac{\ln 2}{2}$. Doing the integral,

$$(n\ln n-n)-T_n\to\alpha$$

and so taking the exponential,

$$\frac{n^n}{e^n n!n^{-1/2}}\to e^{\alpha}$$

This has proved the following lemma.

Lemma 9.8.2 *$\lim_{n\to\infty}\frac{n!e^n}{n^{n+1/2}}=c$ for some positive number c.*

In many applications, the above is enough. However, the constant can be found. There are various ways to show that this constant c equals $\sqrt{2\pi}$. The version given here also includes a formula which is interesting for its own sake.

Using integration by parts, it follows that whenever n is a positive integer larger than 1,

$$\int_0^{\pi/2} \sin^n(x)\,dx = \frac{n-1}{n}\int_0^{\pi/2} \sin^{n-2}(x)\,dx$$

Lemma 9.8.3 *For $m \geq 1$,*

$$\begin{aligned}\int_0^{\pi/2} \sin^{2m}(x)\,dx &= \frac{(2m-1)(2m-3)\cdots 1}{2m(2m-2)\cdots 2}\frac{\pi}{2}, \\ \int_0^{\pi/2} \sin^{2m+1}(x)\,dx &= \frac{(2m)(2m-2)\cdots 2}{(2m+1)(2m-1)\cdots 3}.\end{aligned}$$

Proof: Consider the first formula in the case where $m = 1$. From beginning calculus, $\int_0^{\pi/2} \sin^2(x)\,dx = \frac{\pi}{4} = \frac{1}{2}\frac{\pi}{2}$ so the formula holds in this case. Suppose it holds for m. Then from the above reduction identity and induction,

$$\begin{aligned}\int_0^{\pi/2} \sin^{2m+2}(x)\,dx &= \frac{2m+1}{2(m+1)}\int_0^{\pi/2} \sin^{2m}(x)\,dx \\ &= \frac{2m+1}{2(m+1)}\frac{(2m-1)\cdots 1}{2m(2m-2)\cdots 2}\frac{\pi}{2}.\end{aligned}$$

The second claim is proved similarly. ■

Then using the reduction identity and the above,

$$\frac{2m+1}{2m} \geq \frac{\int_0^{\pi/2} \sin^{2m}(x)\,dx}{\frac{2m}{2m+1}\int_0^{\pi/2} \sin^{2m-1}(x)\,dx} = \frac{\int_0^{\pi/2} \sin^{2m}(x)\,dx}{\int_0^{\pi/2} \sin^{2m+1}(x)\,dx} =$$

$$= \frac{\pi}{2}(2m+1)\frac{(2m-1)^2(2m-3)^2\cdots 1}{2^{2m}(m!)^2} \geq 1$$

It follows from the squeezing theorem that

$$\lim_{m\to\infty}\frac{1}{2m+1}\frac{2^{2m}(m!)^2}{(2m-1)^2(2m-3)^2\cdots 1} = \frac{\pi}{2}$$

This exceedingly interesting formula is Wallis' formula.

Now multiply both the top and the bottom of the expression on the left by

$$(2m)^2(2(m-1))^2\cdots 2^2$$

which is $2^{2m}(m!)^2$. This is another version of the Wallis formula.

$$\frac{\pi}{2} = \lim_{m\to\infty}\frac{2^{2m}}{2m+1}\frac{2^{2m}(m!)^2(m!)^2}{((2m)!)^2}$$

It follows that

$$\sqrt{\frac{\pi}{2}} = \lim_{m\to\infty}\frac{2^{2m}}{\sqrt{2m+1}}\frac{(m!)^2}{(2m)!} = \lim_{m\to\infty}\frac{2^{2m}}{\sqrt{2m}}\frac{(m!)^2}{(2m)!} \tag{9.10}$$

Now with this result, it is possible to find c in Stirling's formula. Recall

$$\lim_{m\to\infty}\frac{m!}{m^{m+(1/2)}e^{-m}c}=1=\lim_{m\to\infty}\frac{m^{m+(1/2)}e^{-m}c}{m!}$$

In particular, replacing m with $2m$,

$$\lim_{m\to\infty}\frac{(2m)!}{(2m)^{2m+(1/2)}e^{-2m}c}=\lim_{m\to\infty}\frac{(2m)^{2m+(1/2)}e^{-2m}c}{(2m)!}=1$$

Therefore, from 9.10,

$$\sqrt{\frac{\pi}{2}}=\lim_{m\to\infty}\frac{\left(\frac{m^{m+(1/2)}e^{-m}c}{m!}\right)^2}{\left(\frac{(2m)^{2m+(1/2)}e^{-2m}c}{2m!}\right)}\left(\frac{2^{2m}}{\sqrt{2m}}\frac{(m!)^2}{(2m)!}\right)=\lim_{m\to\infty}\frac{2^{2m}}{\sqrt{2m}}\frac{\left(m^{m+(1/2)}e^{-m}c\right)^2}{\left((2m)^{2m+(1/2)}e^{-2m}c\right)}$$

$$=c\lim_{m\to\infty}\frac{2^{2m}}{\sqrt{2m}}\frac{m^{2m+1}}{2^{2m+1/2}\left(m^{2m+(1/2)}\right)}=c\lim_{m\to\infty}\frac{1}{2}\frac{m^{2m+1}}{m^{2m+1}}=\frac{c}{2}$$

so $c=\sqrt{2\pi}$. This proves Stirling's formula.

Theorem 9.8.4 $\lim_{m\to\infty}\frac{m!}{m^{m+(1/2)}e^{-m}}=\sqrt{2\pi}$.

9.9 Fubini's Theorem an Introduction

Fubini's theorem has become the name of a theorem which involves interchanging the order of integration in iterated integrals. You may have seen it mentioned in a beginning calculus course. It is actually an incredibly deep result, much more so than what will be indicated here. Here I will only consider enough to allow what will be done in this book. It turns out that iterated integrals are what occur naturally in many situations, and each integral in an iterated integral is a one dimensional notion, so it is natural to consider the interchange of iterated integrals in a book on single variable calculus. All of this depends on the theorems about continuous functions defined on a subset of $\mathbb{F}^p$. In the case considered here, $p=2$.

The following theorem is just like an earlier one for functions of one variable.

Theorem 9.9.1 *Let f be increasing and let g be continuous on $[a,b]$. Then there exists $c\in[a,b]$ such that $\int_a^b g\,df=g(c)(f(b)-f(c))$.*

Proof: If f is constant, there is nothing to prove so assume $f(b)>f(a)$. Let $M\equiv\max\{g(x):x\in[a,b]\},m\equiv\min\{g(x):x\in[a,b]\}$. Then in a Riemann sum for $\int_a^b g\,df$, if g is replaced by M, the resulting Riemann sum will increase and if it is replaced with m, the resulting sum will decrease. Therefore,

$$m(f(b)-f(a))=\int_a^b m\,df\leq\int_a^b g\,df\leq\int_a^b M\,df=M(f(b)-f(a))$$

and so $m\leq\frac{1}{f(b)-f(a)}\int_a^b g\,df\leq M$. Therefore, by the intermediate value theorem, there is $c\in[a,b]$ such that $\frac{1}{f(b)-f(a)}\int_a^b g\,df=g(c)$. ■

Lemma 9.9.2 *Let $f:[a,b]\times[c,d]\to\mathbb{R}$ be continuous at every point so it is uniformly continuous. Let α,β be increasing on $[a,b],[c,d]$ respectively. Then*

$$x\to\int_c^d f(x,y)\,d\beta(y),\ y\to\int_a^b f(x,y)\,d\alpha(x)$$

are both continuous functions.

Proof: Consider the first. The other is exactly similar.

$$\left|\int_c^d f(x,y)\,d\beta(y)-\int_c^d f(\hat{x},y)\,d\beta(y)\right|=\left|\int_c^d (f(x,y)-f(\hat{x},y))\,d\beta(y)\right|$$
$$\leq\int_c^d |f(x,y)-f(\hat{x},y)|\,d\beta(y)$$

But by uniform continuity, if $|x-\hat{x}|$ is small enough, then $|f(x,y)-f(\hat{x},y)|<\varepsilon$ and so the integral in the above is no larger than $\varepsilon(\beta(d)-\beta(c))$. Since ε is arbitrary, this shows the claim. ■

Note that, since these are continuous functions, it follows from Theorem 9.3.7 that it makes perfect sense to write the iterated integrals

$$\int_a^b\int_c^d f(x,y)\,d\beta(y)\,d\alpha(x),\ \int_c^d\int_a^b f(x,y)\,d\alpha(x)\,d\beta(y)$$

Of course the burning question is whether these two numbers are equal. This is the next theorem.

Theorem 9.9.3 *Let $f:[a,b]\times[c,d]\to\mathbb{R}$ be continuous and let β,α be increasing functions on $[c,d],[a,b]$ respectively. Then*

$$\int_a^b\int_c^d f(x,y)\,d\beta(y)\,d\alpha(x)=\int_c^d\int_a^b f(x,y)\,d\alpha(x)\,d\beta(y)$$

Proof:

$$\int_a^b\int_c^d f(x,y)\,d\beta(y)\,d\alpha(x)=\sum_{i=1}^n\int_{x_{i-1}}^{x_i}\int_c^d f(x,y)\,d\beta(y)\,d\alpha(x)$$

$$=\sum_{i=1}^n\int_{x_{i-1}}^{x_i}\sum_{j=1}^m\int_{y_{j-1}}^{y_j} f(x,y)\,d\beta(y)\,d\alpha(x)=\sum_{i=1}^n\sum_{j=1}^m\int_{x_{i-1}}^{x_i}\int_{y_{j-1}}^{y_j} f(x,y)\,d\beta(y)\,d\alpha(x)$$

By the mean value theorem for integrals, Theorem 9.9.1, this is

$$\begin{aligned}&\sum_{i=1}^n\sum_{j=1}^m\int_{x_{i-1}}^{x_i}\left(\beta(y_j)-\beta\left(y_{j-1}\right)\right)f(x,t_j)\,d\alpha(x)\\ =\ &\sum_{i=1}^n\sum_{j=1}^m\left(\beta(y_j)-\beta\left(y_{j-1}\right)\right)(\alpha(x_i)-\alpha(x_{i-1}))\,f(s_i,t_j)\end{aligned}$$

Also, by the same reasoning,

$$\int_c^d\int_a^b f(x,y)\,d\alpha(x)\,d\beta(y)=\sum_{j=1}^m\sum_{i=1}^n\left(\beta(y_j)-\beta\left(y_{j-1}\right)\right)(\alpha(x_i)-\alpha(x_{i-1}))\,f\left(s_i',t_j'\right)$$

and now because of uniform continuity, it follows that if the partition points are close enough,

$$\left|f\left(s_j',t_j'\right)-f\left(s_j,t_j\right)\right|<\frac{\varepsilon}{\left(\beta\left(d\right)-\beta\left(c\right)\right)\left(\alpha\left(b\right)-\alpha\left(a\right)\right)}$$

and so $\left|\int_c^d\int_a^b f\left(x,y\right)d\alpha\left(x\right)d\beta\left(y\right)-\int_a^b\int_c^d f\left(x,y\right)d\beta\left(y\right)d\alpha\left(x\right)\right|<\varepsilon$. Since ε is arbitrary, this shows the two iterated integrals are equal. ■

The following is concerning a very important formula. First recall the arctan function. Restricting tan to $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$, this function is one to one and has an inverse function called arctan. Thus $\arctan\left(y\right)=x$ where $x\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ and $\tan x=y$. Then, using the theory of the derivative of inverse functions, it follows that arctan is differentiable and $\arctan'\left(y\right)y'\left(x\right)=1$ and so $\arctan'\left(y\right)=\frac{1}{\sec^2\left(x\right)}=\frac{1}{1+\tan^2\left(x\right)}=\frac{1}{1+y^2}$. Also, $\arctan\left(0\right)=0$ obviously because $\tan\left(0\right)=0$. Therefore, $\arctan\left(y\right)=\int_0^y\frac{1}{1+u^2}du$.

Incidentally, this nice formula can be used to obtain all the trig functions. Note that $\arctan\left(1\right)=\pi/4$ because, from the above development, $\tan\left(\pi/4\right)=1$.

Theorem 9.9.4 *The following holds.* $\lim_{x\to\infty}\int_0^x e^{-t^2}dt=\frac{\sqrt{\pi}}{2}$.

Proof: Using the theorems about the integral obtained earlier, in particular the fundamental theorem of calculus,

$$\begin{aligned}\frac{d}{dx}\left(\int_0^x e^{-t^2}dt\right)^2&=2\left(\int_0^x e^{-t^2}dt\right)e^{-x^2}=2x\left(\int_0^1 e^{-x^2u^2}du\right)e^{-x^2}\\&=2x\int_0^1 e^{-x^2\left(u^2+1\right)}du\end{aligned}$$

Then, integrating both sides and interchanging the order of integration with Fubini's theorem, Theorem 9.9.3,

$$\begin{aligned}\left(\int_0^x e^{-t^2}dt\right)^2&=\int_0^x 2v\int_0^1 e^{-v^2\left(u^2+1\right)}dudv=\int_0^1\int_0^x 2ve^{-v^2\left(u^2+1\right)}dvdu\\&=\int_0^1\frac{-e^{-v^2\left(u^2+1\right)}}{u^2+1}|_0^x du=\int_0^1\left(\frac{1}{u^2+1}-\frac{e^{-x^2\left(u^2+1\right)}}{u^2+1}\right)du\end{aligned}$$

Hence

$$\int_0^x e^{-t^2}dt=\sqrt{\int_0^1\left(\frac{1}{u^2+1}-\frac{e^{-x^2\left(u^2+1\right)}}{u^2+1}\right)du}$$

Now the integrand on the right converges uniformly to $\frac{1}{u^2+1}$ as $x\to\infty$ and so we can pass to a limit as $x\to\infty$ and obtain $\lim_{x\to\infty}\int_0^x e^{-t^2}dt=\sqrt{\int_0^1\frac{1}{u^2+1}du}=\frac{\sqrt{\pi}}{2}$. ■

9.10 Geometric Length of a Curve in $\mathbb{R}^p$

I think that the right way to consider length is in terms of one dimensional Hausdorff measure. However, this is not a topic for this book. In this section, I am using the Euclidean norm because this is the one which corresponds to the usual notion of distance.

Definition 9.10.1 *A set of points $\gamma^* \subseteq \mathbb{R}^p$ is an oriented piecewise smooth curve if there is an oriented interval $[a,b]$ and $\gamma^* \equiv \gamma([a,b])$, and there are intermediate points between a and $b, z_1, z_2, ..., z_n$ such that $(b-a)(z_k - z_{k-1}) > 0$ and the following hold:*

1. *γ is one to one on $[a,b)$ and γ is one to one on $(a,b]$*

2. *$\gamma = \gamma_k$ on (z_{k-1}, z_k) with γ_k' continuous on $[z_{k-1}, z_k]$, where $\gamma_k'(z_k), \gamma'(z_{k-1})$ is an appropriate one sided derivative.*

3. *$\gamma_k' \neq 0$ on (z_{k-1}, z_k).*

Here γ' is defined in the usual way, $\gamma(t+h) = \gamma(t) + \gamma'(t)h + o(h)$. See Problem 16 on Page 155. Then letting P be an ordered partition of $[a,b]$, $P = \{a = t_0, t_1, ..., t_n = b\}$ where $(b-a)(t_k - t_{k-1}) > 0$, and letting $L(P)$ denote the sum $\sum_{k=1}^n |\gamma(t_k) - \gamma(t_{k-1})|$, the length of γ^* denoted as L is defined as

$$\sup\{L(P) : P \text{ is an ordered partition of } [a,b]\}.$$

Note that this gives an intrinsic definition of length depending only on γ^* and not on the particular parametrization because it picks a particular order along the curve γ^* and expresses the length as the sup of the lengths of all polygonal approximations of this curve.

Proposition 9.10.2 *With the above definition of length, $L = \int_a^b |\gamma'(t)| \, dt$.*

Proof: Whenever considering P, one of these ordered partitions, there is no loss of generality in assuming that the intermediate points $z_1, z_2, ..., z_n$ are in P because $L(P)$ only gets larger when points are added in to P. I will tacitly assume this in all that follows. Let $f : [z_{k-1}, z_k] \times [z_{k-1}, z_k]$

$$f(s,t) \equiv \begin{cases} \frac{(\gamma(t) - \gamma(s)) - \gamma'(s)(t-s)}{t-s} \text{ if } t \neq s \\ 0 \text{ if } t = s \end{cases}$$

Then f is uniformly continuous due to continuity of γ' and compactness. Therefore there exists $\delta_k > 0$ such that if $|t-s| < \delta_k$, then $|(\gamma(t) - \gamma(s)) - \gamma'(s)(t-s)| < \frac{\varepsilon}{b-a}|t-s|$. Now let $\|P\| < \delta \equiv \min\{\delta_k, k = 1, 2, ..., n\}$ and always P includes the z_k. Then by the triangle inequality, for such P,

$$\left|\gamma'(t_{k-1})\right| |t_k - t_{k-1}| - \frac{\varepsilon}{b-a} |t_k - t_{k-1}| \leq |\gamma(t_k) - \gamma(t_{k-1})|$$

$$\leq \left|\gamma'(t_{k-1})\right| |t_k - t_{k-1}| + \frac{\varepsilon}{b-a} |t_k - t_{k-1}| \tag{9.11}$$

Thus, for $\|P\| < \delta$, $L(|\gamma'|, P) - \varepsilon \leq L(P)$. Recall also the upper sums get smaller when points are added and lower sums get larger. Therefore, there exists P with $\|P\| < \delta$ and $U(|\gamma'|, P) - L(|\gamma'|, P) < \varepsilon$. In particular, from the above inequality, $L(P) \leq \sum_P |\gamma'| + \varepsilon \leq \int_a^b |\gamma'| \, dx + 2\varepsilon$ so $L \leq \int_a^b |\gamma'| \, dx + 2\varepsilon$. Thus, there exists possibly another P, with the above holding and also $L - \varepsilon < L(P) \leq L$. Then, from 9.11,

$$\int_a^b \left|\gamma'\right| dx - 2\varepsilon \leq L\left(\left|\gamma'\right|, P\right) - \varepsilon \leq L(P) \leq L \leq \int_a^b \left|\gamma'\right| dx + 2\varepsilon$$

and so $L - \int_a^b |\gamma'| \, dx \in [-2\varepsilon, 2\varepsilon]$. Since ε is arbitrary, it follows that $L = \int_a^b |\gamma'| \, dx$. ■

There are exactly two directions of motion over γ^*. In tracing out γ^*, one can either let t go from a to b or from b to a and these are the only possibilities if γ is to be one to one. Indeed, if $\hat{\gamma}$ maps the interval to γ^* and is continuous and one to one, then $\hat{\gamma}^{-1} \circ \gamma$ is either strictly increasing or strictly decreasing by Lemma 6.4.3. Increasing means same direction and decreasing, the opposite direction.

9.11 Exercises

1. In the chapter, upper and lower sums were considered. Suppose g is an increasing function and you are considering upper and lower sums for approximating $\int_a^b f dg$. Show that when you add in a point to the partition, the upper sum which results is no larger but the lower sum is no smaller.

2. Let $f(x) = 1 + x^2$ for $x \in [-1, 3]$ and let $P = \left\{-1, -\frac{1}{3}, 0, \frac{1}{2}, 1, 2\right\}$. Find $U(f, P)$ and $L(f, P)$ for $F(x) = x$ and for $F(x) = x^3$.

3. Let $P = \left\{1, 1\frac{1}{4}, 1\frac{1}{2}, 1\frac{3}{4}, 2\right\}$ and $F(x) = x$. Find upper and lower sums for the function $f(x) = \frac{1}{x}$ using this partition. What does this tell you about $\ln(2)$?

4. If $f \in R([a,b], F)$ with $F(x) = x$ and f is changed at finitely many points, show the new function is also in $R([a,b], F)$. Is this still true for the general case where F is only assumed to be an increasing function? Explain.

5. In the case where $F(x) = x$, define a "left sum" as $\sum_{k=1}^n f(x_{k-1})(x_k - x_{k-1})$ and a "right sum", $\sum_{k=1}^n f(x_k)(x_k - x_{k-1})$. Also suppose that all partitions have the property that $x_k - x_{k-1}$ equals a constant, $(b-a)/n$ so the points in the partition are equally spaced, and define the integral to be the number these right and left sums get close to as n gets larger and larger. Show that for f given in 9.6, $\int_0^x f(t) \, dt = x$ if x is rational and $\int_0^x f(t) \, dt = 0$ if x is irrational. It turns out that the correct answer should always equal zero for that function, regardless of whether x is rational. This illustrates why this method of defining the integral in terms of left and right sums is total nonsense. Show that even though this is the case, it makes no difference if f is continuous.

6. The function $F(x) \equiv \lfloor x \rfloor$ gives the greatest integer less than or equal to x. Thus $F(1/2) = 0, F(5.67) = 5, F(5) = 5$, etc. If $F(x) = \lfloor x \rfloor$ as just described, find $\int_0^{10} x dF$. More generally, find $\int_0^n f(x) \, dF$ where f is a continuous function.

7. Suppose f is a bounded function on $[0,1]$ and for each $\varepsilon > 0$, $\int_\varepsilon^1 f(x) \, dx$ exists. Can you conclude $\int_0^1 f(x) \, dx$ exists?

8. A differentiable function f defined on $(0, \infty)$ satisfies $f(xy) = f(x) + f(y)$, $f'(1) = 1$. Find f and sketch its graph.

9. Does there exist a function which has two continuous derivatives but the third derivative fails to exist at any point? If so, give an example. If not, explain why.

10. Suppose f is a continuous function on $[a,b]$ and $\int_a^b f^2 dF = 0$ where F is a strictly increasing integrator function. Show that then $f(x) = 0$ for all x. If F is not strictly increasing, is the result still true?

11. Suppose f is a continuous function and $\int_a^b f(x) x^n dx = 0$ for $n = 0, 1, 2, 3 \cdots$. Show using Problem 10 that $f(x) = 0$ for all x. **Hint:** You might use the Weierstrass approximation theorem.

12. Here is a function:
$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$$
Show this function has a derivative at every point of $\mathbb{R}$. Does it make any sense to write $\int_0^1 f'(x) dx = f(1) - f(0) = f(1)$? Explain.

13. Let $f(x) = \begin{cases} \sin\left(\frac{1}{x}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$. Is f Riemann integrable with respect to the integrator on the interval $[0, 1]$?

14. Recall that for a power series, $\sum_{k=0}^{\infty} a_k (x - c)^k$ you could differentiate term by term on the interval of convergence. Show that if the radius of convergence of the above series is $r > 0$ and if $[a, b] \subseteq (c - r, c + r)$, then
$$\int_a^b \sum_{k=0}^{\infty} a_k (x - c)^k dx = a_0 (b - a) + \sum_{k=1}^{\infty} \frac{a_k}{k} (b - c)^{k+1} - \sum_{k=1}^{\infty} \frac{a_k}{k} (a - c)^{k+1}$$
In other words, you can integrate term by term.

15. Find $\sum_{k=1}^{\infty} \frac{2^{-k}}{k}$.

16. Let f be Riemann integrable on $[0, 1]$. Show directly that $x \to \int_0^x f(t) dt$ is continuous. **Hint:** It is always assumed that Riemann integrable functions are bounded.

17. Suppose f, g are two functions which are continuous with continuous derivatives on $[a, b]$. Show using the fundamental theorem of calculus and the product rule the integration by parts formula. Also explain why all the terms make sense.
$\int_a^b f'(t) g(t) dt = f(b) g(b) - f(a) g(a) - \int_a^b f(t) g'(t) dt$

18. Show $\frac{1}{1+x^2} = \sum_{k=0}^{n} (-1)^k x^{2k} + \frac{(-1)^{n+1} x^{2n+2}}{1+x^2}$.Now use this to find a series which converges to $\arctan(1) = \pi/4$. Recall $\arctan(x) = \int_0^x \frac{1}{1+t^2} dt$. For which values of x will your series converge? For which values of x does the above description of arctan in terms of an integral make sense? Does this help to show the inferiority of power series?

19. Define $F(x) \equiv \int_0^x \frac{1}{1+t^2} dt$. Of course $F(x) = \arctan(x)$ as mentioned above but just consider this function in terms of the integral. Sketch the graph of F using only its definition as an integral. Show there exists a constant M such that $-M \leq F(x) \leq M$. Next explain why $\lim_{x \to \infty} F(x)$ exists and show this limit equals $-\lim_{x \to -\infty} F(x)$.

20. In Problem 19 let the limit defined there be denoted by $\pi/2$ and define $T(x) \equiv F^{-1}(x)$ for $x \in (-\pi/2, \pi/2)$. Show $T'(x) = 1 + T(x)^2$ and $T(0) = 0$. As part of this, you must explain why $T'(x)$ exists. For $x \in [0, \pi/2]$ let $C(x) \equiv 1/\sqrt{1 + T(x)^2}$ with $C(\pi/2) = 0$ and on $[0, \pi/2]$, define $S(x)$ by $\sqrt{1 - C(x)^2}$. Show both $S(x)$ and $C(x)$ are differentiable on $[0, \pi/2]$ and satisfy $S'(x) = C(x)$ and $C'(x) = -S(x)$. Find

the appropriate way to define $S(x)$ and $C(x)$ on all of $[0,2\pi]$ in order that these functions will be $\sin(x)$ and $\cos(x)$ and then extend to make the result periodic of period 2π on all of $\mathbb{R}$. Note this is a way to define the trig. functions which is independent of plane geometry and also does not use power series. See the book by Hardy, [14] for this approach, if I remember right.

21. Show $\arcsin(x) = \int_0^x \frac{1}{\sqrt{1-t^2}} dt$. Now use the binomial theorem to find a power series for $\arcsin(x)$.

22. The initial value problem from ordinary differential equations is of the form $y' = f(y)$, $y(0) = y_0$. Suppose f is a continuous function of y. Show that a function $t \to y(t)$ solves the above initial value problem if and only if $y(t) = y_0 + \int_0^t f(y(s))\, ds$.

23. Let $p, q > 1$ and satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Consider the function $x = t^{p-1}$. Then solving for t, you get $t = x^{1/(p-1)} = x^{q-1}$. Explain this. Now let $a, b \geq 0$. Sketch a picture to show why $\int_0^b x^{q-1} dx + \int_0^a t^{p-1} dt \geq ab$. Now do the integrals to obtain a very important inequality $\frac{b^q}{q} + \frac{a^p}{p} \geq ab$. When will equality hold in this inequality?

24. Suppose f, g are two Riemann Stieltjes integrable functions on $[a,b]$ with respect to F, an increasing function. Verify Holder's inequality.

$$\int_a^b |f|\,|g|\, dF \leq \left(\int_a^b |f|^p\, dF\right)^{1/p} \left(\int_a^b |g|^q\, dF\right)^{1/q}, \frac{1}{p} + \frac{1}{q} = 1, p > 1$$

Hint: Do the following. Let

$$A = \left(\int_a^b |f|^p\, dF\right)^{1/p}, B = \left(\int_a^b |g|^q\, dF\right)^{1/q}.$$

Then let $a = \frac{|f|}{A}, b = \frac{|g|}{B}$ and use the wonderful inequality of Problem 23.

25. Let $F(x) = \int_{x^2}^{x^3} \frac{t^5+7}{t^7+87t^6+1} dt$. Find $F'(x)$.

26. Let $F(x) = \int_2^x \frac{1}{1+t^4} dt$. Sketch a graph of F and explain why it looks the way it does.

27. Let a and b be positive numbers and consider the function

$$F(x) = \int_0^{ax} \frac{1}{a^2+t^2} dt + \int_b^{a/x} \frac{1}{a^2+t^2} dt.$$

Show that F is a constant.

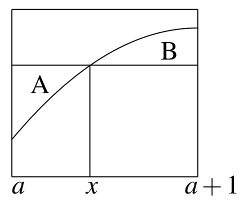
28. Solve the following initial value problem from ordinary differential equations which is to find a function y such that

$$y'(x) = \frac{x^4 + 2x^3 + 4x^2 + 3x + 2}{x^3 + x^2 + x + 1}, \; y(0) = 2.$$

29. If $F, G \in \int f(x)\, dx$ for all $x \in \mathbb{R}$, show $F(x) = G(x) + C$ for some constant, C. Use this to give a different proof of the fundamental theorem of calculus which has for its conclusion $\int_a^b f(t)\, dt = G(b) - G(a)$ where $G'(x) = f(x)$.

30. Suppose f is continuous on $[a,b]$. Show there exists $c \in (a,b)$ such that
$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx.$$
Hint: You might consider the function $F(x) \equiv \int_a^x f(t)\,dt$ and use the mean value theorem for derivatives and the fundamental theorem of calculus. In a sense, this was done in the chapter, but this one is more specific and note that here $c \in (a,b)$, the open interval.

31. Use the mean value theorem for integrals, Theorem 9.1.5 or the above problem to conclude that $\int_a^{a+1} \ln(t)\,dt = \ln(x) \leq \ln\left(a+\frac{1}{2}\right)$ for some $x \in (a,a+1)$. **Hint:** Consider the shape of the graph of $\ln(x)$ in the following picture. Explain why if x is the special value between a and $a+1$, then the area of A is equal to area of B. Why should $x < a+\frac{1}{2}$?



Now use this to obtain the inequality 9.9.

32. Suppose f and g are continuous functions on $[a,b]$ and that $g(x) \neq 0$ on (a,b). Show there exists $c \in (a,b)$ such that $f(c)\int_a^b g(x)\,dx = \int_a^b f(x)\,g(x)\,dx$. **Hint:** Define $F(x) \equiv \int_a^x f(t)\,g(t)\,dt$ and let $G(x) \equiv \int_a^x g(t)\,dt$. Then use the Cauchy mean value theorem on these two functions.

33. Consider the function
$$f(x) \equiv \begin{cases} \sin\left(\frac{1}{x}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}.$$
Is f Riemann integrable on $[0,1]$? Explain why or why not.

34. The Riemann integral is only defined for bounded functions on bounded intervals. When f is Riemann integrable on $[a,R]$ for each $R > a$ define an "improper" integral as follows. $\int_a^\infty f(t)\,dt \equiv \lim_{R\to\infty}\int_a^R f(t)\,dt$ whenever this limit exists. Show $\int_0^\infty \frac{\sin x}{x}dx$ exists. Here the integrand is defined to equal 1 when $x = 0$, not that this matters.

35. Show $\int_0^\infty \sin\left(t^2\right)dt$ exists.

36. The most important of all differential equations is the first order linear equation, $y' + p(t)\,y = f(t)$ where p, f are continuous. Show the solution to the initial value problem consisting of this equation and the initial condition, $y(a) = y_a$ is given by $y(t) = e^{-P(t)}y_a + e^{-P(t)}\int_a^t e^{P(s)}f(s)\,ds$, where $P(t) = \int_a^t p(s)\,ds$. Give conditions under which everything is correct. **Hint:** You use the integrating factor approach. Multiply both sides by $e^{P(t)}$, verify the left side equals $\frac{d}{dt}\left(e^{P(t)}y(t)\right)$,and then take the integral, $\int_a^t$ of both sides.

37. Suppose f is a continuous function which is not equal to zero on $[0,b]$. Show that
$$\int_0^b \frac{f(x)}{f(x)+f(b-x)}dx = \frac{b}{2}.$$
Hint: First change the variables to obtain the integral equals
$$\int_{-b/2}^{b/2} \frac{f(y+b/2)}{f(y+b/2)+f(b/2-y)}dy$$
Next show by another change of variables that this integral equals
$$\int_{-b/2}^{b/2} \frac{f(b/2-y)}{f(y+b/2)+f(b/2-y)}dy.$$
Thus the sum of these equals b.

38. Let there be three equally spaced points, $x_{i-1}, x_{i-1}+h \equiv x_i$, and $x_i + 2h \equiv x_{i+1}$. Suppose also a function f, has the value f_{i-1} at x, f_i at $x+h$, and f_{i+1} at $x+2h$. Then consider
$$\begin{aligned} g_i(x) \quad \equiv \quad & \frac{f_{i-1}}{2h^2}(x-x_i)(x-x_{i+1}) - \frac{f_i}{h^2}(x-x_{i-1})(x-x_{i+1}) \\ & + \frac{f_{i+1}}{2h^2}(x-x_{i-1})(x-x_i). \end{aligned}$$
Check that this is a second degree polynomial which equals the values f_{i-1}, f_i, and f_{i+1} at the points x_{i-1}, x_i, and x_{i+1} respectively. The function g_i is an approximation to the function f on the interval $[x_{i-1}, x_{i+1}]$. Also,
$$\int_{x_{i-1}}^{x_{i+1}} g_i(x)\, dx$$
is an approximation to $\int_{x_{i-1}}^{x_{i+1}} f(x)\, dx$. Show $\int_{x_{i-1}}^{x_{i+1}} g_i(x)\, dx$ equals $\frac{hf_{i-1}}{3} + \frac{hf_i 4}{3} + \frac{hf_{i+1}}{3}$. Now suppose n is even and $\{x_0, x_1, \cdots, x_n\}$ is a partition of the interval, $[a,b]$ and the values of a function f defined on this interval are $f_i = f(x_i)$. Adding these approximations for the integral of f on the succession of intervals,
$$[x_0, x_2], [x_2, x_4], \cdots, [x_{n-2}, x_n],$$
show that an approximation to $\int_a^b f(x)\, dx$ is
$$\frac{h}{3}[f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_2 + \cdots + 4f_{n-1} + f_n].$$
This is called Simpson's rule. Use Simpson's rule to compute an approximation to $\int_1^2 \frac{1}{t} dt$ letting $n = 4$.

39. Suppose $x_0 \in (a,b)$ and that f is a function which has $n+1$ continuous derivatives on this interval. Consider the following.
$$\begin{aligned} f(x) &= f(x_0) + \int_{x_0}^x f'(t)\, dt \\ &= f(x_0) + (t-x) f'(t)|_{x_0}^x + \int_{x_0}^x (x-t) f''(t)\, dt \\ &= f(x_0) + f'(x_0)(x-x_0) + \int_{x_0}^x (x-t) f''(t)\, dt. \end{aligned}$$

Explain the above steps and continue the process to eventually obtain Taylor's formula,

$$f(x)=f(x_0)+\sum_{k=1}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{1}{n!}\int_{x_0}^{x}(x-t)^n f^{(n+1)}(t)\,dt$$

where $n!\equiv n(n-1)\cdots 3\cdot 2\cdot 1$ if $n\geq 1$ and $0!\equiv 1$.

40. In the above Taylor's formula, use Problem 32 on Page 222 to obtain the existence of some z between x_0 and x such that

$$f(x)=f(x_0)+\sum_{k=1}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{f^{(n+1)}(z)}{(n+1)!}(x-x_0)^{n+1}.$$

Hint: You might consider two cases, the case when $x>x_0$ and the case when $x<x_0$.

41. There is a general procedure for constructing methods of approximate integration like the trapezoid rule and Simpson's rule. Consider $[0,1]$ and divide this interval into n pieces using a uniform partition, $\{x_0,\cdots,x_n\}$ where $x_i-x_{i-1}=1/n$ for each i. The approximate integration scheme for a function f, will be of the form

$$\left(\frac{1}{n}\right)\sum_{i=0}^{n}c_i f_i\approx\int_0^1 f(x)\,dx$$

where $f_i=f(x_i)$ and the constants, c_i are chosen in such a way that the above sum gives the exact answer for $\int_0^1 f(x)\,dx$ where $f(x)=1,x,x^2,\cdots,x^n$. When this has been done, change variables to write

$$\begin{aligned}\int_a^b f(y)\,dy &= (b-a)\int_0^1 f(a+(b-a)x)\,dx\\ &\approx \frac{b-a}{n}\sum_{i=1}^{n}c_i f\left(a+(b-a)\left(\frac{i}{n}\right)\right)=\frac{b-a}{n}\sum_{i=1}^{n}c_i f_i\end{aligned}$$

where $f_i=f\left(a+(b-a)\left(\frac{i}{n}\right)\right)$. Consider the case where $n=1$. It is necessary to find constants c_0 and c_1 such that

$$c_0+c_1=1=\int_0^1 1\,dx,\ 0c_0+c_1=1/2=\int_0^1 x\,dx.$$

Show that $c_0=c_1=1/2$, and that this yields the trapezoid rule. Next take $n=2$ and show the above procedure yields Simpson's rule. Show also that if this integration scheme is applied to any polynomial of degree 3 the result will be exact. That is,

$$\frac{1}{2}\left(\frac{1}{3}f_0+\frac{4}{3}f_1+\frac{1}{3}f_2\right)=\int_0^1 f(x)\,dx$$

whenever $f(x)$ is a polynomial of degree three. Show that if f_i are the values of f at $a,\frac{a+b}{2}$, and b with $f_1=f\left(\frac{a+b}{2}\right)$, it follows that the above formula gives $\int_a^b f(x)\,dx$ exactly whenever f is a polynomial of degree three. Obtain an integration scheme for $n=3$.

42. Let f have four continuous derivatives on $[x_{i-1}, x_{i+1}]$ where $x_{i+1} = x_{i-1} + 2h$ and $x_i = x_{i-1} + h$. Show using Problem 40, there exists a polynomial of degree three, $p_3(x)$, such that
$$f(x) = p_3(x) + \frac{1}{4!} f^{(4)}(\xi)(x - x_i)^4$$
Now use Problem 41 and Problem 38 to conclude
$$\left| \int_{x_{i-1}}^{x_{i+1}} f(x)\, dx - \left(\frac{hf_{i-1}}{3} + \frac{hf_i 4}{3} + \frac{hf_{i+1}}{3} \right) \right| < \frac{M}{4!} \frac{2h^5}{5},$$
where M satisfies, $M \geq \max\left\{ \left| f^{(4)}(t) \right| : t \in [x_{i-1}, x_i] \right\}$. Now let $S(a, b, f, 2m)$ denote the approximation to $\int_a^b f(x)\, dx$ obtained from Simpson's rule using $2m$ equally spaced points. Show
$$\left| \int_a^b f(x)\, dx - S(a, b, f, 2m) \right| < \frac{M}{1920} (b - a)^5 \frac{1}{m^4}$$
where $M \geq \max\left\{ \left| f^{(4)}(t) \right| : t \in [a, b] \right\}$. Better estimates are available in numerical analysis books but these also have the error in the form $C\left(1/m^4\right)$.

43. A **regular Sturm Liouville problem** involves the differential equation, for an unknown function of x which is denoted here by y,
$$\left(p(x) y'\right)' + (\lambda q(x) + r(x)) y = 0,\ x \in [a, b]$$
and it is assumed that $p(t), q(t) > 0$ for any t along with boundary conditions,
$$C_1 y(a) + C_2 y'(a) = 0,\ C_3 y(b) + C_4 y'(b) = 0$$
where $C_1^2 + C_2^2 > 0$, and $C_3^2 + C_4^2 > 0$. There is an immense theory connected to these important problems. The constant, λ is called an eigenvalue. Show that if y is a solution to the above problem corresponding to$\lambda = \lambda_1$ and if z is a solution corresponding to $\lambda = \lambda_2 \neq \lambda_1$, then
$$\int_a^b q(x) y(x) z(x)\, dx = 0. \tag{9.12}$$
Hint: Do something like this:
$$\begin{aligned} \left(p(x) y'\right)' z + (\lambda_1 q(x) + r(x)) yz &= 0, \\ \left(p(x) z'\right)' y + (\lambda_2 q(x) + r(x)) zy &= 0. \end{aligned}$$
Now subtract and either use integration by parts or show
$$\left(p(x) y'\right)' z - \left(p(x) z'\right)' y = \left(\left(p(x) y'\right) z - \left(p(x) z'\right) y\right)'$$
and then integrate. From the boundary conditions, show $y'(a) z(a) - z'(a) y(a) = 0$ and $y'(b) z(b) - z'(b) y(b) = 0$. The formula, 9.12 is called an orthogonality relation and it makes possible an expansion in terms of certain functions called eigenfunctions.

44. Letting $[a,b] = [-\pi,\pi]$, consider an example of a regular Sturm Liouville problem which is of the form $y'' + \lambda y = 0, y(-\pi) = 0, y(\pi) = 0$. Show that if $\lambda = n^2$ and $y_n(x) = \sin(nx)$ for n a positive integer, then y_n is a solution to this regular Sturm Liouville problem. In this case, $q(x) = 1$ and so from Problem 43, it must be the case that $\int_{-\pi}^{\pi} \sin(nx)\sin(mx)\,dx = 0$ if $n \neq m$. Show directly using integration by parts that the above equation is true.

45. Suppose g is increasing and f is continuous and of bounded variation. By the theorems in the chapter, $\int_a^b f dg$ exists and so $\int_a^b g df$ exists also. See Theorem 9.4.1. $g \in R([a,b], f)$. Show there exists $c \in [a,b]$ such that

$$\int_a^b g df = g(a)\int_a^c df + g(b)\int_c^b df$$

This is called the second mean value theorem for integrals. **Hint:** Use integration by parts.

$$\int_a^b g df = -\int_a^b f dg + f(b)g(b) - f(a)g(a)$$

Now use the first mean value theorem, the result of Theorem 9.9.1 to substitute something for $\int_a^b f dg$ and then simplify.

46. Generalize the result of Theorem 9.9.3 to the situation where α and β are only of bounded variation.

47. This problem is in Apostol [2]. Explain why whenever f is continuous on $[a,b]$

$$\lim_{n\to\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\left(\frac{b-a}{n}\right)\right) = \int_a^b f dx.$$

Apply this to $f(x) = \frac{1}{1+x^2}$ on the interval $[0,1]$ to obtain the very interesting formula $\frac{\pi}{4} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{n}{n^2+k^2}$.

48. Suppose $f : [a,b] \times (c,d) \to \mathbb{R}$ is continuous. Recall the meaning of the partial derivative from calculus,

$$\frac{\partial f}{\partial x}(t,x) \equiv \lim_{h\to 0} \frac{f(t,x+h) - f(t,x)}{h}$$

Suppose also $\frac{\partial f}{\partial x}(t,x)$ exists and for some K independent of t,

$$\left|\frac{\partial f}{\partial x}(t,z) - \frac{\partial f}{\partial x}(t,x)\right| < K|z-x|.$$

This last condition happens, for example if $\frac{\partial^2 f(t,x)}{\partial x^2}$ is uniformly bounded on $[a,b] \times (c,d)$. (Why?) Define $F(x) \equiv \int_a^b f(t,x)\,dt$. Take the difference quotient of F and show using the mean value theorem that $F'(x) = \int_a^b \frac{\partial f(t,x)}{\partial x} dt$. Is there a version of this result with dt replaced with $d\alpha$ where α is an increasing function? How about α a function of bounded variation?

49. I found this problem in Apostol's book [2]. This is a very important result and is obtained very simply by differentiating under an integral. Read it and fill in any missing details. Let $g(x) \equiv \int_0^1 \frac{e^{-x^2(1+t^2)}}{1+t^2}dt$ and $f(x) \equiv \left(\int_0^x e^{-t^2}dt\right)^2$. Note $\frac{\partial}{\partial x}\left(\frac{e^{-x^2(1+t^2)}}{1+t^2}\right) = -2xe^{-x^2(1+t^2)}$ and

$$\frac{\partial^2}{\partial x^2}\left(\frac{e^{-x^2(1+t^2)}}{1+t^2}\right) = -2e^{-x^2(1+t^2)} + 4x^2e^{-x^2(1+t^2)} + 4x^2e^{-x^2(1+t^2)}t^2$$

which is bounded for $t \in [0,1]$ and $x \in (-\infty,\infty)$. Explain why this is so. Also show the conditions of Problem 48 are satisfied so that

$$g'(x) = \int_0^1 \left(-2xe^{-x^2(1+t^2)}\right)dt.$$

Now use the chain rule and the fundamental theorem of calculus to find $f'(x)$. Then change the variable in the formula for $f'(x)$ to make it an integral from 0 to 1 and show $f'(x)+g'(x) = 0$.Now this shows $f(x)+g(x)$ is a constant. Show the constant is $\pi/4$ by assigning $x=0$. Next take a limit as $x \to \infty$ to obtain the following formula for the improper integral, $\int_0^\infty e^{-t^2}dt$, $\left(\int_0^\infty e^{-t^2}dt\right)^2 = \pi/4$. In passing to the limit in the integral for g as $x \to \infty$ you need to justify why that integral converges to 0. To do this, argue the integrand converges uniformly to 0 for $t \in [0,1]$ and then explain why this gives convergence of the integral. Thus $\int_0^\infty e^{-t^2}dt = \sqrt{\pi}/2$.

50. To show you the power of Stirling's formula, find whether the series $\sum_{n=1}^\infty \frac{n!e^n}{n^n}$ converges. The ratio test falls flat but you can try it if you like. Now explain why, if n is large enough, $n! \geq \frac{1}{2}\sqrt{\pi}\sqrt{2}e^{-n}n^{n+(1/2)} \equiv c\sqrt{2}e^{-n}n^{n+(1/2)}$.

51. The Riemann integral $\int_a^b f(x)\,dt$ for integrator function $F(t) = t$ is only defined if f is bounded. This problem discusses why this is the case. From the definition of the Riemann integral, there is a $\delta > 0$ such that if $\|P\| < \delta$, then the Riemann sum $S_P(f)$ must satisfy $\left|S_P(f) - \int_a^b f\,dt\right| < 1$. Pick such a partition $P = \{a = x_0 < \cdots < x_n = b\}$ and say $S_P(f) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1})$. Suppose that f is unbounded on $[x_{j-1}, x_j]$. Then you can modify the points t_i, keeping all the same except for $t_j \in [x_{j-1}, x_j]$ and let this one be $\hat{t}_j$ where this is chosen so large that

$$|f(\hat{t}_j)(x_j - x_{j-1})| - \left(\left|\sum_{i\neq j} f(t_i)(x_i - x_{i-1})\right| + \left|\int_a^b f\,dt\right| + 1\right) > 100$$

Show this is a contradiction. Hence f must be bounded.

52. Does the above conclusion that f is bounded hold in case of an arbitrary Riemann Stieltjes integral assuming the integrator function F is strictly increasing?

53. Use Theorem 9.9.1 and Lemma 9.9.2 to justify the following argument. Let f be continuous on $[a,b]\times[c,d]$. Let

$$F(x) \equiv \int_a^x \int_c^d f(t,y)\,dy\,dt - \int_c^d \int_a^x f(t,y)\,dt\,dy.$$

Then F is continuous on $[a,b]$. Also $F(a)=0$ and

$$F'(x)=\int_c^d f(x,y)\,dy-\int_c^d f(x,y)\,dy=0$$

and so $F(b)=0$ so Fubini's theorem holds.

54. Let $\{a_n\}$ be a sequence of positive numbers such that $\lim_{n\to\infty} na_n=0$ and for all $n\in\mathbb{N}$, $na_n\geq(n+1)a_{n+1}$. 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 partial summation formula on $\sum ka_k\frac{\sin kx}{k}$ and show the partial sums of $\sum\frac{\sin kx}{k}$ are bounded independent of x. To do this, you might argue the maximum value of the partial sums of this series occur when $\sum_{k=1}^{n}\cos kx=0$. Sum this series by considering the real part of the geometric series, $\sum_{k=1}^{q}\left(e^{ix}\right)^k$ and then show the partial sums of $\sum\frac{\sin kx}{k}$ are Riemann sums for a certain finite integral.

55. The problem in Apostol's book mentioned in Problem 54 does not require na_n to be decreasing and is as follows. Let $\{a_k\}_{k=1}^{\infty}$ be a decreasing sequence of non-negative numbers which satisfies $\lim_{n\to\infty}na_n=0$. Then $\sum_{k=1}^{\infty}a_k\sin(kx)$ converges uniformly on $\mathbb{R}$. You can find this problem worked out completely in Jones [19]. Fill in the details. This is a very remarkable observation. It says for example that $\sum_{k=1}^{\infty}\frac{1}{k^{1+\ln(k)}}\sin(kx)$ converges uniformly.

    Always let p be so large that $(p-1)a_{p-1}<\varepsilon$. Also, note that $|\sin x|\leq|x|$ for all x and for $x\in(0,\pi/2)$, $\sin x\geq\frac{x}{2\pi}$. (You could just graph $\sin x-\frac{x}{2\pi}$ to verify this.) Also, we can assume all a_k are positive since there is nothing to show otherwise. Define $b(k)\equiv\sup\left\{ja_j: j\geq k\right\}$. Thus $k\to b(k)$ is decreasing and $b(k)\to 0$ and $b(k)/k\geq a_k$.

    Suppose $x<1/q$ so each $\sin(kx)>0$. Then

$$\left|\sum_{k=p}^{q}a_k\sin(kx)\right|\leq\sum_{k=p}^{q}\frac{b(k)}{k}\sin(kx)\leq\sum_{k=p}^{q}\frac{b(k)}{k}kx\leq b(p)(q-p)\frac{1}{q}\leq b(p) \tag{9.13}$$

    Next recall that

$$\begin{aligned}\sum_{k=1}^{n}\sin(kx) &= \frac{\cos\left(\frac{x}{2}\right)-\cos\left(\left(n+\frac{1}{2}\right)x\right)}{2\sin\left(\frac{x}{2}\right)}\equiv m_n(x),\\ |m_n(x)| &\leq n,\ |m_n(x)|\leq\frac{1}{\sin(x/2)}\text{ if }x\in(0,\pi).\end{aligned}$$

    This is from the process for finding the Dirichlet kernel. Then use the process of summation by parts to obtain in every case that

$$\begin{aligned}\left|\sum_{k=p}^{q}a_k\sin(kx)\right| &\leq \left|a_q m_q(x)-a_{p-1}m_{p-1}(x)\right|+\left|\sum_{k=p}^{q-1}m_k(x)(a_k-a_{k+1})\right|\\ &\leq 2\varepsilon+\sum_{k=p}^{q-1}|m_k(x)|(a_k-a_{k+1})\leq 2\varepsilon+\frac{1}{\sin(x/2)}(a_p-a_q) \end{aligned}\tag{9.14}$$

We will only consider $x \in (0,\pi)$ for the next part. Then for such x, It remains to consider $x \in (0,\pi)$ with $x \geq 1/q$. In this case, choose m such that $q > \frac{1}{x} \geq m \geq \frac{1}{x} - 1$. Thus $x < \frac{1}{m}, \frac{1}{m+1} < x$. Then from 9.14, 9.13,

$$\left|\sum_{k=p}^{q} a_k \sin(kx)\right| \leq \overbrace{\left|\sum_{k=p}^{m} a_k \sin(kx)\right|}^{\leq b(p)} + \left|\sum_{k=m+1}^{q} a_k \sin(kx)\right|$$

$$\begin{aligned}
&\leq b(p) + 2\varepsilon + \frac{1}{\sin(x/2)}(a_{m+1} - a_q) \leq b(p) + 2\varepsilon + \frac{4\pi}{x} a_{m+1} \\
&\leq b(p) + 2\varepsilon + \frac{4\pi}{x}\frac{b(m+1)}{m+1} \leq b(p) + 2\varepsilon + \frac{4\pi}{x} b(m+1)x \\
&\leq b(p) + 2\varepsilon + 4\pi b(p) \leq 2\varepsilon + 17b(p), \ \lim_{p\to\infty} b(p) = 0.
\end{aligned}$$

Since ε is arbitrary, this shows uniform convergence on $(0,\pi)$. Thus the series converges uniformly on $[-\pi,\pi]$ and hence it converges uniformly on $\mathbb{R}$. This series is an example of a Fourier series. Its uniform convergence is very significant.

56. Using only the definition of the integral in the 1700's that $\int_a^b f(t)\,dt = F(b) - F(a)$, show that if $f_n \to f$ uniformly for each f_n continuous, then $\int_a^b f(t)\,dt = \lim_{n\to\infty} \int_a^b f_n(t)\,dt$.

57. Suppose $S'' + S = 0, S(0) = 0, S'(0) = 1$ and $C'' + C = 0$ and $C(0) = 1, C'(0) = 0$. Recall that the power series for $\sin x$ and $\cos x$ respectively satisfy these initial value problems. Show directly from the initial value problems that $S' = C$ and $C' = -S$. Also show that $S^2 + C^2 = 1$ and that $S(t) = \sin t, C(t) = \cos t$ where $\cos t, \sin t$, have the usual geometric descriptions for t the radian measure. **Hint:** Show S' satisfies the same initial value problem as C and use uniqueness. Then show $-C'$ satisfies the same initial value problem as S.

58. Show $\ln'(t) = 1/t$ and that for $x > 0$, $\ln(x) = \int_1^x \frac{1}{t}\,dt$. Use this and the mean value theorem for integrals to show that $\ln\left(\frac{n+1}{n}\right) - \frac{1}{n+1} = (\ln(n+1) - \ln(n)) - \frac{1}{n+1} > 0$. Now show that $n \to \sum_{k=1}^{n} \frac{1}{k} - \ln(n)$ is a decreasing sequence bounded below by 0 so it must converge to some number γ. This is called Euler's constant. To show $\gamma > 0$, consider $\sum_{k=1}^{n-1} \frac{1}{k} - \ln(n)$ for $n \geq 3$. Verify this sequence is increasing and when $n = 3$ it is positive.

59. Suppose $u(t)$ is nonnegative and continuous for $t \in [0,T]$ and for some $K > 0, u(t) \leq u_0 + K\int_0^t u(s)\,ds$. Show that $u(t) \leq u_0 e^{Kt}$. This is called Gronwall's inequality. **Hint:** Fill in the details. Let $w(t) \equiv \int_0^t u(s)\,ds$ so $w'(t) - Kw(t) \leq u_0$. Now from the product rule and chain rule, $\frac{d}{dt}\left(e^{-Kt}w(t)\right) \leq u_0 e^{-Kt}$ and so

$$\begin{aligned}
w(t)e^{-Kt} &\leq \frac{-1}{K}u_0 e^{-Kt} + \frac{1}{K}u_0 \\
w(t) &\leq \frac{1}{K}u_0 e^{Kt} - \frac{1}{K}u_0
\end{aligned}$$

Therefore, $u(t) \leq u_0 + K\left(\frac{1}{K}u_0 e^{Kt} - \frac{1}{K}u_0\right) = u_0 e^{Kt}$

60. Let $\mathbf{f}:\mathbb{R}\times\mathbb{R}^p\rightarrow\mathbb{R}^p$ be continuous and bounded and let $\mathbf{x}_0\in\mathbb{R}^p$. If $\mathbf{x}:[0,T]\rightarrow\mathbb{R}^p$ 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 $h\rightarrow 0$ such that $\mathbf{x}_h\rightarrow\mathbf{x}$ in $C([0,T];\mathbb{R}^p)$. Next argue
$$\mathbf{x}(t)=\mathbf{x}_0+\int_0^t\mathbf{f}(s,\mathbf{x}(s))\,ds$$
and conclude the following theorem. If $\mathbf{f}:\mathbb{R}\times\mathbb{R}^n\rightarrow\mathbb{R}^n$ 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.
$$\mathbf{x}'=\mathbf{f}(t,\mathbf{x}),\ \ t\in[0,T],\ \mathbf{x}(0)=\mathbf{x}_0. \tag{9.15}$$
This is the Peano existence theorem for ordinary differential equations.

61. Show, using Gronwall's inequality of Problem 59 that in the above theorem, if
$$\|\mathbf{f}(t,\mathbf{x})-\mathbf{f}(t,\mathbf{y})\|\leq K\|\mathbf{x}-\mathbf{y}\|,$$
then there is only one solution to the initial value problem 9.15.

62. It you let B_r be the closed ball $\{\mathbf{x}:\|\mathbf{x}\|\leq r\}$ let
$$P_r\mathbf{x}=\begin{cases}\frac{\mathbf{x}}{\|\mathbf{x}\|}r \text{ if } \|\mathbf{x}\|>r\\ \mathbf{x} \text{ if } \|\mathbf{x}\|\leq r\end{cases}$$
Show that P_r is continuous as a map from $\mathbb{R}^p$ to $\mathbb{R}^p$.

63. Using the above problem, show using Problem 60 that there is a local solution to 9.15 valid for $t\in[0,T_0]$ for some $T_0\leq T$ if it is only assumed that $\mathbf{f}$ is continuous, with no assumption that it is bounded. The last four problems contain all that is typically left out in undergraduate differential equations courses which is also that which is of most importance.

64. Suppose f is Riemann integral on the interval $[a,b]$. The integrator function is just $g(t)=t$. Now let $h(u)\equiv f(x-u)$ for $u\in[x-b,x-a]$. Show h is Riemann integrable on this new interval. Do something similar for $h(u)\equiv f(x+u)$.

Chapter 10

Improper Integrals

In everything, it is assumed that f is Riemann integrable on finite intervals, usually piecewise continuous on finite intervals. Thus there is no issue about whether the Riemann integral of the function on a finite interval exists. In this chapter, the integrator function will be $g(x) = x$. Also I will write $a_n \downarrow a$ to mean that $\{a_n\}$ is decreasing and $\lim_{n\to\infty} a_n = a$. The symbol $b_n \uparrow b$ is defined similarly but here b_n is increasing and has limit b.

To begin with, assume the functions are real valued.

Definition 10.0.1 *Let $a \geq -\infty$ and $b \leq \infty$. A function f is improper Riemann integrable if it is Riemann integrable on all intervals $[\alpha,\beta] \subseteq (a,b)$ and if there is a number I such that whenever $a_n \downarrow a$, and $b_n \uparrow b$, it follows that $\lim_{n\to\infty} \int_{a_n}^{b_n} f(t)\,dt = I$. Then I will be denoted as $\int_a^b f(t)\,dt$. A function f is in $L^1(a,b)$ if $|f|$ is improper Riemann integrable.*

Proposition 10.0.2 *Let a,b be as in Definition 10.0.1 and let f be a function Riemann integrable on each $[\alpha,\beta] \subseteq (a,b)$. Then $f \in L^1(a,b)$ if and only if*

$$\sup\left\{\int_\alpha^\beta |f(t)|\,dt : [\alpha,\beta] \subseteq (a,b)\right\} \equiv I < \infty$$

and in this case, $\int_a^b |f(t)|\,dt = I$. Also, whenever f is in $L^1(a,b)$, it is improper Riemann integrable.

Proof: $\Rightarrow$ Say $f \in L^1(a,b)$. If $I = \infty$, then there would exist $\alpha_n \downarrow a, \beta_n \uparrow b$ and $\lim_{n\to\infty} \int_{\alpha_n}^{\beta_n} |f(t)|\,dt = \infty$ contrary to assumption.

$\Leftarrow$ Say $I < \infty$. Then pick α,β, with $[\alpha,\beta] \subseteq (a,b)$ and $I - \varepsilon < \int_\alpha^\beta |f(t)|\,dt < I$. If $a_n \downarrow a, b_n \uparrow b$, then for all n large enough, $\alpha > a_n > a$ and $\beta < b_n < b$ and so also, for all n large enough, $I - \varepsilon < \int_\alpha^\beta |f(t)|\,dt < \int_{a_n}^{b_n} |f(t)|\,dt < I$. Thus $\lim_{n\to\infty} \int_{a_n}^{b_n} |f(t)|\,dt = I$.

Consider $f \in L^1(a,b)$. Letting $a_n \downarrow a, b_n \uparrow b$,

$$\int_{a_n}^{b_n} f(t)\,dt = \int_{a_n}^{b_n} \frac{|f(t)| + f(t)}{2}\,dt - \int_{a_n}^{b_n} \frac{|f(t)| - f(t)}{2}\,dt$$

Now both of those two integrals on the right are increasing in n because the integrands are nonnegative and they are also bounded above by $\int_a^b |f(t)|\,dt$ and so they both converge. Therefore, $\lim_{n\to\infty} \int_{a_n}^{b_n} f(t)\,dt$ also exists. ■

The claims of the following example are easy and are left for you to verify.

Example 10.0.3 *If $f(x) = x^{-p}$, for $p \in (0,1)$, then $f \in L^1(0,1)$. If $f(x) = e^{-x}$, then $f \in L^1(-a,\infty)$ for any $a < \infty$. If $f(x) = 1/x^2$, then $f \in L^1(1,\infty)$. If $f(x) = e^{-t}t^{-p}$ where $p > 0$, then $f \in L^1(0,\infty)$.*

Note that it follows from the above Definition 10.0.1 that if $f \in L^1(-\infty,\infty)$, then if $(a,b) \subseteq (-\infty,\infty)$, then $f\mathscr{X}_{(a,b)} \in L^1((a,b))$. Here $\mathscr{X}_{(a,b)}(x)$ is 1 if $x \in (a,b)$ and 0 if $x \notin (a,b)$. Also clear is that if $f,g \in L^1(a,b)$, then for any scalars $\alpha,\beta, \alpha f + \beta g \in L^1(a,b)$ also. Using Problem 64 on Page 230 it is routine to show that if $f \in L^1(-\infty,\infty)$, then if $g(u) = f(x-u)$ or $f(x+u)$ it follows that $g \in L^1(-\infty,\infty)$ also. Also note that if a is finite and f is bounded near a by M then $\left|\int_a^{b_n} f(t)\,dt - \int_{a_n}^{b_n} f(t)\,dt\right| \leq M\,|a_n - a|$ so there is no harm in simply considering $\int_a^{b_n} f(t)\,dt$ in the definition.

Lemma 10.0.4 *Suppose $f \in L^1(a,b)$. Then for $g(t)$ any bounded continuous function, say $|g(t)| \leq M$, fg is improper Riemann integrable.*

Proof: There is no issue about the existence of $\int_\alpha^\beta |f(t)||g(t)|dt$ for $[\alpha,\beta] \subseteq (a,b)$ thanks to Theorem 9.3.13. Then

$$\sup\left\{\int_\alpha^\beta |f(t)||g(t)|\,dt : [\alpha,\beta]\right\} \leq \sup\left\{\int_\alpha^\beta |f(t)|M\,dt : [\alpha,\beta]\right\}$$

$$= M\sup\left\{\int_\alpha^\beta |f(t)|\,dt : [\alpha,\beta]\right\} = M\int_a^b |f(t)|\,dt$$

Since $|fg| \in L^1(a,b)$, it follows that fg is improper Riemann integrable by Proposition 10.0.2. ■

Note that, from the last claim about computing the improper integral, all the usual algebraic properties of the Riemann integral carry over to these improper integrals of functions in L^1. For example, the integral is linear.

Sometimes it is convenient to define $\lim_{R\to\infty}\int_{-R}^R f(t)\,dt$. This may exist even though $f(t)$ may not be improper integrable. Such limits are called the Cauchy principal value integrals.

Example 10.0.5 $\lim_{R\to\infty}\int_{-R}^R \frac{t}{1+t^2}dt = 0$ *but* $\lim_{n\to\infty}\int_{-R}^{R^2}\frac{t}{1+t^2}dt = \infty$. *Both* $-R \to -\infty$ *and* $R^2 \to \infty$.

This is left as an exercise. Note that if $\int_{-\infty}^\infty f(t)\,dt$ does exist, then you can find it as a Cauchy principal value integral. It is just that sometimes the Cauchy principal value integral exists even though the function is not improper Riemann integrable. In the above example, the function is not in L^1. Indeed, you should show that $\lim_{R\to\infty}\int_{-R}^R \frac{|t|}{1+t^2}dt = \infty$.

When functions have values in $\mathbb{C}$, there is no extra problem. You simply consider the real and imaginary parts. That is,

$$\int_a^b f(t)\,dt \equiv \int_a^b \operatorname{Re} f(t)\,dt + i\int_a^b \operatorname{Im} f(t)\,dt$$

and define a function to be improper Riemann integrable if and only if this is true of its real and imaginary parts.

Here a complex valued function is in $L^1(a,b)$ means both the real and imaginary parts are in $L^1(a,b)$. This is equivalent to saying $|f|$ is in $L^1(a,b)$ because

$$\max(|\operatorname{Re} f|, |\operatorname{Im} f|) \leq |f| \leq |\operatorname{Re} f| + |\operatorname{Im} f|$$

Also for $f \in L^1(a,b)$, I use $\left|\int_a^b f(t)\,dt\right| \leq \int_a^b |f(t)|\,dt$ whenever convenient. This is certainly true if f has real values. However, it is also true for complex valued f. Likely the easiest way to see this is to note that it is true for sums. Since these approximate the integrals, it will be true for the integrals also. You could also do the following. There exists $\theta \in \mathbb{C}$ such that $|\theta| = 1$ and

$$\theta\int_a^b f(t)\,dt = \left|\int_a^b f(t)\,dt\right|$$

You just note $\int_a^b f(t)\,dt \in \mathbb{C}$ makes sense if $f \in L^1$ and then such a θ exists. Then some short computations with intervals $[a_n, b_n] \subseteq (a,b)$ show that the left side is equal to $\int_a^b \theta f(t)\,dt$. Since this is real,

$$\left|\int_a^b f(t)\,dt\right| = \int_a^b \theta f(t)\,dt = \int_a^b \operatorname{Re}(\theta f(t))\,dt \leq \int_a^b |f(t)|\,dt.$$

10.1 The Dirichlet Integral

There is a very important improper integral involving $\sin(x)/x$. You can show with a little estimating that $x \to \sin(x)/x$ is not in $L^1(0,\infty)$. However, one can show that this function is improper Riemann integrable. The following lemma is on a very important improper integral known as the Dirichlet integral after Dirichlet who first used it. See Problem 34 on Page 222 or else Problem 5 on 247 for more hints on how to show this. Here the actual value of this integral is obtained along with its existence. First note that for $x > 0$

$$\begin{aligned}\int_r^\infty e^{-tx}dt &= \lim_{R\to\infty}\int_r^R e^{-tx}dt = \lim_{R\to\infty}\frac{-e^{-tx}}{x}\Big|_r^R \\ &= \lim_{R\to\infty}\left(\frac{-e^{-tR}}{x}+\frac{e^{-rx}}{x}\right) = \frac{e^{-rx}}{x}\end{aligned}$$

Lemma 10.1.1 *The following formula holds.*

$$\frac{\pi}{2} = \int_0^\infty \frac{\sin(x)}{x}dx$$

Here $x \to \frac{\sin x}{x}$ is improper Riemann integrable.

Proof: By $\lim_{x\to 0}\frac{\sin x}{x} = 1$, we can assume $\frac{\sin x}{x}$ is continuous on $[0,1]$. Now let $[a_n, b_n] \subseteq (0,\infty)$. Then

$$\int_{a_n}^{b_n}\frac{\sin x}{x}dx = \int_{a_n}^{b_n}\sin(x)\overbrace{\int_0^\infty e^{-tx}dt}^{=1/x}dx = \int_0^{b_n}\sin(x)\overbrace{\int_0^\infty e^{-tx}dt}^{=1/x}dx + e(n)$$

where $\lim_{n\to\infty} e(n) \equiv \lim_{n\to\infty}\left(-\int_0^{a_n}\frac{\sin x}{x}dx\right) = 0$. Thus

$$\begin{aligned}\int_{a_n}^{b_n}\frac{\sin x}{x}dx &= e(n) + \int_0^{b_n}\sin(x)\int_0^{b_n}e^{-tx}dtdx + \int_0^{b_n}\sin(x)\int_{b_n}^\infty e^{-tx}dtdx \\ &= e(n) + \int_0^{b_n}\sin(x)\int_0^{b_n}e^{-tx}dtdx + E(n) \qquad (10.1)\end{aligned}$$

where $|E(n)| \leq \int_0^{b_n}|\sin(x)|\frac{e^{-b_n x}}{x}dx$. Now

$$|E(n)| \leq \int_0^{b_n}\overbrace{\frac{|\sin x|}{x}}^{<1}e^{-b_n x}dx < \frac{1}{b_n}.$$

Now interchange the order of integration in 10.1 using the Fubini theorem presented earlier, Theorem 9.9.3.

$$\int_{a_n}^{b_n}\frac{\sin x}{x}dx = e(n) + E(n) + \int_0^{b_n}\int_0^{b_n}e^{-tx}\sin(x)\,dxdt$$

Some tedious integration by parts on the inside integral on the right gives

$$\begin{aligned}
\int_{a_n}^{b_n} \frac{\sin x}{x} dx &= e(n) + E(n) \\
&\quad + \int_0^{b_n} \left[\frac{1}{t^2+1} \left(\left(1 - (\cos b_n) e^{-b_n t} + t (\sin b_n) e^{-b_n t} \right) \right) \right] dt
\end{aligned}$$

$$\begin{aligned}
&= e(n) + E(n) + \int_0^{b_n} \frac{1}{1+t^2} dt - \int_0^{b_n} e^{-b_n t} \frac{(\cos b_n) + t (\sin b_n)}{1+t^2} dt \\
&= e(n) + E(n) + \int_0^{b_n} \frac{1}{1+t^2} dt - \int_0^{b_n} e^{-b_n t} \frac{1}{\sqrt{1+t^2}} \cos(b_n - \phi_t) dt
\end{aligned}$$

where ϕ_t is a phase shift. This is because

$$\left(\frac{1}{\sqrt{1+t^2}}, \frac{t}{\sqrt{1+t^2}} \right) = (\cos \phi_t, \sin \phi_t) \text{ some } \phi_t$$

so the formula for $\cos(b_n - \phi)$ is used. This last integral satisfies

$$\left| \int_0^{b_n} e^{-b_n t} \frac{1}{\sqrt{1+t^2}} \cos(b_n - \phi_t) dt \right| \leq \int_0^{b_n} e^{-b_n t} dt \leq \frac{1}{b_n}$$

Therefore, include it in $E(n)$ and it follows that

$$\int_{a_n}^{b_n} \frac{\sin x}{x} dx = e(n) + E(n) + \int_0^{b_n} \frac{1}{1+t^2} dt$$

where $\lim_{n\to\infty} (e(n) + E(n)) = 0$. Taking a limit,

$$\lim_{n\to\infty} \int_{a_n}^{b_n} \frac{\sin x}{x} dx = \lim_{n\to\infty} \left(e(n) + E(n) + \int_0^{b_n} \frac{1}{1+t^2} dt \right) = \pi/2 \blacksquare \tag{10.2}$$

A much shorter way to verify this identity is in the exercises but it depends on a theorem which has not been discussed yet and to use it, you need to know the existence of the Dirichlet integral which is obtained here as part of the argument.

For I an interval let

$$\mathscr{X}_I(t) \equiv \begin{cases} 1 \text{ if } t \in I \\ 0 \text{ if } t \notin I \end{cases}$$

Lemma 10.1.2 *Suppose f is Riemann integrable on $[a,b]$. Then for each $\varepsilon > 0$, there is a step function s which satisfies $|s(x)| \leq |f(x)|$ and*

$$\int_a^b |f(x) - s(x)| dx < \varepsilon$$

Also there exists a continuous function h which is 0 at a and b such that $|h| \leq |f|$ and

$$\int_a^b |f(x) - h(x)|^2 dx < \varepsilon$$

Proof: First suppose $f(x) \geq 0$. Then by Theorem 9.3.10, there is a lower sum $L(f,P)$ such that

$$\left|\int_a^b f\,dt - L(f,P)\right| \leq (U(f,P) - L(f,P)) < \varepsilon$$

Let s correspond to this lower sum. That is, if $L(f,P) = \sum_{k=1}^m m_k (x_k - x_{k-1})$, you let $s(x) \equiv \sum_{k=1}^m m_k \mathscr{X}_{I_k}(x)$ where $I_1 = [x_0, x_1], I_k = (x_{k-1}, x_k]$ for $k \geq 2$. Then

$$\int_a^b |f(x) - s(x)|\,dx = \int_a^b (f(x) - s(x))\,dx = \int_a^b f\,dt - L(f,P) < \varepsilon$$

In general, you break down f. $f(x) = \frac{|f(x)|+f(x)}{2} - \frac{|f(x)|-f(x)}{2} \equiv f_+ - f_-$ where both f_+ and f_- are nonnegative functions. These are both integrable thanks to Theorem 9.3.13. Then let $s_+ \leq f_+, s_- \leq f_-$ with $\int_a^b (f_+ - s_+)\,dx, \int_a^b (f_- - s_-)\,dx$ both less than $\varepsilon/2$. Then letting $s = s_+ - s_-$, it is also a step function and

$$\int_a^b |f - s|\,dx = \int_a^b |(f_+ - f_-) - (s_+ - s_-)|\,dx = \int_a^b |(f_+ - s_+) + (s_- - f_-)|\,dx$$

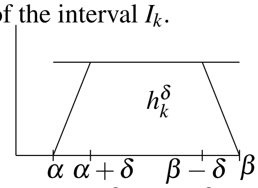
$$\leq \int_a^b |f_+ - s_+|\,dx + \int_a^b |f_- - s_-|\,dx < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

In addition to this,

$$|s(x)| = |s_+(x) - s_-(x)| \leq s_+(x) + s_-(x) \leq f_+(x) + f_-(x) = |f(x)|.$$

This takes care of the case where f is real. If it is complex valued, then to say it is Riemann integrable means simply that the real and imaginary parts are Riemann integrable. You apply what was just shown to these real and imaginary parts of f.

Now consider the last claim. First let $f \geq 0$ and let s be as chosen earlier with $s(x) \leq f(x)$ and $\int_a^b |f(x) - s(x)|\,dx \leq \eta$. Say $s(x) = \sum_{k=1}^n m_k \mathscr{X}_{I_k}(x)$ where I_k is an interval. Consider the following picture which approximates $\mathscr{X}_{I_k}$ with a continuous function called h_k which is zero at the ends of the interval I_k.



Then replace $s(x)$ with $\sum_{k=1}^n m_k h_k^\delta(x) \equiv h^\delta(x)$. Then h^δ is a continuous function which equals zero at a, b. Also, $\int_a^b \left|s(x) - h^\delta(x)\right| dx < \eta$ if δ is small enough. (Why?). Then let δ be this small and denote the resulting function by h. Thus for M an upper bound to $\sup\{|f(x) - h(x)| : x \in [a,b]\}$,

$$\begin{aligned}\int_a^b |f(x) - h(x)|^2\,dx &\leq M\int_a^b |f(x) - h(x)|\,dx \\ &\leq M\left(\int_a^b |f(x) - s(x)| + |s(x) - h(x)|\,dx\right) \leq 2M\eta\end{aligned}$$

Since η is arbitrary, this proves the second assertion in the case that $f \geq 0$. To get the result in general, do what was done in the first part. Let $f = f_+ - f_-$ where these two functions are nonnegative and Riemann integrable thanks to Theorem 9.3.13. Let $h_+ \leq f_+$ and $h_- \leq f_-$ go with f_+ and f_- respectively as in the above such that

$$\int_a^b |f_+(x) - h_+(x)|^2 dx, \int_a^b |f_-(x) - h_-(x)|^2 dx \leq \varepsilon/4$$

Then for $h = h_+ - h_-$, $\int_a^b |f-h|^2 dx = \int_a^b |f_+ - f_- - (h_+ - h_-)|^2 dx$

$$\begin{aligned} &\leq \int_a^b (|f_+ - h_+| + |f_- - h_-|)^2 dx \leq 2\int_a^b \left(|f_+ - h_+|^2 + |f_- - h_-|^2\right) dx \\ &\leq 4(\varepsilon/4) = \varepsilon \end{aligned}$$

and since ε is arbitrary, this yields continuous h, zero at end points such that $|h| \leq |f|$ as before. In case f has complex values, apply this that was just shown to the real and imaginary parts as was just done. ■

10.2 Convergence

The pointwise convergence of Fourier series was first successfully shown by Dirichlet in 1829. Here this important result, discussed later, is obtained from the very remarkable Riemann Lebesgue lemma.

Theorem 10.2.1 *The following hold*

1. $\int_0^\infty \frac{\sin u}{u} du = \frac{\pi}{2} = \int_0^\infty \frac{\sin(ru)}{u} du$ *for any* $r > 0$

2. $\lim_{r\to\infty} \int_\delta^\infty \frac{\sin(ru)}{u} du = 0$ *whenever* $\delta > 0$.

3. *If* $f \in L^1(\mathbb{R})$, *then* $\lim_{r\to\infty} \int_{-\infty}^\infty \sin(ru) f(u) du = 0$. *Note this implies that for any finite interval,* $[a,b]$,
$$\lim_{r\to\infty} \int_a^b \sin(ru) f(u) du = 0.$$
You just apply the first part to the function which is extended as 0 off $[a,b]$.

Proof: The first part 1. is Lemma 10.1.1.

Next consider 2. First note that $\int_0^\infty \frac{\sin(ru)}{u} du = \int_0^\infty \frac{\sin(t)}{t} r \frac{1}{r} dt = \int_0^\infty \frac{\sin(t)}{t} dt$.Now consider the truncated integral $\int_\delta^\infty \frac{\sin(ru)}{u} du$. It equals $\int_0^\infty \frac{\sin(ru)}{u} du - \int_0^\delta \frac{\sin(ru)}{u} du$ which can be seen from the definition of what the improper integral means. Also, you can change the variable. Let $ru = t$ so $rdu = dt$ and the above reduces to

$$\int_0^\infty \frac{\sin(t)}{t} r \frac{1}{r} dt - \int_0^{r\delta} \frac{\sin(t)}{t} dt = \int_\delta^\infty \frac{\sin(ru)}{u} du$$

Thus $\frac{\pi}{2} - \int_0^{r\delta} \frac{\sin(t)}{t} dt = \int_\delta^\infty \frac{\sin(ru)}{u} du$ and so $\lim_{r\to\infty} \int_\delta^\infty \frac{\sin(ru)}{u} du = 0$ from the first part.

Now consider the third claim, the Riemann Lebesgue lemma. Then for $f \in L^1$, let $f_{R,r}(t) \equiv \mathscr{X}_{[-r,R]}(t) f(t)$. Then for R, r large,

$$\int_{-\infty}^\infty |f(t) - f_{R,r}(t)| dt = \int_R^\infty |f(t)| dt + \int_{-\infty}^{-r} |f(t)| dt < \varepsilon \tag{10.3}$$

Now $f_{R,r}$ is Riemann integrable and so there is a step function $s(t) = \sum_{i=1}^{n} a_i \mathscr{X}_{I_i}(t)$ such that $|s(t)| \leq |f_{R,r}(t)|$ and

$$\int_{-r}^{R} |f_{R,r}(t) - s(t)|\, dt = \int_{-\infty}^{\infty} |f_{R,r}(t) - s(t)|\, dt < \varepsilon \tag{10.4}$$

This follows from Lemma 10.1.2. From 10.4 and 10.3, $\int_{-\infty}^{\infty} |s(t) - f(t)|\, dt < 2\varepsilon$. Now

$$\begin{aligned}\left|\int_{-\infty}^{\infty} f(t)\sin(rt)\, dt\right| &\leq \int_{-\infty}^{\infty} |(f(t) - s(t))\sin(rt)|\, dt + \left|\int_{-\infty}^{\infty} s(t)\sin(rt)\, dt\right| \\ &\leq 2\varepsilon + \left|\int_{-\infty}^{\infty} s(t)\sin(rt)\, dt\right| \end{aligned} \tag{10.5}$$

It remains to verify that $\lim_{r\to\infty} \int_{-\infty}^{\infty} s(t)\sin(rt)\, dt = 0$. Since $s(t)$ is a sum of scalars times $\mathscr{X}_I$ for I an interval, it suffices to verify that $\lim_{r\to\infty} \int_{-\infty}^{\infty} \mathscr{X}_{[a,b]}(t)\sin(rt)\, dt = 0$ However, this integral is $\int_a^b \sin(rt)\, dt = \frac{-1}{r}\cos(rb) + \frac{1}{r}\cos(ra)$ which clearly converges to 0 as $r \to \infty$. Therefore, for r large enough, 10.5 implies

$$\left|\int_{-\infty}^{\infty} f(t)\sin(rt)\, dt\right| < 3\varepsilon$$

Since ε is arbitrary, this shows that 3. holds. ■

Another proof is in Problem 12 on Page 249.

A simple repeat of the above argument shows the following slightly more general version of the Riemann Lebesgue lemma.

Corollary 10.2.2 *If $f \in L^1(\mathbb{R})$, then $\lim_{r\to\infty} \int_{-\infty}^{\infty} \sin(ru + c) f(u)\, du = 0$. Also,*

$$\lim_{r\to\infty} \int_{-\infty}^{\infty} \cos(ru + c) f(u)\, du = 0.$$

Proof: If you do the first part, which is in the exercises, the second claim comes right away from the observation that $\sin(x + \pi/2) = \cos(x)$. Thus

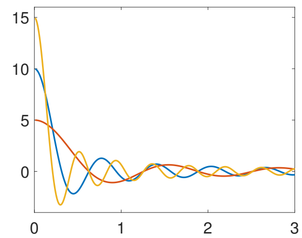
$$\cos(ru + c) = \sin(ru + c + \pi/2). \blacksquare$$

The case of most interest here is that of piecewise continuous functions.

Definition 10.2.3 *The following notation will be used assuming the limits exist.*

$$\lim_{u\to 0+} g(x + u) \equiv g(x+), \quad \lim_{u\to 0+} g(x - u) \equiv g(x-)$$

The convergence of Fourier series is discussed a little later. It will be based on the following theorem and a corollary which follow from the above Riemann Lebesgue lemma. Here is a graph of $\sin(nx)/x$ for a few values of n. Note that $\lim_{x\to 0+} \frac{\sin(nx)}{x} = n$ and that the graph shrinks and wriggles very fast as x increases for n large. This suggests that if f is smooth enough, then $\lim_{n\to\infty} \int_0^{\infty} f(x) \frac{\sin(nx)}{x} dx$ might be expected to depend on $f(0+)$. At least it is not unreasonable that this should happen. This is part of the following major theorem.



Lemma 10.2.4 *Let $f \in L^1((0,\infty))$ and suppose $|f(0) - f(u)| < Ku^\gamma$ for some $\gamma \in (0,1]$ whenever $0 < u \le \delta$. Then*

$$\lim_{r\to\infty} \frac{2}{\pi}\int_0^\infty \frac{\sin(ru)}{u} f(u)\,du = f(0)$$

Proof: From the theorem about the Dirichlet integral, Theorem 10.2.1,

$$\frac{2}{\pi}\int_0^\infty \frac{\sin(ru)}{u} f(u)\,du - f(0) = \frac{2}{\pi}\int_0^\delta \frac{\sin(ru)}{u}(f(u) - f(0))\,du \tag{10.6}$$

$$+\frac{2}{\pi}\int_\delta^\infty \frac{\sin(ru)}{u} f(u)\,du - \frac{2}{\pi}\int_\delta^\infty \frac{\sin(ru)}{u} f(0)\,du \tag{10.7}$$

Now $\frac{f(u)}{u} \in L^1(\delta,\infty)$ because $\left|\frac{f(u)}{u}\right| \le \frac{1}{\delta}|f(u)|$ and so, by the Riemann Lebesgue lemma, the first integral of 10.7 converges to 0 as $r \to \infty$. The second integral of 10.7 converges to 0 as $r \to \infty$ because of the second part of Theorem 10.2.1. Now consider the integral in 10.6. $s \to \frac{f(u)-f(0)}{u}$ is in $L^1([0,\delta])$ because $\left|\frac{f(u)-f(0)}{u}\right| \le u^{\gamma-1}$ which has finite integral since $\gamma > 0$. Therefore, $\lim_{r\to\infty} \frac{2}{\pi}\int_0^\delta \frac{\sin(ru)}{u}(f(u) - f(0))\,du = 0$ by the Riemann Lebesgue lemma. ■

Theorem 10.2.5 *Suppose that $g \in L^1(\mathbb{R})$ and that at some x, g is locally Holder continuous from the right and from the left. This means there exist constants $K, \delta > 0$ and $r \in (0,1]$ such that for $|x-y| < \delta$,*

$$|g(x+) - g(y)| < K|x-y|^r \tag{10.8}$$

for $y > x$ and

$$|g(x-) - g(y)| < K|x-y|^r \tag{10.9}$$

for $y < x$. Then

$$\lim_{r\to\infty} \frac{2}{\pi}\int_0^\infty \frac{\sin(ur)}{u}\left(\frac{g(x-u) + g(x+u)}{2}\right)du = \frac{g(x+) + g(x-)}{2}.$$

Proof: The function $u \to \frac{g(x-u)+g(x+u)}{2} \equiv f(u)$ is in $L^1(0,\infty)$ as noted earlier. Also for $f(0)$ defined as $\frac{g(x+)+g(x-)}{2}$, the conditions of Lemma 10.2.4 are obtained. Therefore, from that lemma,

$$\begin{aligned}\lim_{r\to\infty} \frac{2}{\pi}\int_0^\infty \frac{\sin(ur)}{u} f(u)\,du &= \lim_{r\to\infty} \frac{2}{\pi}\int_0^\infty \frac{\sin(ur)}{u}\left(\frac{g(x-u) + g(x+u)}{2}\right)du \\ &= \frac{g(x+) + g(x-)}{2}\end{aligned}$$

as claimed. ■

There is an other condition which will allow the same conclusion as the above condition. It is that g is of bounded variation on $[x-\delta, x+\delta]$ for some $\delta > 0$. This is called the Jordan condition whereas the more common assumption used above is the Dini condition. This Jordan condition implies that $u \to (g(x-u) - g(x-) + g(x+u) - g(x+))$ is of bounded variation on $[0, \delta]$.

First, here is a little review. If H is continuous, Theorem 9.9.1, the mean value theorem for integrals, implies that if g is increasing, then

$$\int_a^b H(x)\, dg = H(c)(g(b) - g(a))$$

for some $c \in [a,b]$. Suppose then that $H(x) \equiv \int_a^x h(t)\, dt$ where h is continuous and g is increasing on $[a,b]$.

Suppose $H(x) \equiv \int_a^x h(t)\, dt$ where h is continuous and g is increasing on $[a,b]$. Then by integration by parts, Theorem 9.4.1,

$$\int_a^b g dH + \int_a^b H dg = g(b) H(b)$$

From the first mean value theorem for integrals, there is $c \in [a,b]$ such that $\int_a^b H dg = H(c)(g(b) - g(a))$. Then

$$\begin{aligned}
\int_a^b g dH &= \int_a^b g h dt = -H(c)(g(b) - g(a)) + g(b) H(b) \\
&= g(b) \int_c^b h(t)\, dt + g(a) \int_a^c h(t)\, dt
\end{aligned}$$

This is sometimes called the second mean value theorem for integrals. Sufficient conditions are that h is continuous and g increasing. This is stated as the following lemma.

Lemma 10.2.6 *Let h be continuous and g increasing. Then there is $c \in [a,b]$ such that*

$$\int_a^b g(x) h(x)\, dx = g(b) \int_c^b h(x)\, dx + g(a) \int_a^c h(x)\, dx$$

The conclusion is exactly the same if $g(a)$ is replaced with $g(a+)$ with maybe a different $c \in [a,b]$.

Proof: The last claim follows from a repeat of the above argument using $\tilde{g}(x)$ defined as $g(x)$ for $x > a$ and $g(a+)$ when $x = a$. Such a change does nothing to the Riemann integral on the left in the above formula and $\tilde{g}$ is still increasing. Hence, for some $c \in [a,b]$,

$$\int_a^b g(x) h(x)\, dx = \int_a^b \tilde{g}(x) h(x)\, dx = g(b) \int_c^b h(x)\, dx + g(a+) \int_a^c h(x)\, dx \quad ■$$

Lemma 10.2.7 *Suppose g is of bounded variation on $[0, \delta]$, $\delta > 0$ and suppose $g \in L^1(0, a)$ where $\delta < a \leq \infty$. Then*

$$\lim_{r \to \infty} \frac{2}{\pi} \int_0^a g(t) \frac{\sin(rt)}{t}\, dt = g(0+)$$

Proof: Since every bounded variation function is the difference of two increasing functions, it suffices to assume that g is increasing on $[0,\delta]$ and so this will be assumed. Note that this also shows that $g(0+)$ exists. Recall the second mean value theorem of Problem 45 on Page 226 or the above Lemma 10.2.6 applied to $g(0)$ defined as $g(0+)$.

From the material on the Dirichlet integral, there exists C with $\left|\int_0^r \frac{\sin t}{t} dt\right| < C$ independent of r. Indeed, $\left|\int_0^r \frac{\sin t}{t} dt - \frac{\pi}{2}\right| < 1$ if r is sufficiently large. For r not this large, one has the integral of a continuous function.

Let $h \in (0,\delta)$ be such that if $t \leq h, |g(t) - g(0+)| < \frac{\varepsilon}{2C+1}$. Then split up the integral as follows.

$$\frac{2}{\pi}\int_0^a g(t)\frac{\sin(rt)}{t}dt = \overbrace{\frac{2}{\pi}\int_0^h (g(t) - g(0+))\frac{\sin(rt)}{t}dt}^{I_1} + g(0+)\overbrace{\frac{2}{\pi}\int_0^h \frac{\sin(rt)}{t}dt}^{I_2}$$
$$+\overbrace{\int_h^a g(t)\frac{\sin(rt)}{t}dt}^{I_3}$$

Use the second mean value theorem on I_1. It equals

$$(g(h) - g(0+))\frac{2}{\pi}\int_{c_r}^h \frac{\sin(rt)}{t}dt = (g(h) - g(0+))\frac{2}{\pi}\int_{rc_r}^{rh}\frac{\sin(u)}{u}du$$

the integral is $\left(\int_0^{hr}\frac{\sin u}{u}du - \int_0^{rc_r}\frac{\sin u}{u}du\right)$ and both terms are bounded by some constant C so the integral is bounded independent of large r by $2C$. Then

$$\left|(g(h) - g(0+))\frac{2}{\pi}\int_{rc_r}^{rh}\frac{\sin(u)}{u}du\right| \leq \frac{\varepsilon}{2C+1}2C < \varepsilon$$

Now consider $g(0+)I_2$. It equals $g(0+)\frac{2}{\pi}\int_0^{rh}\frac{\sin(u)}{u}du$ so its limit as $r \to \infty$ is $g(0+)$. It is just the Dirichlet integral again.

Finally, consider I_3. For $t \geq h$, $\frac{g(t)}{t}$ is in $L^1(h,a)$ and so, the 0 extension off $[h,\infty)$ is in $L^1([0,a))$. By the Riemann Lebesgue lemma, of Theorem 10.2.1, this integral I_3 converges to 0 as $r \to \infty$. ■

With this, here is a different version of Theorem 10.2.5.

Corollary 10.2.8 *Suppose that $g \in L^1(\mathbb{R})$ and that at some x, g is of finite total variation on $[x-\delta, x+\delta]$ for some $\delta > 0$. Then*

$$\lim_{r\to\infty}\frac{2}{\pi}\int_0^\infty \frac{\sin(ur)}{u}\left(\frac{g(x-u)+g(x+u)}{2}\right)du = \frac{g(x+)+g(x-)}{2}.$$

Proof: This follows from Lemma 10.2.7 applied to $u \to \frac{g(x-u)+g(x+u)}{2}$. ■

10.3 The Gamma Function

Recall the definition of an improper integral specialized to $(0,\infty)$. You let $a_n \downarrow 0, b_n \uparrow \infty$ and $\int_0^\infty f(t)\,dt = \lim_{n\to\infty}\int_{a_n}^{b_n} f(t)\,dt$.

Definition 10.3.1 *Whenever $\alpha > 0$, $\Gamma(\alpha) \equiv \int_0^\infty e^{-t}t^{\alpha-1}\,dt$.*

Lemma 10.3.2 *The improper integral $\int_0^\infty e^{-t}t^{\alpha-1}\,dt$ exists for each $\alpha>0$.*

Proof: Let
$$f(t)=\begin{cases} t^{\alpha-1} \text{ if } t\leq 1 \\ Ce^{-t/2} \text{ if } t>1\end{cases}$$
where C is chosen large enough that for $t>1, Ce^{-t/2}>e^{-t}t^{\alpha-1}$. Then obviously $f\in L^1(0,\infty)$ and $f(t)\geq e^{-t}t^{\alpha-1}$. Also, if $[\alpha,\beta]\subseteq(0,\infty)$,
$$\int_\alpha^\beta e^{-t}t^{\alpha-1}\,dt\leq\int_0^\infty f(t)\,dt<\infty$$
and so $t\to e^{-t}t^{\alpha-1}$ is in $L^1(0,\infty)$ so the improper integral exists as claimed thanks to Proposition 10.0.2. ■

This gamma function has some fundamental properties described in the following proposition. In case the improper integral exists, we can obviously compute it in the form $\lim_{\delta\to 0+}\int_\delta^{1/\delta} f(t)\,dt$ which is used in what follows. Thus also the usual algebraic properties of the Riemann integral are inherited by the improper integral.

Proposition 10.3.3 *For n a positive integer, $n!=\Gamma(n+1)$. In general, one has the following identity: $\Gamma(1)=1,\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$*

Proof: First of all, $\Gamma(1)=\lim_{\delta\to 0}\int_\delta^{\delta^{-1}} e^{-t}dt=\lim_{\delta\to 0}\left(e^{-\delta}-e^{-(\delta^{-1})}\right)=1$. Next, for $\alpha>0$,
$$\Gamma(\alpha+1)=\lim_{\delta\to 0}\int_\delta^{\delta^{-1}} e^{-t}t^\alpha dt=\lim_{\delta\to 0}\left[-e^{-t}t^\alpha|_\delta^{\delta^{-1}}+\alpha\int_\delta^{\delta^{-1}} e^{-t}t^{\alpha-1}dt\right]$$
$$=\lim_{\delta\to 0}\left(e^{-\delta}\delta^\alpha-e^{-(\delta^{-1})}\delta^{-\alpha}+\alpha\int_\delta^{\delta^{-1}} e^{-t}t^{\alpha-1}dt\right)=\alpha\Gamma(\alpha)$$
Now it is defined that $0!=1$ and so $\Gamma(1)=0!$. Suppose that $\Gamma(n+1)=n!$, what of $\Gamma(n+2)$? Is it $(n+1)!$? if so, then by induction, the proposition is established. From what was just shown, $\Gamma(n+2)=\Gamma(n+1)(n+1)=n!(n+1)=(n+1)!$ and so this proves the proposition. ■

The properties of the gamma function also allow for a fairly easy proof about differentiating under the integral in a Laplace transform. First is a definition.

Definition 10.3.4 *A function ϕ has exponential growth on $[0,\infty)$ if there are positive constants λ, C such that $|\phi(t)|\leq Ce^{\lambda t}$ for all $t\geq 0$.*

Theorem 10.3.5 *Let $f(s)=\int_0^\infty e^{-st}\phi(t)\,dt$ where $t\to\phi(t)e^{-st}$ is improper Riemann integrable for all s large enough and ϕ has exponential growth, $|\phi(t)|\leq Ce^{\lambda t}$. Then for s large enough, $f^{(k)}(s)$ exists and equals $\int_0^\infty(-t)^k e^{-st}\phi(t)\,dt$.*

Proof: Suppose true for some $k\geq 0$. By definition it is so for $k=0$. Then always assuming $s>\lambda, |h|<s-\lambda$, where $|\phi(t)|\leq Ce^{\lambda t},\lambda\geq 0$,
$$\frac{f^{(k)}(s+h)-f^{(k)}(s)}{h}=\int_0^\infty(-t)^k\frac{e^{-(s+h)t}-e^{-st}}{h}\phi(t)\,dt$$

$$= \int_0^\infty (-t)^k e^{-st} \left(\frac{e^{-ht} - 1}{h} \right) \phi(t)\, dt = \int_0^\infty (-t)^k e^{-st} \left((-t) e^{\theta(h,t)} \right) \phi(t)\, dt$$

where $\theta(h,t)$ is between $-ht$ and 0, this by the mean value theorem. Thus by mean value theorem again,

$$\left| \frac{f^{(k)}(s+h) - f^{(k)}(s)}{h} - \int_0^\infty (-t)^{k+1} e^{-st} \phi(t)\, dt \right|$$

$$\leq \int_0^\infty |t|^{k+1} C e^{\lambda t} e^{-st} \left| e^{\theta(h,t)} - 1 \right| dt \leq \int_0^\infty t^{k+1} C e^{\lambda t} e^{-st} e^{\alpha(h,t)} |ht|\, dt$$

$$\leq \int_0^\infty t^{k+2} C e^{\lambda t} e^{-st} |h| e^{t|h|} dt = C|h| \int_0^\infty t^{k+2} e^{-(s-(\lambda+|h|))t} dt$$

Let $u = (s - (\lambda + |h|))t, du = (s - (\lambda + |h|))\, dt$ and changing the variable, you see that the right side converges to 0 as $h \to 0$ so $f^{(k+1)}(t)$ has the correct form. This proves the theorem. ■

The function $f(s)$ just defined is called the Laplace transform of ϕ.

Incidentally, $f^{(k)}(s)$ exists for each k and $s \in \mathbb{C}$ with $\text{Re}(s) > \lambda$ by the same argument. This will be used later when the computation of inverse Laplace transforms is considered.

10.4 Laplace Transforms

It will be assumed here that $t \to f(t) e^{-st}$ is in $L^1(0,\infty)$ for all s large enough.

Definition 10.4.1 *We say that a function defined on $[0,\infty)$ has exponential growth if for some $\lambda \geq 0$, and $C > 0, |f(t)| \leq Ce^{\lambda t}$.*

Note that this condition is satisfied if $|f(t)| \leq a + be^{\lambda t}$. You simply pick $C > \max(a,b)$ and observe that $a + be^{\lambda t} \leq 2Ce^{\lambda t}$.

Proposition 10.4.2 *Let f have exponential growth and be continuous except for finitely many points in $[0,R]$ for each R. Then $\lim_{R\to\infty} \int_0^R f(t) e^{-st} dt \equiv \mathscr{L} f(s)$ exists for every $s > \lambda$ where $|f(t)| \leq e^{\lambda t}$. That limit is denoted as $\int_0^\infty f(t) e^{-st} dt$.*

Proof: It is clear that f has exponential growth implies $t \to f(t) e^{-st}$ is in $L^1(\mathbb{R})$ and so the improper integral above exists. ■

Certain properties are obvious. For example,

1. If a, b scalars and if g, f have exponential growth, then for all s large enough,

$$\mathscr{L}(af + bg)(s) = a\mathscr{L}(f)(s) + b\mathscr{L}(g)(s)$$

2. If $f'(t)$ exists and has exponential growth, and so does $f(t)$ then for s large enough,

$$\mathscr{L}\left(f'\right)(s) = -f(0) + s\mathscr{L}(f)(s)$$

One can also compute Laplace transforms of many standard functions without much difficulty. One of the most important properties of the Laplace transform is the convolution.

Definition 10.4.3 *Let f, g have exponential growth and be continuous except for finitely many points in each $[0,R]$. Then $f * g(t) \equiv \int_0^t f(t-u) g(u)\, du$.*

Observation 10.4.4 *$f * g = g * f$. This merely involves changing the variable. Let $v = t - u$.*

The following proposition will involve an assumption that the functions are continuous. This is not necessary but I have not developed the necessary machinery for Fubini's theorem to do this in full generality and an ad hoc approach sufficient to include discontinuous functions would be tedious. Therefore, here is an easy lemma.

Lemma 10.4.5 *Let f, g be continuous on $[0, R]$. Then*

$$\int_0^R \int_0^t f(t-u) g(u) \, du dt = \int_0^R \int_u^R f(t-u) g(u) \, dt du \tag{10.10}$$

Proof: First note the following. For $F(t) \equiv \int_0^t f(u) \, du$ and $G(t)$ defined similarly,

$$\int_0^R \int_0^t f(t) g(u) \, du dt = \int_0^R f(t) \int_0^t g(u) \, du dt =$$

$$F(t) \int_0^t g(u) \, du|_0^R - \int_0^R F(t) g(t) \, dt = F(R) G(R) - \int_0^R F(t) g(t) \, dt$$

$$\begin{aligned}
\int_0^R \int_u^R f(t) g(u) \, dt du &= \int_0^R g(u) \int_u^R f(t) \, dt du = \int_0^R g(u) (F(R) - F(u)) \, du \\
&= G(R) F(R) - \int_0^R F(u) g(u) \, du
\end{aligned}$$

If f is a polynomial, then $f(t-u) g(u)$ is a sum of terms of the form $\hat{f}(t) \hat{g}(u)$ and so 10.10 holds. Now by Weierstrass approximation theorem, there is a sequence of polynomials $\{f_n\}$ which converges uniformly to f on $[0, R]$. Hence, since 10.10 holds for each f_n replacing f it continues to hold in the limit. ■

Proposition 10.4.6 *Let f, g have exponential growth and be continuous. Then $f * g$ has the same properties. Also $\mathscr{L}(f * g)(s) = \mathscr{L}(f)(s) \mathscr{L}(g)(s)$ for all s large enough.*

Proof: Consider the second claim. Say $|f(t)| \leq Ce^{\lambda t}, |g(t)| \leq \hat{C}e^{\hat{\lambda} t}$. Letting $\mu \geq \max\left(\lambda, \hat{\lambda}\right)$,

$$\left| \int_0^t f(t-u) g(u) \, du \right| \leq \int_0^t C\hat{C} e^{\mu(t-u)} e^{\mu u} du \leq C\hat{C} t e^{\mu t} \leq C\hat{C} e^{2\max(\mu,1)t}$$

Now consider the claim about the convolution. Let $s > 2\max(\mu, 1) \equiv \gamma$. By Lemma 10.4.5,

$$\int_0^R e^{-st} \int_0^t f(t-u) g(u) \, du dt = \int_0^R \int_u^R e^{-st} f(t-u) g(u) \, dt du$$

$$= \int_0^R e^{-su} g(u) \int_u^R e^{-s(t-u)} f(t-u) \, dt du = \int_0^R e^{-su} g(u) \int_0^{R-u} e^{-sv} f(v) \, dv du$$

$$= \int_0^R e^{-su} g(u) \int_0^R e^{-sv} f(v) \, dv - \left(\int_{R-u}^R e^{-sv} f(v) \, dv \right) du$$

Therefore,

$$\int_0^R e^{-st}(f*g)(t)\,dt = \int_0^R e^{-su}g(u)\,du\int_0^R e^{-sv}f(v)\,dv - \int_0^R e^{-su}g(u)\int_{R-u}^R e^{-sv}f(v)\,dvdu \tag{10.11}$$

Now

$$\left|\int_0^R e^{-su}g(u)\int_{R-u}^R e^{-sv}f(v)\,dvdu\right| \leq \int_0^R \hat{C}e^{-(s-\gamma)u}\int_{R-u}^R Ce^{-(s-\gamma)v}dvdu$$

$$\equiv \int_0^R \hat{C}e^{-\lambda u}\int_{R-u}^R Ce^{-\lambda v}dvdu \leq C\hat{C}\int_0^R e^{-\lambda u}e^{-\lambda(R-u)}du = C\hat{C}e^{-\lambda R}R$$

which converges to 0 as $R\to\infty$. Therefore, the last term on the right in 10.11 converges to 0 and so, taking a limit as $R\to\infty$ in 10.11 yields $\mathscr{L}(f*g)(s) = \mathscr{L}(f)(s)\mathscr{L}(g)(s)$. ■

That which is most certainly not obvious is the following major theorem. This is omitted from virtually all ordinary differential equations books, and it is this very thing which justifies the use of Laplace transforms in solving various equations. Without it or something like it, the whole method is nonsense. I am following Widder [26]. This theorem says that if you know the Laplace transform, this will determine the function it came from at every point of continuity. The proof only involves the theory of the integral which was presented in this chapter and Stirling's formula. Also, it would easily generalize to functions having values in some normed vector space. First is a lemma.

Lemma 10.4.7 *Let $a>0$ and $b>1$. Then* $\lim_{k\to\infty}\frac{1}{k!}\int_{bk}^\infty v^k e^{-\left(1-\frac{a}{k}\right)v}dv = 0$.

Proof: First change variables $\left(1-\frac{a}{k}\right)v = kx$. A few computations show that the above integral is $\frac{1}{k!}\frac{k^{k+1}}{\left(1-\frac{a}{k}\right)^{k+1}}\int_{\left(1-\frac{a}{k}\right)b}^\infty x^k e^{-kx}dx$. Now let $1<\hat{b}<b$. Then for k large enough, $\left(1-\frac{a}{k}\right)b>\hat{b}$, so the above integral is dominated by $\frac{1}{k!}\frac{k^{k+1}}{\left(1-\frac{a}{k}\right)^{k+1}}\int_{\hat{b}}^\infty x^k e^{-kx}dx$. Using integration by parts and Stirling's formula which implies that for large $k, k! > \frac{1}{2}\sqrt{2\pi}k^{k+(1/2)}e^{-k}$, and also that $\left(1+\frac{a}{k}\right)^k \leq e^a$, this is dominated for large k by

$$\frac{1}{k!}\frac{k^{k+1}}{\left(1-\frac{a}{k}\right)^{k+1}}\frac{e^{-k\hat{b}}}{k}\sum_{j=0}^k \hat{b}^j = \frac{2}{\sqrt{2\pi}\sqrt{k}}\frac{1}{1-\frac{a}{k}}\frac{e^a}{\left(1-\frac{a^2}{k^2}\right)^k}e^{-k\left(\hat{b}-1\right)}\left(\frac{\hat{b}^{k+1}-1}{\hat{b}-1}\right)$$

$$< e^{-k\left(\hat{b}-1\right)}\left(\frac{\hat{b}^{k+1}-1}{\hat{b}-1}\right)$$

which converges to 0 as $k\to\infty$ since $\hat{b}>1$. ■

Theorem 10.4.8 *Let ϕ have exponential growth, $|\phi(t)|\leq Ce^{mt}$ where we can let $m\geq 0$. Suppose also that ϕ is integrable on every interval $[0,R]$ and let $f(s)\equiv\mathscr{L}(\phi)(s)$. Then if t is a point of continuity of ϕ, it follows that*

$$\phi(t) = \lim_{k\to\infty}\frac{(-1)^k}{k!}\left[f^{(k)}\left(\frac{k}{t}\right)\right]\left(\frac{k}{t}\right)^{k+1}.$$

Thus $\phi(t)$ is determined by its Laplace transform at every point of continuity.

Proof: $f(s) \equiv \int_0^\infty e^{-su}\phi(u)\,du$ so $f^{(k)}(s) = \int_0^\infty (-u)^k e^{-su}\phi(u)\,du$. This is valid for all s large enough and the exponential growth of $\phi(t)$ thanks to Theorem 10.3.5. Formally, you differentiate under the integral. Then, always assuming k is sufficiently large,

$$\frac{(-1)^k}{k!}\left[f^{(k)}\left(\frac{k}{t}\right)\right]\left(\frac{k}{t}\right)^{k+1} = \frac{(-1)^k}{k!}\left(\int_0^\infty (-u)^k e^{-\frac{k}{t}u}\phi(u)\,du\right)\left(\frac{k}{t}\right)^{k+1}$$

Now let $v = \frac{ku}{t}$ so this becomes

$$\frac{(-1)^k}{k!}\left(\int_0^\infty \left(-\frac{tv}{k}\right)^k e^{-v}\phi\left(\frac{tv}{k}\right)\frac{t}{k}dv\right)\left(\frac{k}{t}\right)^{k+1} = \frac{1}{k!}\int_0^\infty v^k e^{-v}\phi\left(\frac{tv}{k}\right)dv$$

$\int_0^\infty \frac{1}{k!}v^k e^{-v}dv = 1$ by Proposition 10.3.3 and so the above equals

$$= \phi(t) + \frac{1}{k!}\int_0^\infty v^k e^{-v}\left(\phi\left(\frac{tv}{k}\right) - \phi(t)\right)dv$$

Suppose now that ϕ is continuous at $t > 0, 0 < \delta < t$. To say that $\left|\frac{tv}{k} - t\right| < \delta$ is to say that $v \in \left(\frac{t-\delta}{t}k, \frac{t+\delta}{t}k\right)$. Split the integral into one which goes from 0 to $\frac{t-\delta}{t}k$, one from $\frac{t-\delta}{t}k$ to $\frac{t+\delta}{t}k$, and one from $\frac{t+\delta}{t}k$ to ∞ where δ is small enough that when $\left|\frac{tv}{k} - t\right| < \delta, \left|\phi\left(\frac{tv}{k}\right) - \phi(t)\right| < \varepsilon$. Then the middle integral

$$\left|\frac{1}{k!}\int_{\frac{t-\delta}{t}k}^{\frac{t+\delta}{t}k} v^k e^{-v}\left(\phi\left(\frac{tv}{k}\right) - \phi(t)\right)dv\right| \leq \frac{1}{k!}\int_{\frac{t-\delta}{t}k}^{\frac{t+\delta}{t}k} v^k e^{-v}\varepsilon dv \leq \varepsilon$$

It remains to consider the other two integrals.

$$\frac{1}{k!}\int_0^{\frac{t-\delta}{t}k} v^k e^{-v}\left(\phi\left(\frac{tv}{k}\right) - \phi(t)\right)dv + \frac{1}{k!}\int_{\frac{t+\delta}{t}k}^\infty v^k e^{-v}\left(\phi\left(\frac{tv}{k}\right) - \phi(t)\right)dv \qquad (*)$$

Now $v \to v^k e^{-v}$ is increasing for $v < k$. The first of these, on this interval, $\frac{tv}{k} \leq t - \delta$ and so there is a constant $C(t)$ such that $\left|\phi\left(\frac{tv}{k}\right) - \phi(t)\right| < C(t)$. Thus, using Stirling's formula, the first integral is dominated by

$$\begin{aligned} C(t)\frac{1}{k!}\int_0^{\frac{t-\delta}{t}k} v^k e^{-v}dv &\leq 2C(t)\frac{1}{\sqrt{2\pi}k^{k+1/2}e^{-k}}\left(\frac{t-\delta}{t}k\right)\left(\frac{t-\delta}{t}k\right)^k e^{-\left(\frac{t-\delta}{t}k\right)} \\ &< 2C(t)\frac{1}{\sqrt{2\pi}}\sqrt{k}\left(e^{\frac{\delta}{t}}\left(1 - \frac{\delta}{t}\right)\right)^k \equiv 2C(t)\frac{1}{\sqrt{2\pi}}\sqrt{k}r^k \end{aligned}$$

where $r \equiv e^{\frac{\delta}{t}}\left(1 - \frac{\delta}{t}\right)$ which is less than 1 since $\delta < t$. Then $\lim_{k\to\infty}\sqrt{k}r^k = 0$ by a use of the root test. In the second integral of $*$, $\left|\phi\left(\frac{tv}{k}\right) - \phi(t)\right| \leq C(t)e^{m\frac{tv}{k}}$. Then, simplifying this second integral, it is dominated by $\frac{C(t)}{k!}\int_{\frac{t+\delta}{t}k}^\infty v^k e^{-\left(1-\frac{mt}{k}\right)v}dv$ which converges to 0 as $k \to \infty$ by Lemma 10.4.7. ■

I think the approach given above is especially interesting because it gives an explicit description of $\phi(t)$ at most points[1]. I will next give a proof based on the Weierstrass approximation theorem to prove this major result which shows that the function is determined by

[1] If ϕ is piecewise continuous on every finite interval, this is obvious. However, the method will end up showing this is true more generally.

its Laplace transform. I think it is easier to follow but lacks the explicit description given above. Later in the book is a way to actually compute the function with given Laplace transform using a contour integral and the method of residues from complex analysis.

Lemma 10.4.9 *Suppose q is a continuous function defined on $[0,1]$. Also suppose that for all $n = 0,1,2,\cdots$, $\int_0^1 q(x)x^n dx = 0$. Then it follows that $q = 0$.*

Proof: By assumption, for $p(x)$ any polynomial, $\int_0^1 q(x)p(x)dx = 0$. Now let $\{p_n(x)\}$ be a sequence of polynomials which converge uniformly to $q(x)$ by Theorem 6.10.2. Say $\max_{x\in[0,1]} |q(x) - p_n(x)| < \frac{1}{n}$ Then

$$\begin{aligned}\int_0^1 q^2(x)\,dx &= \int_0^1 q(x)(q(x)-p_n(x))\,dx + \overbrace{\int_0^1 q(x)\,p_n(x)\,dx}^{=0} \\ &\leq \int_0^1 |q(x)(q(x)-p_n(x))|\,dx \leq \int_0^1 |q(x)|\,dx\frac{1}{n}\end{aligned}$$

Since n is arbitrary, it follows that $\int_0^1 q^2(x)dx = 0$ and by continuity, it must be the case that $q(x) = 0$ for all x since otherwise, there would be a small interval on which $q^2(x)$ is positive and so the integral could not have been 0 after all. ■

Lemma 10.4.10 *Suppose $|\phi(t)| \leq Ce^{-\delta t}$ for some $\delta > 0$ and all $t > 0$ and also that ϕ is continuous. Suppose that $\int_0^\infty e^{-st}\phi(t)dt = 0$ for all $s > 0$. Then $\phi = 0$.*

Proof: First note that $\lim_{t\to\infty}|\phi(t)| = 0$. Next change the variable letting $x = e^{-t}$ and so $x \in [0,1]$. Then this reduces to $\int_0^1 x^{s-1}\phi(-\ln(x))dx$. Now if you let $q(x) = \phi(-\ln(x))$, it is not defined when $x = 0$, but $x = 0$ corresponds to $t \to \infty$. Thus $\lim_{x\to 0+} q(x) = 0$. Defining $q(0) \equiv 0$, it follows that it is continuous and for all $n = 0,1,2,\cdots$, $\int_0^1 x^n q(x)dx = 0$ and so $q(x) = 0$ for all x from Lemma 10.4.9. Thus $\phi(-\ln(x)) = 0$ for all $x \in (0,1]$ and so $\phi(t) = 0$ for all $t \geq 0$. ■

Now suppose only that $|\phi(t)| \leq Ce^{\lambda t}$ so ϕ has exponential growth and that for all s sufficiently large, $\mathscr{L}(\phi) = 0$. Does it follow that $\phi = 0$? Say this holds for all $s \geq s_0$ where also $s_0 > \lambda$. Then consider $\hat{\phi}(t) \equiv e^{-s_0 t}\phi(t)$. Then if $s > 0$,

$$\int_0^\infty e^{-st}\hat{\phi}(t)\,dt = \int_0^\infty e^{-st}e^{-s_0 t}\phi(t)\,dt = \int_0^\infty e^{-(s+s_0)t}\phi(t)\,dt = 0$$

because $s + s_0$ is large enough for this to happen. It follows from Lemma 10.4.10 that $\hat{\phi} = 0$. But this implies that $\phi = 0$ also. This proves the following fundamental theorem.

Theorem 10.4.11 *Suppose ϕ has exponential growth and is continuous on $[0,\infty)$. Suppose also that for all s large enough, $\mathscr{L}(\phi)(s) = 0$. Then $\phi = 0$.*

This proves the case where ϕ is continuous. Can one still recover ϕ at points of continuity? Suppose ϕ is continuous at every point but finitely many on each interval $[0,t]$ and has exponential growth and $\mathscr{L}(\phi)(s) = 0$ for all s large enough. Does it follow that $\phi(t) = 0$ for t a point of continuity of ϕ? Approximating with finite intervals $[0,R]$ in place of $[0,\infty)$ and then taking a limit, (details left to you.)

$$0 = \int_0^\infty e^{-st}\phi(t)\,dt = e^{-st}\int_0^t \phi(u)\,du\Big|_{t=0}^\infty + s\int_0^\infty e^{-st}\left(\int_0^t \phi(u)\,du\right)dt$$

The boundary term is 0 for large s because

$$\left|\int_0^t \phi(u)\,du\right| \leq \int_0^t |\phi(u)|\,du \leq \int_0^t Ce^{\lambda u}du = \frac{C}{\lambda}\left(e^{\lambda t}-1\right) \leq \frac{C}{\lambda}e^{\lambda t}$$

Therefore, $0=\int_0^\infty e^{-st}\left(\int_0^t \phi(u)\,du\right)dt$ and by Theorem 10.4.11, $\int_0^t \phi(u)\,du = 0$ for all $t>0$. Then by the fundamental theorem of calculus, Corollary 9.5.6, $g'(t)=\phi(t)=0$ at every point of continuity. This proves the following theorem.

Theorem 10.4.12 *Suppose ϕ has exponential growth having finitely many points of discontinuity on every interval of the form $[0,R]$. Suppose also that for all s large enough, $\mathscr{L}(\phi)(s)=0$. Then $\phi(t)=0$ whenever ϕ is continuous at t.*

Note that this implies that if $\mathscr{L}\phi=\mathscr{L}\psi$ then, $\mathscr{L}(\phi-\psi)=0$ so $\phi=\psi$ at all points of continuity.

10.5 Exercises

1. Prove Lemma 10.4.5 by considering

$$F(z) \equiv \int_0^z\int_0^t f(t-u)\,g(u)\,du\,dt - \int_0^z\int_u^z f(t-u)\,g(u)\,dt\,du.$$

   Explain why $F(0)=0$ and use the first mean value theorem for integrals to show that $F'(z)=0$ for $z>0$. Generalize to f,g are piecewise continuous.

2. Find $\Gamma\left(\frac{1}{2}\right)$. **Hint:** $\Gamma\left(\frac{1}{2}\right)\equiv\int_0^\infty e^{-t}t^{-1/2}dt$. Explain why this equals $2\int_0^\infty e^{-u^2}du$. Then use Problem 49 on Page 227 or Theorem 9.9.4. Find a formula for $\Gamma\left(\frac{3}{2}\right),\Gamma\left(\frac{5}{2}\right)$, etc.

3. For $p,q>0$, $B(p,q)\equiv\int_0^1 x^{p-1}(1-x)^{q-1}dx$. This is called the beta function. Show $\Gamma(p)\Gamma(q)=B(p,q)\Gamma(p+q)$. **Hint:** You might want to adapt and use the Fubini theorem presented earlier in Theorem 9.9.3 on Page 216 about iterated integrals.

4. Verify that $\mathscr{L}(\sin(t))=\frac{1}{1+s^2}$.

5. Show directly that $\int_0^\infty \frac{\sin x}{x}dx$ exists. **Hint:** For large R and small ε, consider

$$\int_\varepsilon^R \frac{\sin t}{t}dt = \int_\varepsilon^{\pi/2}\frac{\sin t}{t}dt + \int_{\pi/2}^R \frac{\sin t}{t}dt$$

   and show that the limit of the first as $\varepsilon\to 0$ exists and the limit of the second as $R\to\infty$ also exists. **Hint:** The second integral is

$$\frac{-\cos t}{t}\Big|_{\pi/2}^R - \int_{\pi/2}^R \frac{\cos t}{t^2}dt = \frac{-\cos(R)}{R} - \int_{\pi/2}^R \frac{\cos t}{t^2}dt$$

   Then $t\to\frac{\cos t}{t^2}$ is in $L^1\left(\frac{\pi}{2},\infty\right)$ so the limit exists as $R\to\infty$. Why? Then the first term on the right converges to 0 as $R\to\infty$. Now consider $\int_\varepsilon^{\pi/2}\frac{\sin t}{t}dt$. Explain why the integrand is positive and bounded above by $1/\sqrt{t}$. Then compare with $\int_\varepsilon^{\pi/2}\frac{1}{\sqrt{x}}dx = 2\sqrt{\pi/2}-2\sqrt{\varepsilon}$. Argue that

$$\lim_{\varepsilon\to 0+}\int_\varepsilon^{\pi/2}\frac{\sin t}{t}dt = \sup\left\{\int_\varepsilon^{\pi/2}\frac{\sin t}{t}dt : 0<\varepsilon<\pi/2\right\} \leq 2\sqrt{\pi/2}.$$

6. For $s \geq 0$, define $F(s) \equiv \int_0^\infty e^{-sx} \frac{\sin x}{x} dx$. Thus
$$F(0) - F(s) = \lim_{R \to \infty} \int_0^R \left(1 - e^{-sx}\right) \frac{\sin x}{x} dx.$$
Show $\lim_{s \to 0+} F(s) = F(0)$. **Hint:** You might try to justify the following steps or something similar.
$$\begin{aligned} |F(0) - F(s)| &\leq \left| \int_0^M \left(1 - e^{-st}\right) \right| + \left| -\frac{\cos t}{t} |_M^\infty \right| + \left| \int_M^\infty \frac{1}{t^2} dt \right| \\ &\leq \left| \int_0^M \left(1 - e^{-st}\right) \right| + \frac{1}{M} + \frac{1}{M} \end{aligned}$$
Now pick M very large and then when it is chosen in an auspicious manner, let $s \to 0+$ and show the first term on the right converges to 0 as this happens.

7. It was shown that $\mathscr{L}(\sin(t)) = \frac{1}{1+s^2}$. Show that it makes sense to take $\mathscr{L}\left(\frac{\sin t}{t}\right)$. Show that $\int_0^\infty \frac{\sin(t)}{t} e^{-st} dt = \frac{\pi}{2} - \int_0^s \frac{1}{1+u^2} du$. To do this, let $F(s) = \int_0^\infty \frac{\sin(t)}{t} e^{-st} dt$ and show using Theorem 10.3.5 that $F'(s) = -\frac{1}{1+s^2}$ so $F(s) = -\arctan(s) + C$. Then argue that as $s \to \infty, F(s) \to 0$. Use this to determine C. Then when you have done this, you will have an interesting formula valid for all positive s. To finish it, let $s = 0$. From Problem 6 $F(0) = \lim_{s \to 0+} F(s)$, this gives $\int_0^\infty \frac{\sin x}{x} dx$, the Dirichlet integral. Another derivation is given earlier in the chapter.

8. Verify the following short table of Laplace transforms. Here $F(s)$ is $\mathscr{L} f(s)$.

$f(t)$	$F(s)$	$f(t)$	$F(s)$	$f(t)$	$F(s)$
$t^n e^{at}$	$\frac{n!}{(s-a)^{n+1}}$	$t^n, n \in \mathbb{N}$	$\frac{n!}{s^{n+1}}$	$e^{at} \sin bt$	$\frac{b}{(s-a)^2+b^2}$
$e^{at} \cos bt$	$\frac{s-a}{(s-a)^2+b^2}$	$f * g(t)$	$F(s) G(s)$		

9. Let r be a positive integer. Then if $f(x) = \frac{1}{\Gamma(r/2) 2^{r/2}} x^{(r/2)-1} e^{-x/2}$, this function is called a chi-squared density, denoted as $\mathscr{X}^2(r)$. Show for each r, $\int_0^\infty f(x) dx = 1$. This particular function is the basis for a large part of mathematical statistics.

10. The Fresnel integrals are $\int_0^x \sin\left(t^2\right) dt, \int_0^x \cos\left(t^2\right) dt$ for $x > 0$. This problem is on the limit of these as $x \to \infty$. In an earlier problem this limit was shown to exist. This limit is probably most easily done in the context of contour integrals from complex analysis. However, here is a real analysis way. Justify the following steps. Let $F(x) \equiv \left(\int_0^x e^{it^2} dt\right)^2$
$$F'(x) = 2\left(\int_0^x e^{it^2} dt\right)\left(e^{ix^2}\right) = 2x\left(\int_0^1 e^{ix^2 t^2} dt\right)\left(e^{ix^2}\right) = 2x\left(\int_0^1 e^{ix^2\left(t^2+1\right)} dt\right)$$
$$\begin{aligned} F(x) &= 2\int_0^x \int_0^1 y e^{iy^2\left(t^2+1\right)} dt dx = \int_0^1 \int_0^x 2y e^{iy^2\left(t^2+1\right)} dx dt \\ &= \int_0^1 \left(\frac{-ie^{iy^2\left(t^2+1\right)}}{t^2+1} |_0^x\right) dt = \int_0^1 \left(i\frac{1}{t^2+1} - i\frac{e^{ix^2\left(t^2+1\right)}}{t^2+1}\right) dt \end{aligned}$$

Let $u = t^2 + 1$ so $t = \sqrt{u-1}$ and $du = 2\sqrt{u-1}dt$. Then in terms of u, the second integral is $i\frac{1}{2}\int_1^2 \frac{e^{ix^2 u}}{u(u-1)^{1/2}}du$. This converges to 0 as $x \to \infty$ by the Riemann Lebesgue lemma and the observation that the improper integral $\int_1^2 \frac{1}{u(u-1)^{1/2}}du < \infty$. Therefore,

$$\left(\lim_{x\to\infty}\int_0^x e^{it^2}dt\right)^2 = i\frac{\pi}{4} = i\int_0^1 \frac{1}{t^2+1}dt$$

and so $\int_0^\infty e^{it^2}dt = \int_0^\infty \cos\left(t^2\right)dt + i\int_0^\infty \sin\left(t^2\right)dt = \frac{\sqrt{2}}{2}\frac{\sqrt{\pi}}{2} + i\frac{\sqrt{2}}{2}\frac{\sqrt{\pi}}{2}$ or $-\frac{\sqrt{2}}{2}\frac{\sqrt{\pi}}{2} - i\frac{\sqrt{2}}{2}\frac{\sqrt{\pi}}{2}$. The first of the two alternatives will end up holding. You can see this from observing that $\int_0^\infty \sin\left(t^2\right)dt > 0$ from numerical experiments. Indeed, $\int_0^{10}\sin\left(t^2\right)dt = 0.583\,67$ and for t larger than 10, the contributions to the integral will be small because of the rapid oscillation of the function between -1 and 1.

11. Let $a = x_0 < x_1 < \cdots < x_n = b$ and let $y_i \in [x_{i-1}, x_i] \equiv I_i$. If $\delta : \mathbb{R} \to (0,\infty)$, then the collection $\{(I_i, y_i)\}$ is called a "δ fine division" if for each i,

$$I_i \subseteq (y_i - \delta(y_i), y_i + \delta(y_i)).$$

Show that for any such function δ, there exists a δ fine division. **Hint:** If not, then there would not be one for one of $\left[a, \frac{a+b}{2}\right], \left[\frac{a+b}{2}, b\right]$. Use nested interval lemma to get a contradiction.

12. Show directly, using the Weierstrass approximation theorem, that if f is piecewise continuous on $[a,b]$, then $\lim_{n\to\infty}\int_a^b \sin(nx+c)f(x)dx = 0$. **Hint:** Show it suffices to suppose f is continuous. For f continuous, let g be a polynomial such that $\|f-g\| \equiv \max_{x\in[a,b]}|f(x)-g(x)| < \frac{\varepsilon}{2(b-a)}$. Then $\int_a^b \sin(nx+c)f(x)dx =$

$$\int_a^b \sin(nx+c)(f(x)-g(x))dx + \int_a^b \sin(nx+c)g(x)dx.$$

Now the first integral is small. Use integration by parts in the second.

13. Carefully fill in the details of Lemma 10.4.7.

14. Suppose you have f defined, positive, and decreasing on $[0,\infty)$. Then show that f is in $L^1(0,\infty)$ if and only if $\sum_{k=1}^\infty f(k)$ is a convergent series. This is called the integral test.

15. Assume all the integrals make sense as ordinary or improper integrals on (a,b) where $-\infty \le a < b \le \infty$. Also let $\phi : (a,b) \to \mathbb{R}$ be convex and differentiable. Convexity here means that ϕ' is an increasing function. Thus the graph of ϕ "smiles" and ϕ is always at least as large as any tangent line. Suppose $\int_a^b f(t)dt = 1$. Show that

$$\phi\left(\int_a^b g(t)f(t)dt\right) \le \int_a^b \phi(g(t))f(t)dt$$

This is a case of Jensen's inequality. **Hint:** Since ϕ is convex,

$$\begin{aligned}\phi(g(t)) \quad &\ge \quad \phi\left(\int_a^b g(s)f(s)ds\right) + \\ &\qquad \phi'\left(\int_a^b g(s)f(s)ds\right)\left(\phi(g(t)) - \int_a^b g(s)f(s)ds\right).\end{aligned}$$

Now multiply by $f(t)$ and do $\int_a^b dt$ using $\int_a^b f(t)\,dt = 1$.

16. For $n \in \mathbb{N}$, Stirling's formula says $\lim_{n\to\infty} \frac{\Gamma(n+1)e^n}{n^{n+(1/2)}} = \sqrt{2\pi}$. Here $\Gamma(n+1) = n!$. The idea here is to show that you get the same result if you replace n with $x \in (0,\infty)$. To do this, show

    (a) $n \to \frac{\Gamma(n+1)e^n}{n^{n+(1/2)}}$ is decreasing on the positive integers. This follows from the properties of the Gamma function and a little work.

    (b) Show that $x \to \frac{\Gamma(x+1)e^x}{x^{x+(1/2)}}$ is decreasing on $(m,m+1)$ for $m \in \mathbb{N}$. This is a little harder.

Hint: For $x \in (m,m+1)$, $\ln\left(\frac{\Gamma(x+1)e^x}{x^{x+(1/2)}}\right) =$

$$\begin{aligned} & x+\ln\Gamma(x+1)-\left(x+\frac{1}{2}\right)\ln x \\ = \quad & x+\ln\left(x(x-1)(x-2)\cdots(x-m+1)\Gamma(x-m)\right)-\left(x+\frac{1}{2}\right)\ln x \\ = \quad & x+\sum_{k=0}^{m-1}\ln(x-k)+\ln(\Gamma(x-m))-\left(x+\frac{1}{2}\right)\ln x \end{aligned}$$

Now differentiate and try to show that the derivative is negative for $x \in (m,m+1)$. Thus the desired derivative is

$$\left(\sum_{k=0}^{m-1}\frac{1}{x-k}-\ln x\right)+\frac{1}{\Gamma(x-m)}\int_0^\infty \ln(t)\,t^{x-(m+1)}e^{-t}dt-\frac{1}{2x}$$

The first term is negative from the definition of $\ln(x)$. The derivative being negative will be shown if it is shown that the integral in the above is negative. Do an integration by parts on this integral and split the integrals to obtain

$$\begin{aligned} \int_0^\infty \ln(t)\,t^{x-(m+1)}e^{-t}dt \quad = \quad & -\int_0^1 t^\sigma e^{-t}dt+\int_0^1 (t-\sigma)\,e^{-t}t^\sigma \ln(t)\,dt \\ & +\int_1^\infty t^\sigma e^{-t}\left(1-(t-\sigma)\ln(t)\right)dt \end{aligned}$$

where $\sigma = (x-m)-1 \in (-1,0)$ so $-\sigma > 0$. The last integral is negative because $(t-\sigma) = t+(-\sigma) > 1$. The first two integrals on the right are obviously negative.

Chapter 11

Functions of One Complex Variable

In the nineteenth century, complex analysis developed along with real analysis, the latter being the main topic of this book in which one considers functions of one real variable. However, many difficult real improper integrals can be best considered using contour integrals so some introduction to this very important topic is useful. It is not intended to be a full course on complex analysis, just an introduction to some of the main ideas. Cauchy was the principal originator of the study of complex analysis in the early 1800's. Historically, the main theorems came from the Cauchy Riemann equations and a version of Green's theorem. The Cauchy Riemann equations are considered later in a problem, but Green's theorem is part of multivariable calculus which is not being discussed in this book. However, using the ideas of Goursat (1858-1936) it is possible to present the main theory in terms of functions of a single variable, this time a single complex variable.

11.1 Contour Integrals

This is about contour integrals in $\mathbb{C}$. First is the definition of an oriented C^1 contour. Contours are sets of points in $\mathbb{C}$. Smooth ones require the notion of the derivative of a function of one real variable having values in $\mathbb{C}$. There is also the concept of a differentiable function of a complex variable which is defined on an open set of $\mathbb{C}$.

Definition 11.1.1 *The derivative is defined as before. If $\gamma : [a,b] \to \mathbb{C}$, then $\gamma'(t)$ is said to exist exactly when $\gamma(t+h) - \gamma(t) = ah + o(h)$ for some $a \in \mathbb{C}$ and sufficiently small real h and in this case, $a \equiv \gamma'(t)$. This is not any different than the earlier material on the derivative other than γ having values in $\mathbb{C}$. Derivatives from right and left are similar to before. Also, if f is defined on an open subset of $\mathbb{C}$ then it is differentiable at z means $f(z+h) - f(z) = ah + o(h)$ and $a \equiv f'(z)$. Here $h \in \mathbb{C}$ since the derivative is on an open subset of $\mathbb{C}$ rather than a subset of $\mathbb{R}$. This has already been dealt with in Theorem 8.2.1 on Page 159. $f'(z) \equiv \lim_{h\to 0} \frac{f(z+h)-f(z)}{h}$ exactly as in the case of a real variable.*

All properties of Theorem 7.5.1 continue to apply for the derivative just defined. The proofs are exactly the same. In particular, the chain rule holds. You just have to use $\mathbb{C}$ rather than $\mathbb{R}$ for values of the function. The necessary complex arithmetic is in Section 2.13.

Next is the idea of an oriented C^1 curve.

Definition 11.1.2 *A set of points γ^* in $\mathbb{C}$ is called an oriented C^1 curve or contour if the following conditions hold.*

1. *There exists γ a continuous function mapping some interval $[a,b]$ to $\mathbb{C}$ such that $\gamma^* \equiv \gamma([a,b])$.*
2. *This γ is one to one on $[a,b)$ and the derivative γ' is continuous and exists on $[a,b]$, being defined in terms of right or left derivatives at the end points. When $\gamma(a) = \gamma(b)$ this is a closed curve.*
3. *If $\gamma : [a,b] \to \gamma^*$ and $\hat{\gamma} : [\hat{a},\hat{b}] \to \gamma^*$ are two parametrizations, then $\hat{\gamma}^{-1} \circ \gamma$ is a continuous function which is one to one and so by Lemma 6.4.3 it is either strictly increasing or strictly decreasing on $[a,b]$. Two parametrizations $\hat{\gamma}, \gamma$ are said to*

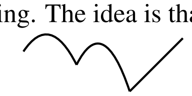
have the same orientation and $\gamma \sim \hat{\gamma}$ if and only if $\hat{\gamma}^{-1} \circ \gamma$ is increasing. Then $\sim$ is easily seen to be an equivalence relation and γ^* together with all C^1 parametrizations having the same orientation is called an oriented C^1 curve.

4. *The curve is called smooth if it has a C^1 parametrization $\gamma : [a,b] \to \mathbb{C}$ such that $\gamma'(t) \neq 0$ for all $t \in (a,b)$.*

To see that $\sim$ is an equivalence relation say $\gamma_1 \sim \gamma_2$ and $\gamma_2 \sim \gamma_3$. Then $\gamma_2^{-1} \circ \gamma_1$ is increasing and also $\gamma_3^{-1} \circ \gamma_2$ is increasing. Hence $\left(\gamma_3^{-1} \circ \gamma_2\right) \circ \left(\gamma_2^{-1} \circ \gamma_1\right) = \gamma_3^{-1} \circ \gamma_1$ is increasing. Thus $\sim$ is transitive. It is symmetric because a function is increasing on an interval is equivalent to its inverse being increasing. Clearly $\gamma \sim \gamma$.

Definition 11.1.3 *An ordered partition of $[p,q]$ will be a sequence of intermediate points, $\alpha_0, \alpha_1, \cdots, \alpha_n$, $p = \alpha_0 < \alpha_1 < \cdots < \alpha_n = q$. A set of points $\gamma^* \subseteq \mathbb{C}$ is an oriented piecewise smooth curve means there is a parametrization γ which is one to one and continuous on $[p,q)$, such that γ restricted to (α_{i-1}, α_i) is C^1, and $\gamma'(t) \neq 0$ on this open interval, and the right and left derivatives exist at the endpoints. Thus $\gamma([\alpha_{i-1}, \alpha_i])$ is a smooth curve.*

Here is a picture of such a thing. The idea is that it has finitely many pointy places.



The above is fussy and technical, We can ignore it because it is included in the case of C^1 curves. Suppose you have $a < b < c$ and $\gamma_1'(t) \neq 0$ on (a,b), $\gamma_2'(t) \neq 0$ on (b,c) but $\gamma_1'(b) \neq \gamma_2'(b)$ although $\gamma_1 = \gamma_2$ at b. Then consider

$$\hat{\gamma}(t) = \begin{cases} \gamma_1\left(b + (t-b)^3 \frac{1}{(b-a)^2}\right), t \in [a,b] \\ \gamma_2\left(b + (t-b)^3 \frac{1}{(c-b)^2}\right), t \in [b,c] \end{cases}$$

Then $\hat{\gamma}(t)$ moves from $\gamma_1(a)$ to $\hat{\gamma}(b)$ in the same direction as γ_1 and γ_2 and is differentiable on all of (a,c) although $\hat{\gamma}'(b) = 0$. Thus this piecewise smooth curve can be expressed as a C^1 curve, not smooth because of the vanishing of the derivative at b. If you wanted, you could define and draw the same conclusions for a piecewise smooth C^2 curve. You would simply feature $(t-b)^5 \frac{1}{(b-a)^4}, (t-b)^5 \frac{1}{(c-b)^4}$ instead. This shows that a piecewise smooth curve, has a C^1 parametrization which gives the same oriented piecewise smooth curve, so one might as well just consider curves which have C^1 parametrizations without insisting they are smooth and forget about the fussy details.

Definition 11.1.4 *A piecewise smooth curve γ^* is an oriented C^1 curve having a C^1 parametrization $\gamma : [a,b] \to \gamma^*$ such that there exists an ordered partition of $[a,b]$, $\alpha_0, \alpha_1, \cdots, \alpha_n$ for which $\gamma'(t) \neq 0$ on (α_{i-1}, α_i) and right and left derivatives exist at the endpoints of this interval. More generally, a C^1 curve is one which has a C^1 parametrization as above. Also define $-\gamma^*$ as follows. If $\gamma^* = \gamma([a,b])$ then let $\eta : [a,b] \to \gamma^*$ be defined by $\eta(t) \equiv \gamma(a+b-t)$. This amounts to going over γ^* in the opposite direction. Then η will be a parametrization for $-\gamma^*$.*

Also, I will define the contour integral for a C^1 curve as follows:

Definition 11.1.5 *Let γ^* be an oriented C^1 curve having a C^1 parametrization $\gamma : [a,b] \to \gamma^*$, then*

$$\begin{aligned}\int_{\gamma^*} f(z)\,dz &\equiv \int_a^b f(\gamma(t))\,\gamma'(t)\,dt \equiv \int_a^b \operatorname{Re}\big(f(\gamma(t))\,\gamma'(t)\big)\,dt \\ &+i\int_a^b \operatorname{Im}\big(f(\gamma(t))\,\gamma'(t)\big)\,dt\end{aligned}$$

This is well defined because all the functions are continuous. Then $\int_{\gamma^*} f(z)\,dz$ is called a contour integral.

Proposition 11.1.6 *The above contour integral is well defined and for γ^* an oriented curve, $f \to \int_{\gamma^*} f(z)\,dz$ is a complex linear map meaning that for $a,b \in \mathbb{C}$,*

$$\int_{\gamma^*} (af(z) + bg(z))\,dz = a\int_{\gamma^*} f(z)\,dz + b\int_{\gamma^*} g(z)\,dz$$

Also $\int_{\gamma^*} f(z)\,dz = -\int_{-\gamma^*} f(z)\,dz$. In addition to this, if $M \geq \max\{|f(z)| : z \in \gamma^*\}$, one obtains the estimate $\left|\int_{\gamma^*} f(z)\,dz\right| \leq ML$ where L is the length of γ^* defined as $\int_a^b |\gamma'(t)|\,dt$ for γ a parametrization for γ^*. This number L is well defined. If f_n converges uniformly to f on γ^*, then $\lim_{n\to\infty}\int_\gamma f_n(z)\,dz = \int_\gamma f(z)\,dz$

Proof: The claim about the contour integral being linear is a routine computation from doing arithmetic for complex numbers and the above definition. This is obvious for a,b real. In case $b = 0$ and $a = i$,

$$\begin{aligned}i\int_{\gamma^*} f(z)\,dz &\equiv i\left(\int_a^b \operatorname{Re}\big(f(\gamma(t))\,\gamma'(t)\big)\,dt + i\int_a^b \operatorname{Im}\big(f(\gamma(t))\,\gamma'(t)\big)\,dt\right) \\ &= i\int_a^b \operatorname{Re}\big(f(\gamma(t))\,\gamma'(t)\big)\,dt - \int_a^b \operatorname{Im}\big(f(\gamma(t))\,\gamma'(t)\big)\,dt \\ \int_{\gamma^*} if(z)\,dz &\equiv \int_a^b \big(i\operatorname{Re}\big(f(\gamma(t))\,\gamma'(t)\big) - \operatorname{Im}\big(f(\gamma(t))\,\gamma'(t)\big)\big)\,dt\end{aligned}$$

which is the same thing because it holds for Riemann sums approximating the various integrals.

From consideration of real and imaginary parts, the usual change of variables formula holds. If $\gamma, \hat{\gamma}$ are two equivalent parametrizations giving the same orientation, $\gamma : [a,b] \to \gamma^*$ and $\hat{\gamma} : [c,d] \to \gamma^*$. I need to show these give the same thing for the contour integral. Let $s = \hat{\gamma}^{-1} \circ \gamma(t)$ so $ds = \left(\hat{\gamma}^{-1} \circ \gamma\right)'(t)\,dt$. Also $\gamma(t) = \hat{\gamma}\left(\hat{\gamma}^{-1} \circ \gamma(t)\right)$ so $\gamma'(t) = \hat{\gamma}'\left(\hat{\gamma}^{-1} \circ \gamma(t)\right)\left(\hat{\gamma}^{-1} \circ \gamma\right)'(t)$

$$\int_c^d f(\hat{\gamma}(s))\,\hat{\gamma}'(s)\,ds = \int_a^b f\left(\hat{\gamma}\left(\hat{\gamma}^{-1} \circ \gamma(t)\right)\right)\hat{\gamma}'\left(\hat{\gamma}^{-1} \circ \gamma(t)\right)\left(\hat{\gamma}^{-1} \circ \gamma\right)'(t)\,dt$$

$$= \int_a^b f(\gamma(t))\,\gamma'(t)\,dt$$

Now if η is the above parametrization corresponding to $-\gamma^*, \eta'(t) = -\gamma'(a+b-t)$. Thus letting $s = a+b-t$

$$\begin{aligned}\int_a^b f(\eta(t))\eta'(t)\,dt &= -\int_a^b f(\gamma(a+b-t))\gamma'(a+b-t)\,dt \\ &= \int_b^a f(\gamma(s))\gamma'(s)\,dt = -\int_a^b f(\gamma(t))\gamma'(t)\,dt\end{aligned}$$

As to the estimate. $\left|\int_a^b f(\gamma(t))\gamma'(t)\,dt\right| = \omega\int_a^b f(\gamma(t))\gamma'(t)\,dt$ for a suitable complex number, $|\omega| = 1$ and the last expression must be $\int_a^b \operatorname{Re}(\omega f(\gamma(t))\gamma'(t))\,dt \leq \int_a^b M|\gamma'(t)| = ML$.

Why is the length well defined? Say $\gamma, \hat{\gamma}$ are two parametrizations yielding the same orientation $\gamma : [a,b] \to \gamma^*$ and $\hat{\gamma} : [c,d] \to \gamma^*$. Then let $s = \hat{\gamma}^{-1} \circ \gamma(t)$ then by the same change of variables result,

$$\begin{aligned}\int_c^d |\hat{\gamma}'(s)|\,ds &= \int_a^b \left|\hat{\gamma}\left(\hat{\gamma}^{-1}\circ\gamma(t)\right)\right|\left(\hat{\gamma}^{-1}\circ\gamma\right)'(t)\,dt \\ &= \int_a^b \left|\hat{\gamma}\left(\hat{\gamma}^{-1}\circ\gamma(t)\right)\right|\left(\hat{\gamma}^{-1}\circ\gamma\right)'(t)\,dt \\ &= \int_a^b \left|\hat{\gamma}\left(\hat{\gamma}^{-1}\circ\gamma(t)\left(\hat{\gamma}^{-1}\circ\gamma\right)'(t)\right)\right|dt = \int_a^b |\gamma'(t)|\,dt\end{aligned}$$

The last claim follows right away from the estimate. If $f_n \to f$ uniformly on γ^*, then

$$\left|\int_{\gamma^*} f_n(z)\,dz - \int_{\gamma^*} f(z)\,dz\right| \leq \varepsilon L$$

whenever n is large enough that $\max\{|f_n(z) - f(z)| : z \in \gamma^*\} < \varepsilon$. ■

I will sometimes write $\int_\gamma f(z)\,dz$ instead of $\int_{\gamma^*} f(z)\,dz$ where it is understood that γ symbolizes any of the similar parametrizations of γ^* for one of the two orientations.

11.2 Cauchy Goursat, Cauchy Integral Theorem

In calculus, every continuous function has an antiderivative thanks to the fundamental theorem of calculus. However, the situation is not at all the same for functions of a complex variable. This is why we have the following definition using a different word.

Definition 11.2.1 *A function F with $F' = f$ is called a **primitive** of f.*

So what if a function has a primitive? It turns out that it becomes very easy to compute the contour integrals.

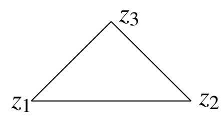
Theorem 11.2.2 *Suppose γ^* is an oriented C^1 curve Suppose $f : \gamma^* \to \mathbb{C}$ is continuous and has a primitive F. Thus $F'(z) = f(z)$ for some $\Omega \supseteq \gamma^*$. Then $\int_\gamma f(z)\,dz = F(\gamma(b)) - F(\gamma(a))$.*

Proof: By definition and the chain rule for derivatives,

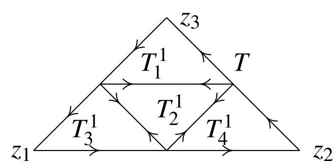
$$\int_\gamma f(z)\,dz = \int_a^b f(\gamma(t))\gamma'(t)\,dt = \int_a^b \frac{d}{dt}(F(\gamma(t)))\,dt = F(\gamma(b)) - F(\gamma(a)) \ ■$$

Many of these theorems in this section were first done by Cauchy using a version of Green's theorem which is not discussed in this book because it is on functions of one variable. Later, this other approach presented here was formulated by Goursat.

If you have two points in $\mathbb{C}$, z_1 and z_2, you can consider $\gamma(t) \equiv z_1 + t(z_2 - z_1)$ for $t \in [0,1]$ to obtain a smooth curve from z_1 to z_2. More generally, if $z_1, \cdots, z_m$ are points in $\mathbb{C}$ you can obtain a piecewise smooth 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 provided it does not intersect itself. 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.



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.



Thus

$$\int_{\partial T} f(z)\,dz = \sum_{k=1}^{4} \int_{\partial T_k^1} f(z)\,dz. \tag{11.1}$$

On the "inside lines" the integrals cancel because there are two integrals going in opposite directions for each of these inside lines.

Theorem 11.2.3 *(Cauchy Goursat) Let $f : \Omega \to \mathbb{C}$, where Ω is an open subset of $\mathbb{C}$ have the property that $f'(z)$ exists for all $z \in \Omega$ and let T be a triangle contained in Ω with the inside of the triangle also contained in Ω. Then $\int_{\partial T} f(w)\,dw = 0$.*

Proof: Suppose not. Then $\left|\int_{\partial T} f(w)\,dw\right| = \alpha \neq 0$. From 11.1 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, $\left|\int_{\partial T_1} f(w)\,dw\right| \geq \frac{\alpha}{4}$. 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, diam means diameter. It would be the length of the longest side.

$$T_k \supseteq T_{k+1}, \operatorname{diam}(T_k) \leq \operatorname{diam}(T)\,2^{-k},$$

and $\left|\int_{\partial T_k} f(w)\,dw\right| \geq \frac{\alpha}{4^k}$.

If you pick $z_k \in T_k$, then $\{z_k\}$ is a Cauchy sequence converging to some $z \in \mathbb{C}$. However, each of these triangles is a closed set so $z \in T_k$ for each k. Thus $z \in \cap_{k=1}^{\infty} T_k$. By assumption, $f'(z)$ exists. Therefore, for all k large enough,

$$\int_{\partial T_k} f(w)\,dw = \int_{\partial T_k} \left(f(z) + f'(z)(w-z) + o(w-z)\right) dw$$

where $|o(w-z)| < \varepsilon |w-z|$. Now observe that $w \to f(z) + f'(z)(w-z)$ has a primitive, namely, $F(w) = f(z)w + f'(z)(w-z)^2/2$. Then, by Theorem 11.2.2, $\int_{\partial T_k} f(w)\,dw = \int_{\partial T_k} o(w-z)\,dw$. From Theorem 11.1.6,

$$\begin{aligned}\frac{\alpha}{4^k} &\leq \left|\int_{\partial T_k} o(w-z)\,dw\right| \leq \varepsilon \operatorname{diam}(T_k)(\text{length of } \partial T_k) \\ &\leq \varepsilon 2^{-k} (\text{length of } \partial T) \operatorname{diam}(T) 2^{-k},\end{aligned}$$

and so $\alpha \leq \varepsilon (\text{length of } \partial T) \operatorname{diam}(T)$. Since ε is arbitrary, this shows $\alpha = 0$, a contradiction. Thus $\int_{\partial T} f(w)\,dw = 0$ as claimed. ■

Now we use this to construct a primitive.

Definition 11.2.4 *A set $\Omega \subseteq \mathbb{C}$ is convex if, whenever $z, w \in \Omega$, it follows that $tz + (1-t)w \in \Omega$ for all $t \in [0,1]$. In other words, if two points are in Ω then so is the line segment joining them.*

Lemma 11.2.5 *Suppose Ω is a convex set. Then so is the open set $\Omega + B(0,\delta)$. Here $\Omega + B(0,\delta) \equiv \cup_{z\in\Omega}\{z + B(0,\delta)\} = \cup_{z\in\Omega} B(z,\delta)$.*

Proof: First note that $z + B(0,\delta) = B(z,\delta)$ because $z+y$ is in the left if and only if $z + y - z \in B(0,\delta)$ if and only if $z + y \in B(z,\delta)$ and if $w \in B(z,\delta)$, then letting $y = w - z$, it follows that $z + y = w \in z + B(0,\delta)$. Thus $\Omega + B(0,\delta)$ is an open set because it is the union of open sets. If $z+y, \hat{z}+\hat{y}$ are in this set with $y,\hat{y}$ in $B(0,\delta)$, and $z,\hat{z} \in \Omega$, then if $t \in [0,1]$,

$$t(z+y) + (1-t)(\hat{z}+\hat{y}) = (tz + (1-t)\hat{z}) + (ty + (1-t)\hat{y})$$

The first term is in Ω and the second is in $B(0,\delta)$ because both sets are convex. ■

Theorem 11.2.6 *(Morera[1]) Let Ω be a convex open set and let $f'(z)$ exist for all $z \in \Omega$. Then f has a primitive on Ω*

Proof: Pick $z_0 \in \Omega$. Define $F(w) \equiv \int_{\gamma(z_0,w)} f(u)\,du$. Then by the Cauchy Goursat theorem, and $w \in \Omega$, it follows that for $|h|$ small enough,

$$\begin{aligned}\frac{F(w+h) - F(w)}{h} &= \frac{1}{h}\int_{\gamma(w,w+h)} f(u)\,du = \frac{1}{h}\int_0^1 f(w+th)\,h\,dt \\ &= \int_0^1 f(w+th)\,dt\end{aligned}$$

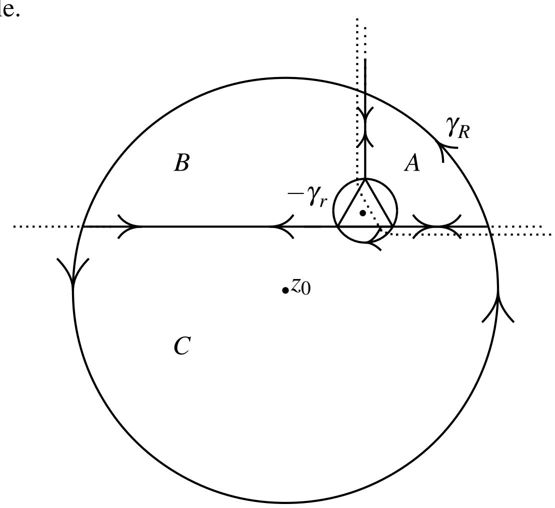
which converges to $f(w)$ due to the continuity of f at w. ■

You can get by with less rather easily.

Definition 11.2.7 *An open set U is star shaped if there is a point $p \in U$ called the star center such that if $z \in U$ is any other point, then the line segment $t \to p + t(z-p)$ for $t \in [0,1]$ is contained in U.*

[1] Giancinto Morera 1856-1909. This theorem or one like it dates from around 1886

The following picture is to illustrate the arguments which follow. z is the point in the center of the triangle.



The following corollary follows right away from Theorem 11.2.6. You just repeat the proof, but this time, you use the given star center rather than having the freedom to pick any point in the set.

Corollary 11.2.8 *Let Ω be a star shaped open set and suppose $f'(z)$ exists on Ω. Then f has a primitive on Ω.*

The above is a picture of piecewise smooth curves smashed together as shown in the picture. Also suppose $f'(z)$ exists for all $z \in \Omega$, an open convex set containing the large circle along with its inside, possibly $\Omega = \mathbb{C}$. γ_R corresponds to the large circle and $-\gamma_r$ is the parametrization for the small circle centered at z. The large circle is oriented counter clockwise and the small one is oriented clockwise. Thus γ_r would be oriented counter clockwise.

There are three contours sharing sides which are straight lines. Orient each of these three contours in the counter clockwise orientation as suggested by the arrows. Thus the integrals over the horizontal and vertical lines will cancel because they have opposite orientations. Now there are three regions labelled with A, B, C these are convex and bounded by the line segments consisting of the vertical and horizontal lines which are extended indefinitely, along with those line segments which, taken together, form the small triangle that encloses the point z. Thus these are convex sets of points. Letting $0 < \delta$ so that the distance from each of these convex sets to z is more than δ, Consider the open regions

$$A + B(0, \delta), B + B(0, \delta), C + B(0, \delta)$$

where the notation $E + F$ means $\{x + y : x \in E, y \in F\}$. I have illustrated $A + B(0, \delta)$ in the above picture. The three contours are contained in the open convex sets

$$(A + B(0, \delta)) \cap \Omega, (B + B(0, \delta)) \cap \Omega, (C + B(0, \delta)) \cap \Omega$$

Thus $f'(z)$ exists in each of these convex open sets. Hence, $w \to \frac{f(w)}{w-z} \equiv g(w)$ has a primitive in each of these convex open sets. Just use the quotient rule which holds for the same reason as it does for functions of a real variable. The only place there is a problem is where

$w=z$ and this is avoided. Since g has a primitive, each contour integral is 0. Since the integrals over the straight lines cancel, this reduces to

$$\int_{\gamma_R}\frac{f(w)}{w-z}dw+\int_{-\gamma_r}\frac{f(w)}{w-z}dw=0$$

or more conveniently,

$$\int_{\gamma_R}\frac{f(w)}{w-z}dw=\int_{\gamma_r}\frac{f(w)}{w-z}dw.$$

Now the integral on the right equals

$$\int_0^{2\pi}\frac{f\left(z+re^{i\theta}\right)}{re^{i\theta}}rie^{i\theta}d\theta=i\int_0^{2\pi}f\left(z+re^{i\theta}\right)d\theta$$

By continuity of f at z, the limit of this last integral as $r\to 0$ is

$$i\int_0^{2\pi}f(z)\,d\theta=2\pi i f(z)$$

This proves the most important theorem in complex analysis in the case of a circle, the Cauchy integral formula.

Theorem 11.2.9 *Suppose $f'(z)$ exists on an open set in $\mathbb{C}$ containing $D(z_0,R)\equiv\{z\in\mathbb{C}$ such that $|z-z_0|\le R\}$. Then if $z\in B(z_0,R)$, and γ_R is the oriented curve around the boundary of $B(z_0,R)$ oriented counter clockwise, then*

$$\frac{1}{2\pi i}\int_{\gamma_R}\frac{f(w)}{w-z}dw=f(z)$$

Remember how in Theorem 8.2.1 a function given by a power series had a derivative for all z in some open disk. With Theorem 11.2.9 it follows that if a function has a single complex derivative in an open set, then it has all of them because it is given by a power series. This is shown next.

Definition 11.2.10 *A function is called analytic on U an open subset of $\mathbb{C}$ if it has a derivative on U. This is also referred to as holomorphic.*

This definition is equivalent to the earlier use of the word "analytic" having to do with being representable with a power series which is the content of the following corollary.

Corollary 11.2.11 *Suppose f has a derivative on an open set containing the closed disk $D\left(z_0,\tilde{R}\right)$. Then there are a_k such that $f(z)=\sum_{k=0}^{\infty}a_k(z-z_0)^k$ for all z in this disk. Furthermore, convergence is absolute and uniform. Also,*

$$\frac{1}{2\pi i}\int_{\gamma_R}\frac{f(w)}{(w-z_0)^{k+1}}dw=\frac{f^{(k)}(z_0)}{k!}=a_k \tag{11.2}$$

Proof: By assumption, there is $\delta>0$ such that $f'(z)$ exists if $|z-z_0|\le\delta$. Let $R=\tilde{R}+\delta$. From Theorem 11.2.9, if γ_R is the circle of radius R which is centered at z_0, and if $z\in D\left(z_0,\tilde{R}\right)$

$$f(z)=\frac{1}{2\pi i}\int_{\gamma_R}\frac{f(w)}{w-z}dw=\frac{1}{2\pi i}\int_{\gamma_R}\frac{f(w)}{(w-z_0)\left(1-\frac{z-z_0}{w-z_0}\right)}dw$$

$$= \frac{1}{2\pi i}\int_{\gamma_R}\sum_{k=0}^{\infty}\frac{f(w)(z-z_0)^k}{(w-z_0)^{k+1}}dw$$

The series converges uniformly by the Weierstrass M test and also absolutely because

$$\left|\frac{f(w)(z-z_0)^k}{(w-z_0)^{k+1}}\right| \leq M\frac{\tilde{R}^k}{(\tilde{R}+\delta)^{k+1}}$$

where M is as large as the maximum value of $|f|$ on the compact set $D(z_0,R)$. It follows from Theorem 11.1.6, that one can interchange the integral with the sum. This yields

$$f(z) = \sum_{k=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma_R}\frac{f(w)}{(w-z_0)^{k+1}}dw\right)(z-z_0)^k$$

Then a use of Theorem 11.1.6 again and the Weierstrass M test shows the series converges uniformly and absolutely for all $|z-z_0|\leq\tilde{R}$. Corollary 8.2.2 shows that

$$\frac{1}{2\pi i}\int_{\gamma_R}\frac{f(w)}{(w-z_0)^{k+1}}dw = \frac{f^{(k)}(z_0)}{k!}.\ \blacksquare$$

In summary, this shows that for $f: U\to\mathbb{C}$ for U an open set in $\mathbb{C}$, it follows that if f' exists on U then near $z_0\in U$, f is given by a power series and has infinitely many derivatives. As to primitives, if F is one, then $F' = f$ and so F and hence f have all derivatives. Thus there is no such thing in this subject as a primitive of a function which is only continuous. It also shows that such differentiable functions of a complex variable are really glorified polynomials and you find eventually that in every way they behave just like polynomials. This was partially observed earlier in the material on power series. The above argument also shows that the power series for a function will converge on increasing disks until the circle bounding the disk encounters a point where the derivative of the function does not exist. This follows from Theorem 8.1.3 and Corollary 8.2.2. This completes the discussion of power series and shows that they are only understandable in the complex plane.

The two kinds of functions of greatest interest in algebra are polynomials and rational functions. The two kinds of interest in complex analysis are analytic functions and meromorphic functions, the latter being a generalization of rational functions just as analytic functions are generalizations of polynomials.

There was nothing in the above argument for the Cauchy integral formula which depended on γ_R^* being a circle.

Corollary 11.2.12 *Suppose γ^* is a C^1 closed curve as described above in Definition 11.1.2. Assume it divides the plane into two open sets such that γ^* is the boundary of each, the inside being the bounded open set. Assume γ^* along with its inside is contained in a convex open set on which a function f is differentiable. Then the conclusion of Theorem 11.2.9 is still valid for a suitable orientation of γ^*. This orientation will be called the counter clockwise orientation.*

Proof: Letting z be the point on the inside of γ^*, consider a small circle as shown above containing z. Orient this small circle in the clockwise direction. Obtain the four contours as in the proof of Theorem 11.2.9 by extending the straight lines till they intersect γ^* for the first time. Then this determines the desired orientation on each of the contours similar to the ones in Theorem 11.2.9. The rest of the argument is the same. There is a primitive for f on each of these contours because f' exists on a convex set containing the contour and its inside, so the integral over the contour is zero. Then adding these together, it is a repeat of the proof of Theorem 11.2.9. ■

Obviously these theorems are not the best possible. For general versions of this, see Rudin's book Real and Complex Analysis or my on line book Calculus of Real and Complex Variables. The latter treatment depends on a very general Green's theorem and uses the Jordan curve theorem. What is given here is sufficient for the applications of interest in this book.

11.3 Isolated Singularities

This is about functions which are analytic near a point z_0 but possibly not analytic at the point. This point z_0 is called an isolated singularity.

Lemma 11.3.1 *Suppose f is analytic on $\hat{B} \equiv B(z_0,r) \setminus \{z_0\}$. Then f can be defined at z_0 such that the resulting function is analytic on $B(z_0,r)$ if and only if $\lim_{z\to z_0}(z-z_0) f(z) = 0$. Such a z_0 is called a removable singularity.*

Proof: It is clear that if $f(z_0)$ can be chosen to make the function analytic then it follows that $\lim_{z\to z_0}(z-z_0) f(z) = 0$.

Suppose then that this limit condition holds. Consider $h(z) \equiv (z-z_0)^2 f(z), h(z_0) \equiv 0$. Then

$$h'(z_0) \equiv \lim_{z\to 0} \frac{h(z)}{z-z_0} = \lim_{z\to 0} \frac{(z-z_0)^2 f(z)}{(z-z_0)} = \lim_{z\to 0}(z-z_0) f(z) = 0$$

Thus $h(z)$ is analytic near z_0 with a power series of the form $\sum_{k=2}^{\infty} a_k (z-z_0)^k$ and so $f(z) = \sum_{k=2}^{\infty} a_k (z-z_0)^{k-2}$ for all $z \neq z_0$ and hence we can take $f(z_0) \equiv a_2$ and the resulting function is given by a power series and is therefore, analytic by Theorem 8.2.1. ■

Theorem 11.3.2 *(Casorati Weierstrass) Suppose f is analytic near z_0 and for $\hat{B} \equiv B(z_0,r) \setminus \{z_0\}$, $f(\hat{B})$is NOT dense in $\mathbb{C}$. This means there is w where $B(w,\delta)$ has no points of $f(\hat{B})$. Then near z_0 there is an analytic function $g(z)$ such that for z near z_0, $f(z) = g(z) + \sum_{k=1}^{m} \frac{b_k}{(z-z_0)^k}$. In words, this says that f has a pole at z_0 or else is equal to an analytic function near z_0 unless $f(\hat{B})$ is dense in $\mathbb{C}$.*

Proof: It is clearly true if $\lim_{z\to z_0}(z-z_0) f(z) = 0$ from the above lemma. This is the case where z_0 is removable, and in this case, there is no sum $\sum_{k=1}^{m} \frac{b_k}{(z-z_0)^k}$, just an analytic function $g(z)$.

Now suppose $B(w,r)$ contains no points of $f(\hat{B})$. Then consider $\frac{1}{f(z)-w}$ which is analytic near z_0. For z close enough to $z_0, |f(z)-w|$ is larger than some δ since otherwise, there would be $z_n \to z_0, f(z_n) \to w$ so from Lemma $f(z_n) \in B(w,r)$ if n sufficiently large. Hence $\lim_{z\to z_0}(z-z_0)\frac{1}{f(z)-w} = 0$ and so from Lemma 11.3.1, $\frac{1}{f(z)-w} = h(z)$ where h is analytic near z_0. Say $h(z) = \sum_{k=0}^{\infty} a_k (z-z_0)^k$.

Case 1. $a_0 \neq 0$. Then for z close to $z_0, h(z)^{-1} = \frac{1}{a_0} + \sum_{k=1}^{\infty} b_k (z-z_0)^k$ since $h(z)^{-1}$ has a derivative, so

$$f(z) = w + \frac{1}{a_0} + \sum_{k=1}^{\infty} b_k (z-z_0)^k = g(z) \text{ analytic}$$

which shows the result in this case.

Case 2. $h(z) = \sum_{k=m}^{\infty} a_k (z-z_0)^k, m > 1$ where a_m is the first nonzero a_j. Then in this case,

$$\begin{aligned}
\frac{h(z)}{(z-z_0)^m} &= \sum_{k=m}^{\infty} a_k (z-z_0)^{k-m}, \\
\frac{(z-z_0)^m}{h(z)} &= (z-z_0)^m (f(z)-w) = \frac{1}{a_m} + \sum_{k=1}^{\infty} c_k (z-z_0)^k,
\end{aligned}$$

for some c_k. Thus $f(z) - w = \frac{1}{(z-z_0)^m}\left(\frac{1}{a_m} + \sum_{k=1}^{\infty} c_k (z-z_0)^k\right) = g(z) + \sum_{k=1}^{m} \frac{b_k}{(z-z_0)^k}$ for some analytic function $g(z)$. ■

This shows that if z_0 is an isolated singularity, then unless something really bizarre happens, ($f(\hat{B})$ dense in $\mathbb{C}$) the function has a pole at the singularity or is equal to an analytic function. How bizarre? If the isolated singularity is not a pole, then $f^{-1}(\beta) \cap \hat{B}$ is infinite for every $\beta \in \mathbb{C}$ with maybe one exception. This is Picard's theorem and is an interesting known result but to see this proved see a more advanced book on complex analysis like Conway [12], Page 300.

11.4 The Logarithm

First of all, we define for $z = x + iy, e^z \equiv e^x (\cos(y) + i\sin(y))$. This agrees with e^x when $y = 0$. Then it is routine to verify that the usual rules of exponents apply. That $\frac{d}{dz}(e^z) = e^z$, let $h = h_1 + ih_2$ and using the power series for $\cos, \sin$,

$$\begin{aligned}
e^{z+h} - e^z &= e^z \left(e^{h_1}(\cos(h_2) + i\sin(h_2)) - 1\right) = e^z \left(e^{h_1} - 1\right) + ie^z e^{h_1} h_2 + o(h) \\
&= e^z h_1 + ie^z e^{h_1} h_2 + o(h) = e^z h + o(h)
\end{aligned}$$

Next I want to define a logarithm which is the inverse of this function. You want to have $e^{\log(z)} = z = |z|(\cos\theta + i\sin\theta)$ where θ is the angle of z. Now $\log(z)$ should be a complex number and so it will have a real and imaginary part. Thus

$$e^{\operatorname{Re}(\log(z)) + i\operatorname{Im}(\log(z))} = |z|(\cos\theta + i\sin\theta) \tag{11.3}$$

where θ is the angle of z. The magnitude of the left side needs to equal the magnitude of the right side. Hence, $e^{\operatorname{Re}(\log(z))} = |z|$ and so it is clear that $\operatorname{Re}(\log(z)) = \ln|z|$. Note that we must exclude $z = 0$ just as in the real case. What about $\operatorname{Im}(\log(z))$? Having found $\operatorname{Re}(\log(z))$, 11.3 is

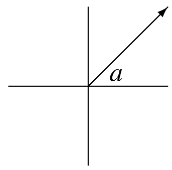
$$|z|(\cos(\operatorname{Im}(\log z)) + i\sin(\operatorname{Im}(\log z))) = |z|(\cos\theta + i\sin\theta) \tag{11.4}$$

which happens if and only if

$$\operatorname{Im}(\log z) = \theta + 2k\pi \tag{11.5}$$

for k an integer. Thus there are many solutions for $\operatorname{Im}(\log z)$ to the above problem. A branch of the logarithm is determined by picking one of them. The idea is that there is only one possible solution for $\operatorname{Im}(\log z)$ in any open interval of length 2π because if you have two different k in 11.5, the two values of $\operatorname{Im}(\log z)$ would differ by at least 2π so they could not both be in an **open** interval of length 2π.

What is done is to consider e^z where if $z = |z| e^{i\theta}$, then $\theta \in (a-\pi, a+\pi)$ for some a. In other words, you consider the ray coming from 0 in the complex plane and including 0 which has angle a. Then regard e^z as being defined for all of $\mathbb{C}$ other than this ray.



This involves restricting the domain of the function to an open set so that it has an inverse. It is like what was done for arctan and other trig. functions, except here we are careful to have the domain be an open set. Then if this restriction is made, there is exactly one solution $\operatorname{Im}(\log z)$ to 11.4. The most common assignment of a is π, so we leave out the negative real axis. However, one could leave out any other ray. If the usual one is left out, this shows that we need to have $\log(z) = \ln(|z|) + i \arg(z)$ where $\arg(z)$ is the angle of z which is in $(-\pi, \pi)$. It is called the principal branch of the logarithm when this is done. If you left out some other ray, then $\arg(z)$ would refer to an angle in some other open interval of length 2π.

Now the above geometric description shows that $z \to \log(z)$ is continuous. Indeed, if $z_n \to z$, then by the triangle inequality,

$$||z_n| - |z|| \leq |z_n - z|$$

and so by continuity of ln, you get $\ln(|z_n|) \to \ln(|z|)$. As to convergence of $\arg(z_n)$ to $\arg(z)$, just note that saying one is close to another is the same as saying that $\arg(z_n)$ is in any open set determined by two rays emanating from 0 which include z. This happens if $z_n \to z$. Is $z \to \log(z)$ differentiable? First recall that $(e^z)' = e^z$ and so

$$h = e^{\log(z+h)} - e^{\log(z)} = e^{\log(z)} (\log(z+h) - \log(z)) + o(\log(z+h) - \log(z)) \tag{11.6}$$

$$\frac{h}{z} = \log(z+h) - \log(z) + o(\log(z+h) - \log(z)) \tag{11.7}$$

By continuity, if h is small enough,

$$|o(\log(z+h) - \log(z))| < \frac{1}{2} |\log(z+h) - \log(z)|$$

Hence $\left|\frac{h}{z}\right| \geq \frac{1}{2} |\log(z+h) - \log(z)|$. This shows that $\frac{|\log(z+h) - \log(z)|}{|h|} \leq \frac{2}{|z|}$. Now

$$\frac{|o(|\log(z+h) - \log(z)|)|}{|h|} = \frac{o(|\log(z+h) - \log(z)|)}{|\log(z+h) - \log(z)|} \frac{|\log(z+h) - \log(z)|}{|h|}$$

and the second term on the right is bounded while the first converges to 0 as $h \to 0$. Therefore, $o(\log(z+h) - \log(z)) = o(h)$ and so it follows from 11.7,

$$\log(z+h) - \log(z) = \left(\frac{1}{z}\right) h + o(h)$$

which shows that, just as in the real variable case $\log'(z) = \frac{1}{z}$.

Definition 11.4.1 *For $a \in \mathbb{R}$, let l be the ray from 0 in the complex plane which includes 0 and consider all complex numbers D_a whose angle is in $(a-\pi, a+\pi)$ and not 0.*

$$\log(z) = \ln(|z|) + i\arg(z)$$

where $\arg(z)$ is the angle for z which is in $(a-\pi, a+\pi)$. This function is one to one and analytic on D_a and $e^{\log(z)} = z$. This is called a branch of the logarithm. It is called the principal branch if the ray defining D_a consists of 0 along with the negative real axis.

Note that $\log(D_a)$, is the open set in $\mathbb{C}$ defined by $\operatorname{Im} z \in (a-\pi, a+\pi)$. Thus there is a one to one and onto analytic map which maps D_a onto

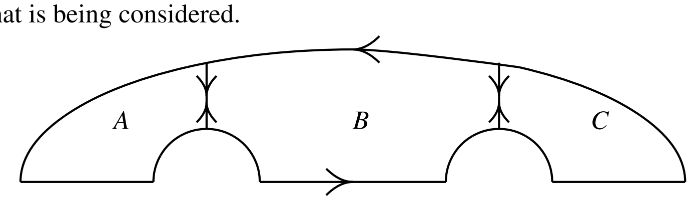
$$\{z \in \mathbb{C} : \operatorname{Im} z \in (a-\pi, a+\pi)\}.$$

This book is not about a detailed study of such conformal maps, (analytic functions with values in $\mathbb{C}$ are called this) but this is an interesting example. Some people find these kind of mappings very useful and they are certainly beautiful when you keep track of level curves of real and imaginary parts. You can have lots of fun by having Matlab graph real and imaginary parts, but this is about functions of two variables so is outside the scope of this book.

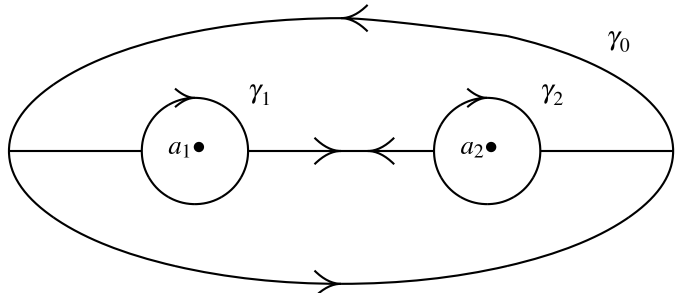
11.5 The Method of Residues

Next is a consideration of various improper and difficult integrals using contour integrals. To do this in maximal generality you should develop the winding number and the Cauchy integral formula for cycles. However, when I do it this way, I sometimes end up wondering what exactly has been shown and how it relates to specific examples. Ultimately in this book, the consideration of examples is of interest more than the most general formulation of the theory.

Here the thing of interest is a function analytic in some open connected set except for finitely many points. The following picture will describe the kind of thing which is meant. The various piecewise smooth closed curves will satisfy the conclusion of the Cauchy integral theorem. Recall that this will be so if the closed curve is contained in a star shaped set on which the function has a derivative. Consider the following pictures which are a paradigm of what is being considered.



You can see from the picture that if you fattened up the contours around the regions A, B, C you would have star shaped open sets containing these contours. Thus, if $f'(z)$ exists for z in some open set containing these contours, then the contour integrals over these three would be 0 and the integrals over the straight vertical lines cancel, so the contour integral over the curve consisting of horizontal lines and half circles with an elliptical curve on top is also 0. Now consider two of these pasted together as shown below.



Assume $f'(z)$ exists for all z in some open set which includes all the contours and their insides other than at the two points shown, a_1 and a_2. Then applying the above to the second piecewise smooth closed curve, we find that on adding the contour integrals over the top and the bottom, yields that the contour integrals over the straight lines cancel and so what finally results is

$$\int_{\gamma_0} f(z)\,dz + \int_{-\gamma_1} f(z)\,dz + \int_{-\gamma_2} f(z)\,dz = 0$$

In other words,

$$\int_{\gamma_0} f(z)\,dz = \int_{\gamma_1} f(z)\,dz + \int_{\gamma_2} f(z)\,dz$$

where γ_i for $i = 1,2$ are oriented counter clockwise.

Obviously you could have many more exceptional points a_i and small circular contours surrounding these, and the same argument would work. This example and its obvious generalizations is a paradigm for the procedure of evaluating contour integrals with the method of residues. First is a definition of what is meant by a pole and a residue.

Definition 11.5.1 *A function f has a pole at a if*

$$f(z) = g(z) + \sum_{k=1}^{n} \frac{b_k}{(z-a)^k}$$

where $b_n \neq 0$ and g is analytic near a. The order of the pole is n. Denote by $S(a,z)$ the term described by the sum. It is called the singular part of f at a. The residue at a is defined to be b_1 and is denoted as $\operatorname{res}(f,a)$. *See Theorem 11.3.2 to see how this takes place.*

How can we find it? Multiply by $(z-a)^n$. This gives

$$f(z)(z-a)^n = g(z)(z-a)^n + b_1(z-a)^{n-1} + \sum_{k=2}^{n} b_k (z-a)^{n-k}$$

Now take the derivative $n-1$ times and then take a limit as $z \to a$. The differentiation will zero out all the terms in the sum on the right. Then the limit will zero out all other terms except the b_1 term which will be $(n-1)!b_1$. This justifies the following procedure.

Procedure 11.5.2 *When f has a pole at a of order n, to find the residue, multiply by $(z-a)^n$, take $n-1$ derivatives, and finally take the limit as $z \to a$. The residue will be this number b divided by $(n-1)!$. Thus*

$$\operatorname{res}(f,a) = \frac{1}{(n-1)!} \lim_{z \to a} \frac{d^{n-1}}{dz^{n-1}} \left(f(z)(z-a)^n \right)$$

In taking the limit, you can use L'Hospital's rule provided you have an indeterminate form because you know the limit exists. Thus you could take a limit along a one dimensional line $a+t$ as $t \to 0$ and you can apply L'Hospital's rule to the real and imaginary parts. Thus, you can get the answer by pretending z is a real variable and using the usual techniques for functions of a real variable.

I want to consider contours like the above and functions which are of the form

$$f(z) = g(z) + \sum_{j=1}^{m} S(a_j, z) \tag{11.8}$$

where $g(z)$ is analytic and the $S(a_j, z)$ are the terms which yield a pole at a_j. Consider γ centered at a with radius r

$$\int_\gamma S(a_j, z)\, dz = \sum_{k=1}^{p} \int_\gamma \frac{b_k}{(z-a_j)^k} dz.$$

One of these terms is

$$\int_\gamma \frac{b_k}{(z-a_j)^k} dz = b_k \int_0^{2\pi} \frac{1}{r^k e^{ikt}} ire^{it} dt = \begin{cases} 0 \text{ if } k > 1 \\ b_1 i \int_0^{2\pi} dt = b_1 2\pi i \text{ if } k = 1 \end{cases}$$

Thus, if the $(z-a_j)^{-1}$ coefficient in 11.8 is b_1^j,

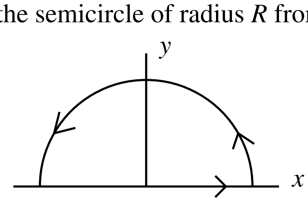
$$\begin{aligned} \int_{\gamma_0} f(z)\, dz &= \sum_{j=1}^{m} \int_{\gamma_j} g(z)\, dz + \sum_{j=1}^{m} \int_{\gamma_j} S(a_j, z)\, dz \\ &= \sum_{j=1}^{m} \int_{\gamma_j} \frac{b_1^j}{z-a_j} dz = 2\pi i \operatorname{res}(f, a_j) \end{aligned}$$

This is the residue method. You can use it to compute very obnoxious improper integrals and this is illustrated next.

Letting $p(x), q(x)$ be polynomials, you can use the above method of residues to evaluate obnoxious integrals of the form

$$\int_{-\infty}^{\infty} \frac{p(x)}{q(x)} dx \equiv \lim_{R\to\infty} \int_{-R}^{R} \frac{p(x)}{q(x)} dx$$

provided the degree of $p(x)$ is two less than the degree of $q(x)$ and the zeros of $q(z)$ involve $\operatorname{Im}(z) > 0$. Of course if the degree of $p(x)$ is larger than that of $q(x)$, you would do long division. The contour to use for such problems is γ_R which goes from $(-R, 0)$ to $(R, 0)$ along the real line and then on the semicircle of radius R from $(R, 0)$ to $(-R, 0)$.



Letting C_R be the circular part of this contour, for large R,

$$\left| \int_{C_R} \frac{p(z)}{q(z)} dz \right| \le \pi R \frac{CR^k}{R^{k+2}}$$

which converges to 0 as $R \to \infty$. Therefore, it is only a matter of taking large enough R to enclose all the roots of $q(z)$ which are in the upper half plane, finding the residues at these points and then computing the contour integral. Then you would let $R \to \infty$ and the part of the contour on the semicircle will disappear leaving the Cauchy principal value integral which is desired. There are other situations which will work just as well. You simply need to have the case where the integral over the curved part of the contour converges to 0 as $R \to \infty$.

Here is an easy example.

Example 11.5.3 *Find* $\int_{-\infty}^{\infty} \frac{1}{x^2+1} dx$

You know from calculus that the answer is π. Lets use the method of residues to find this. The function $\frac{1}{z^2+1}$ has poles at i and $-i$. We don't need to consider $-i$. It seems clear that the pole at i is of order 1 and so all we have to do is take

$$\lim_{z\to i} \frac{x-i}{1+x^2} = \frac{1}{(x-i)(x+i)}(x-i) = \frac{1}{2i}$$

Then the integral equals $2\pi i \left(\frac{1}{2i}\right) = \pi$.

That one is easy. Now here is a genuinely obnoxious integral.

Example 11.5.4 *Find* $\int_{-\infty}^{\infty} \frac{1}{1+x^4} dx$

It will have poles at the roots of $1+x^4$. These roots are

$$\left(\frac{1}{2}-\frac{1}{2}i\right)\sqrt{2}, -\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}, -\left(\frac{1}{2}-\frac{1}{2}i\right)\sqrt{2}, \left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}$$

Using the above contour, we only need consider

$$-\left(\frac{1}{2}-\frac{1}{2}i\right)\sqrt{2}, \left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}$$

Since they are all distinct, the poles at these two will be of order 1. To find the residues at these points, you would need to take

$$\lim_{z\to-\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}} \frac{\left(z-\left(-\left(\frac{1}{2}-\frac{1}{2}i\right)\sqrt{2}\right)\right)}{1+z^4}, \lim_{z\to\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}} \frac{\left(z-\left(\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}\right)\right)}{1+z^4}$$

factoring $1+x^4$ and computing the limit, you could get the answer. However, it is easier to apply L'Hospital's rule to identify the limit you know is there,

$$\lim_{z\to-\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}} \frac{1}{4z^3} = \left(\frac{1}{8}-\frac{1}{8}i\right)\sqrt{2}, \lim_{z\to\left(\frac{1}{2}+\frac{1}{2}i\right)\sqrt{2}} \frac{1}{4z^3} = -\left(\frac{1}{8}+\frac{1}{8}i\right)\sqrt{2}$$

Then the contour integral is

$$2\pi i\left(\left(\frac{1}{8}-\frac{1}{8}i\right)\sqrt{2}\right) + 2\pi i\left(-\left(\frac{1}{8}+\frac{1}{8}i\right)\sqrt{2}\right) = \frac{1}{2}\sqrt{2}\pi$$

You might observe that this is a lot easier than doing the usual partial fractions and trig substitutions etc. Now here is another tedious example.

Example 11.5.5 *Find* $\int_{-\infty}^{\infty} \frac{x+2}{(x^2+1)(x^2+4)^2} dx$

The poles of interest are located at $i, 2i$. The pole at $2i$ is of order 2 and the one at i is of order 1. In this case, the partial fractions expansion is

$$\frac{\frac{1}{9}x+\frac{2}{9}}{x^2+1} - \frac{\frac{1}{3}x+\frac{2}{3}}{(x^2+4)^2} - \frac{\frac{1}{9}x+\frac{2}{9}}{x^2+4}$$

You could do these integrals by elementary methods. However, I will consider the original problem by finding $2\pi i$ times the sum of the residues.

The pole at i would be

$$\lim_{z\to i} \frac{\left(\frac{1}{9}z+\frac{2}{9}\right)(z-i)}{(z+i)(z-i)} = \frac{\left(\frac{1}{9}i+\frac{2}{9}\right)}{(i+i)} = \frac{1}{18} - \frac{1}{9}i$$

Now consider the pole at $2i$ which is a pole of order 2. Using Procedure 11.5.2, it is

$$\lim_{z\to 2i} \frac{d}{dz}\left(\frac{(z-2i)^2(z+2)}{(z^2+1)(z^2+4)^2}\right) = \lim_{z\to 2i} \frac{d}{dz}\left(\frac{(z+2)}{(z^2+1)(z+2i)^2}\right)$$

$$\begin{aligned} &= \lim_{z\to 2i}\left(-\frac{1}{(z^2+1)^2(z+2i)^3}\left(3z^3+(8+2i)z^2+(1+8i)z+4-2i\right)\right) \\ &= -\frac{1}{18}+\frac{11}{144}i \end{aligned}$$

Integral over a large semicircle is

$$2\pi i\left(-\frac{1}{18}+\frac{11}{144}i\right)+2\pi i\left(\frac{1}{18}-\frac{1}{9}i\right) = \frac{5}{72}\pi$$

Letting $R\to\infty$, this is the desired improper integral. More precisely, it is the Cauchy principal value integral, $\lim_{R\to\infty}\int_{-R}^{R} \frac{x+2}{(x^2+1)(x^2+4)^2} dx$. In this case, it is a genuine improper integral.

Sometimes you don't blow up the curves and take limits. Sometimes the problem of interest reduces directly to a complex integral over a closed curve. The integral of rational functions of cosines and sines lead to this kind of thing. Here is an example of this.

Example 11.5.6 *The integral is* $\int_0^{\pi} \frac{\cos\theta}{2+\cos\theta} d\theta$.

This integrand is even and so it equals $\frac{1}{2}\int_{-\pi}^{\pi} \frac{\cos\theta}{2+\cos\theta} d\theta$. For z ousn the unit circle, $z = e^{i\theta}$, $\bar{z} = \frac{1}{z}$ 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 this is done in order to get a contour integral which reduces to the one of interest. It follows that a contour integral which reduces to the integral of interest is

$$\frac{1}{2i}\int_{\gamma} \frac{\frac{1}{2}\left(z+\frac{1}{z}\right)}{2+\frac{1}{2}\left(z+\frac{1}{z}\right)} \frac{dz}{z} = \frac{1}{2i}\int_{\gamma} \frac{z^2+1}{z(4z+z^2+1)} dz$$

where γ is the unit circle oriented counter clockwise. 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,

$$\operatorname{res}(f,0)=\lim_{z\to 0} z\left(\frac{z^2+1}{z(4z+z^2+1)}\right)=1.$$

$$\operatorname{res}\left(f,-2+\sqrt{3}\right)=$$

$$\lim_{z\to -2+\sqrt{3}}\left(z-\left(-2+\sqrt{3}\right)\right)\frac{z^2+1}{z(4z+z^2+1)}=-\frac{2}{3}\sqrt{3}.$$

It follows

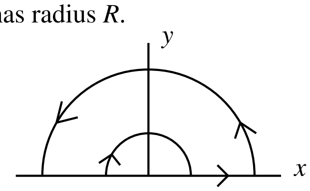
$$\begin{aligned}\int_0^{\pi}\frac{\cos\theta}{2+\cos\theta}d\theta &= \frac{1}{2i}\int_{\gamma}\frac{z^2+1}{z(4z+z^2+1)}dz=\frac{1}{2i}2\pi i\left(1-\frac{2}{3}\sqrt{3}\right)\\ &= \pi\left(1-\frac{2}{3}\sqrt{3}\right).\end{aligned}$$

Other rational functions of the trig functions will work out by this method also.

Sometimes you have to be clever about which version of an analytic function you wish to use. The following is such an example.

Example 11.5.7 *The integral here is* $\int_0^{\infty}\frac{\ln x}{1+x^4}dx$.

It is natural to try and use the contour in the following picture in which the small circle has radius r and the large one has radius R.



However, this will create problems with the log since the usual version of the log is not defined on the negative real axis. This difficulty may be eliminated by simply using another branch of the logarithm as discussed above. Leave out the ray from 0 along the negative y axis and use this example 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 $z=|z|e^{i\theta}$. Of course, with this contour, this will end up finding the integral $\int_{-\infty}^{\infty}\frac{\ln|x|}{1+x^4}dx$. Then the function used is $f(z)\equiv\frac{L(z)}{1+z^4}$. Now the only singularities contained in this contour 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 $\operatorname{res}\left(f,\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}\right)=$

$$\lim_{z\to\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}}\frac{\left(z-\left(\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}\right)\right)(\ln|z|+i\arg_1(z))}{1+z^4}$$

$$= \lim_{z \to \frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}} \frac{(\ln|z| + i\arg_1(z)) + \left(z - \left(\frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}\right)\right)(1/z)}{4z^3}$$

$$= \frac{\ln\left(\sqrt{\frac{1}{2} + \frac{1}{2}}\right) + i\frac{\pi}{4}}{4\left(\frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}\right)^3} = \left(\frac{1}{32} - \frac{1}{32}i\right)\sqrt{2}\pi$$

Similarly

$$\operatorname{res}\left(f, \frac{-1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}\right) = \frac{\ln\left(\sqrt{\frac{1}{2} + \frac{1}{2}}\right) + i\frac{3\pi}{4}}{4\left(-\frac{1}{2}\sqrt{2} + \frac{1}{2}i\sqrt{2}\right)^3} = \frac{3}{32}\sqrt{2}\pi + \frac{3}{32}i\sqrt{2}\pi.$$

Of course it is necessary to consider the integral along the small semicircle of radius r. This reduces to

$$\int_\pi^0 \frac{\ln|r| + it}{1 + (re^{it})^4}\left(rie^{it}\right) dt$$

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 $R \to \infty$,

$$e(R) + 2 \lim_{r \to 0+} \int_r^R \frac{\ln t}{1 + t^4} dt + i\pi \int_{-\infty}^0 \frac{1}{1 + t^4} dt = \left(-\frac{1}{8} + \frac{1}{4}i\right)\pi^2\sqrt{2}$$

where $e(R) \to 0$ as $R \to \infty$. 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 $R \to \infty$,

$$2\int_0^\infty \frac{\ln t}{1 + t^4} dt = \left(-\frac{1}{8} + \frac{1}{4}i\right)\pi^2\sqrt{2} - i\pi\left(\frac{\sqrt{2}}{4}\pi\right) = -\frac{1}{8}\sqrt{2}\pi^2,$$

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 only real analysis. I don't have any idea how to get this one from standard methods of calculus. Perhaps some sort of partial fractions might do the job but even if so, it would be very involved.

11.6 Counting Zeros, Open Mapping Theorem

The open mapping theorem is perhaps a little digression from the main emphasis of this book, but it is such a marvelous result, that it seems a shame not to include it. It comes from a remarkable formula which counts the number of zeros of an analytic function inside a ball.

Definition 11.6.1 *Let U be a nonempty open set in $\mathbb{C}$. Then it is called connected if for any $z, w \in U$, there exists a continuous one to one piecewise linear $\gamma : [0,1] \to U$ such that $\gamma(0) = z$ and $\gamma(1) = w$. A connected open set will be called a region in this section.*

By Theorem 6.5.8 the above implies the usual definition of a connected set. To go the other way, suppose the usual definition and let $z \in U$ be given. Let S denote those points reachable by a continuous one to one piecewise linear curve from z. Show S is open. Now show that those points of U which are not reachable by such a curve is also open. Thus U is the disjoint union of open sets. One of them must be empty if U is connected according to the usual definition. This is really very easy if you use convexity of balls. It is not the purpose of this book to belabor higher dimensional considerations and the above definition is sufficiently descriptive.

Here is a very useful equivalence.

Proposition 11.6.2 *Let U be a region as just defined and suppose f is an analytic function defined on U. Then the following are equivalent*

1. *There exists $z \in U$ such that $f^{(n)}(z) = 0$ for all $n = 0, 1, 2, \ldots$.*
2. *$f(z) = 0$ for all $z \in U$.*

Proof: 1.$\Rightarrow$ 2. Let z be as in 1. Then, since U is open, $\overline{B(z,r)} \subseteq U$ for small enough positive r. It follows from Corollary 11.2.11 that for w in this ball,

$$f(w) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z)}{k!} (w-z)^k = 0$$

and so f is zero on $B(z,r)$. This shows that near z, $f^{(n)}(w) = 0$ for all n an integer larger than or equal to 0. Letting S denote $z \in U$ such that $f^{(n)}(z) = 0$ for all n, this shows that S is an open subset of U. If S is all of U then this was what was to be shown. Otherwise, let $z \in S$ and $w \in U \setminus S$. Let $\gamma(t)$ go from z to $w, \gamma(0) = z, \gamma(1) = w$. Then let $T \equiv \sup\{t \in [0,1] : \gamma(s) \in S \text{ for } s \leq t\}$ and suppose $T < 1$. Then let t_n be an increasing sequence converging to T, $0 = f(\gamma(t_n)) \to f(\gamma(T))$ and so $f(\gamma(T)) = 0$. However, each derivative of f is continuous also and so the same reasoning shows that $f^{(n)}(\gamma(T)) = 0$ for each $n \geq 1$. Hence $\gamma(T) \in S$. But this violates the definition of T because $\gamma(t) \in S$ for all $t \in [T, T+\varepsilon]$ for suitably small ε due to what was just shown that S is open. Hence $T = 1$ and so $w \in S$. Thus $U \setminus S = \emptyset$.

2.$\Rightarrow$ 1. This is obvious. If $f(z) = 0$ for all z, then all derivatives are also 0 on the open set U. ■

Next is one more equivalence.

Theorem 11.6.3 *Suppose f is analytic on a region U (open and connected). The following are equivalent.*

1. *There exists $z \in U$ where $f^{(n)}(z) = 0$ for all n.*

2. *f is 0 on U.*

3. *The set of zeros of f has a limit point in U*

Proof: 1.$\Longleftrightarrow$ 2. The first two are equivalent by Proposition 11.6.2.

2.)$\Rightarrow$ 3.) This is obvious. Since f is 0 everywhere, all derivatives at every point are 0 so every point of $B(z_0, R)$ is a limit point of the set of zeros.

3.)$\Rightarrow$ 1.)Suppose there exists a limit point $z \in U$ of the set of zeros. I will show that $f^{(n)}(z) = 0$ for all n. By continuity $f(z) = 0$. Since z is in U, there exists $r > 0$ such that $B(z,r) \subseteq U$. By Corollary 11.2.11, there are complex numbers a_k such that for $w \in B(z,r)$,

$$f(w) = \sum_{k=1}^{\infty} a_k (w-z)^k .$$

If each $a_k = 0$, then at this z, all derivatives of f equal 0. Otherwise,

$$f(w) = (w-z)^m g(w) = (w-z)^m \sum_{k=m}^{\infty} a_k (w-z)^{k-m},\ m > 0$$

where $a_m \neq 0$. I will show this does not happen. From the above, there exists a sequence of distinct z_n converging to z where $f(z_n) = 0$. Then $0 = f(z_n) = (z_n - z)^m g(z_n)$ so $g(z_n) = 0$. By continuity, $g(z) = \lim_{n\to\infty} g(z_n) = 0$ which requires $a_m = 0$ after all. Thus all $a_k = 0$ after all, a contradiction. It follows that all three conditions are equivalent. ■

The counting zeros theorem is as follows:

Theorem 11.6.4 *Let f be analytic in an open set containing the closed disk*

$$D(z_0, r) \equiv \{z : |z - z_0| \leq r\}$$

and suppose f has no zeros on the circle $C(z_0, r)$, the boundary of $D(z_0, r)$. Then the number of zeros of f counted according to multiplicity which are contained in $D(z_0, r)$ is $\frac{1}{2\pi i}\int_{C(z_0,r)} \frac{f'(z)}{f(z)} dz$ where $C(z_0, r)$ is oriented in the counter clockwise direction.

Proof: There are only finitely many zeros in $D(z_0, r)$. Otherwise, there would exist a limit point of the set of zeros z. If z is in $B(z_0, r)$, then by Theorem 11.6.3, $f = 0$ on $D(z_0, r)$. If it is on $C(z_0, r)$, this would contradict having no zeros on the boundary.

Let these zeros be $\{z_1, ..., z_p\}$. Consider z_k and suppose it is a zero of multiplicity m so $f(z) = (z - z_k)^m g(z)$ where $g(z_k) \neq 0, m \geq 1$. Then

$$\frac{f'(z)}{f(z)} = \frac{m(z-z_k)^{m-1} g(z) + (z-z_k)^m g'(z)}{(z-z_k)^m g(z)} = \frac{m}{(z-z_k)} + \frac{g'(z)}{g(z)}$$

The second term is analytic near z_k and so the residue of $\frac{f'(z)}{f(z)}$ is m the number of times z_k is repeated in the list of zeros. Hence doing this for each of the zeros, gives $2\pi i(p) = \int_{C(z_0,r)} \frac{f'(z)}{f(z)} dz$ and so $p = \frac{1}{2\pi i}\int_{C(z_0,r)} \frac{f'(z)}{f(z)} dz$. ■

Obviously the above theorem applies to more general regions than disks, but the main interest tends to be for balls. Also it generalizes to the situation where there are no poles or zeros on $C(z_0, r)$ and finitely many zeros and poles in $B(z_0, r)$. In this case, you get a count of the number of zeros minus the number of poles. This more general theorem is called Rouche's theorem.

Lemma 11.6.5 *Suppose U is a region and $g : U \to \mathbb{N}$ is continuous. Then g is constant on U.*

Proof: Let $z, w \in U$. Let $h(t) = g(\gamma(t))$ for $t \in [0,1]$ where γ is a smooth curve with $\gamma(0) = z$ and $\gamma(1) = z$. Then by the intermediate value theorem, h can have only one value. Thus $g(z) = g(w)$. ■

Theorem 11.6.6 *(Open mapping theorem) Let Ω be a region (open connected set) in $\mathbb{C}$ and suppose $f : \Omega \to \mathbb{C}$ is analytic. Then $f(\Omega)$ is either a point or a region. In the case where $f(\Omega)$ is a region, it follows that for each $z_0 \in \Omega$, there exists an open set V containing z_0 and $m \in \mathbb{N}$ such that for all $z \in V$,*

$$f(z) = f(z_0) + \phi(z)^m \tag{11.9}$$

where $\phi : V \to B(0, \delta)$ is one to one, analytic and onto, $\phi(z_0) = 0$, $\phi'(z) \neq 0$ on V.

Proof: Suppose $f(\Omega)$ is not a point. Then for $z_0 \in \Omega$ 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,

$$\{z \in \Omega : f(z) - f(z_0) = 0\}$$

which would imply from Theorem 11.6.3 that $f(z) = f(z_0)$ for all $z \in \Omega$. Therefore, making r smaller if necessary, and using the power series of f,

$$f(z) = f(z_0) + (z - z_0)^m g(z) \stackrel{?}{=} (f(z_0) + \left((z - z_0) g(z)^{1/m}\right)^m)$$

where g is analytic near z_0 and $g(z_0) \neq 0$. Does an analytic $g(z)^{1/m}$ exist? By continuity, $g(B(z_0, r)) \subseteq B(g(z_0), \varepsilon)$ where ε is small enough that $0 \notin B(g(z_0), \varepsilon)$, so there exists a branch of the logarithm on $\mathbb{C} \diagdown B(g(z_0), \varepsilon)$. Call it log even though it might not be the principle branch. Then consider $e^{(1/m)\log(g(z))} \equiv g(z)^{1/m}$ and so we can obtain an analytic function denoted by $g(z)^{1/m}$ as in the above formula. Let $\phi(z) = (z - z_0) g(z)^{1/m}$. Then $\phi(z_0) = 0$ and

$$\phi'(z) = e^{(1/m)\log(g(z))} + (z - z_0) e^{(1/m)\log(g(z))} \frac{1}{g(z)} g'(z)$$

so $\phi'(z_0) = e^{(1/m)\log(g(z_0))} \neq 0$. Shrinking r some more if necessary, assume $\phi'(z) \neq 0$ for all $z \in B(z_0, r)$. The representation

$$f(z) = f(z_0) + \phi(z)^m, z \in B(z_0, r)$$

where $\phi'(z) \neq 0$ for all $z \in B(z_0, r)$ and $\phi(z_0) = 0$ has been obtained.

Let δ be small enough that the only zero of $\phi(z) - \phi(z_0)$ is z_0 in $\overline{B(z_0, \delta)}$. If no such small positive δ exists, then the zeroes of $\phi(z) - \phi(z_0)$ would have a limit point and so ϕ would be a constant. This would force f to be constant also. Then $\phi(z_0) \notin \phi(C(z_0, \delta))$ and so if $|w - \phi(z_0)|$ is small enough, then $w \notin \phi(C(z_0, \delta))$ either. Thus there is $\varepsilon > 0$ with $B(\phi(z_0), \varepsilon) \cap \phi(C(z_0, \delta)) = \emptyset$. Consider for $w \in B(\phi(z_0), \varepsilon) = B(0, \varepsilon)$ the formula for counting zeroes.

$$\frac{1}{2\pi i} \int_{C(z_0, \delta)} \frac{\phi'(z)}{\phi(z) - w} dz$$

It is a continuous function of w and equals 1 at $0 = \phi(z_0)$ so, since it is integer valued, it equals 1 on all of $B(0,\varepsilon)$, but this is the number of zeroes of $\phi(z) - w$. Thus $\phi(B(z_0,\delta)) = B(0,\varepsilon)$. Hence, $\phi^m(B(z_0,\delta)) = B(0,\varepsilon^m)$. It follows that

$$f(B(z_0,\delta)) = f(z_0) + B(0,\varepsilon^m) = B(f(z_0),\varepsilon^m)$$

and so this shows that f maps small open balls to open balls. Thus $f(\Omega)$ is a connected open set. ■

11.7 Exercises

1. A fractional linear transformation is one of the form $f(z) = \frac{az+b}{cz+d}$ where $ad - bc \neq 0$ where a,b,c,d are in a field, say $\mathbb{R},\mathbb{Q},\mathbb{C}$. Let M denote the 2×2 invertible matrices having entries in the same field. Denote by F these fractional linear transformations. For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M$, let $\phi(A)(z) \equiv \frac{az+b}{cz+d}$. Show that $\phi(AB)(z) = \phi(A)\circ\phi(B)(z)$. Show that $\phi(I)(z) = z$ and that $\phi : M \to F$ is onto. Show $\phi(A)^{-1} = \phi\left(A^{-1}\right)$ so there is an easy way to invert such a fractional linear transformation. This problem is best for those who have had a beginning course in linear algebra.

2. The modular group consists of functions $f(z) = \frac{az+b}{cz+d}$ where a,b,c,d are integers and $ad - bc = 1$. Surprisingly, each of these has an inverse. $f^{-1}(z) = \frac{dz-b}{-cz+a}$. Verify that this is the case. This means $f^{-1}\circ f(z) = z$. Show also that if f,g are two of these, then $f\circ g$ is another one. This last part might be a little tedious without the above problem.

3. Suppose U is an open set in $\mathbb{C}$ and $f : U \to \mathbb{C}$ is analytic, ($f'(z)$ exists for $z \in U$). For $z = x+iy$, $f(z) = f(x+iy) = u(x,y) + iv(x,y)$, u, and v having real values. These are the real and imaginary parts of f. The partial derivative, u_x is defined by fixing y and considering only the variable x. $u_x(x,y) \equiv \lim_{h\to 0}\frac{u(x+h,y)-u(x,y)}{h}$. Other versions of this notation are similar. Thus partial derivatives are a one variable consideration. Show that the existence of $f'(z)$ implies the Cauchy Riemann equations. $u_x = v_y, u_y = -v_x$. **Hint:** In the difference quotient for finding $f'(z)$, use $h \to 0$ and then $ih \to 0$ for h real.

4. Suppose $t \to z(t) = x(t) + iy(t)$ and $t \to w(t) = \hat{x}(t) + i\hat{y}(t)$ are two smooth curves which intersect when $t = 0$. Then consider the two curves $t \to f(z(t))$ and $t \to f(w(t))$ where f is analytic near $z(0) = w(0)$. Show the cosine of the angle between the resulting two curves is the same as the cosine of the angle of the original two curves when $t = 0$. **Hint:** You should write $f(z) = u(x,y) + iv(x,y)$ and use the Cauchy Riemann equations and the chain rule. This problem is really dependent on knowing a little bit about functions of more than one variable so does not exactly fit in this book. It depends on remembering some elementary multivariable calculus. Also recall that the cosine of the angle between two vectors $\mathbf{u},\mathbf{v}$ is $(\mathbf{u}\cdot\mathbf{v})/|\mathbf{u}||\mathbf{v}|$. That analytic mappings preserve angles is very important to some people.

5. Suppose $f : \mathbb{C} \to \mathbb{C}$ has a derivative at every point and $f'(z) = 0$ for all z. Show that, as in the case of a real variable, $f(z)$ is a constant. Generalize to an arbitrary open connected set. **Hint.** Pick z_0 and for arbitrary z, consider $z_0 + t(z - z_0)$ where

$t \in [0,1]$. Now consider the function of a real variable $f(z_0 + t(z - z_0))$ and consider real and imaginary parts. You could apply the mean value theorem to these.

6. $\uparrow$ Suppose $f : \mathbb{C} \to \mathbb{C}$ and f' exists on $\mathbb{C}$. Such a function is called entire. Suppose f is bounded. Then show f must be constant. This is Liouville's theorem. Note how different this is than what we see for functions of a real variable. **Hint:** Pick z. Use formula 11.2 to describe $f'(z)$ where γ_R^* is a large circle including z on its inside. Thus, $f'(z) = \frac{1}{2\pi i}\int_{\gamma_R} \frac{f(w)}{(w-z)^2} dw$. Now use Theorem 11.1.6 to get an estimate for $|f'(z)|$ which is a constant times $1/R$. However, R is arbitrary. Hence $f'(z) = 0$. Now use the above problem.

7. $\uparrow$ Show that if $p(z)$ is a non-constant polynomial, then there exists z_0 such that $p(z_0) = 0$. **Hint:** If not, then $1/p(z)$ is entire. Just use the quotient rule to see it has a derivative. Explain why it is bounded and use Liouville's theorem to assert that then it is a constant which it obviously isn't. This is the shortest known proof of the fundamental theorem of algebra.

8. Let $f : \mathbb{C} \to \mathbb{C}$ be entire (has a derivative on all of $\mathbb{C}$) and suppose that
$$\max\{|f(z)| : |z| \leq R\} \leq CR^k.$$
Then show that $f(z)$ is actually a polynomial of degree k. **Hint:** Recall the formula for the derivative in terms of the Cauchy integral.

9. Let $D(0,1)$ be the closed unit disk and let f_n be analytic on and near D. Suppose also that $f_n \to f$ uniformly on $D(0,1)$. Show that f is also analytic on $B(0,1)$. If f is an arbitrary continuous function defined on $D(0,1)$, does it follow that there exists a sequence of polynomials which converges uniformly to f on $D(0,1)$? In other words, does the Weierstrass approximation theorem hold in this setting?

10. Suppose you have a sequence of functions $\{f_n\}$ analytic on an open set U. If they converge uniformly to a function f, show that f is also analytic on U.

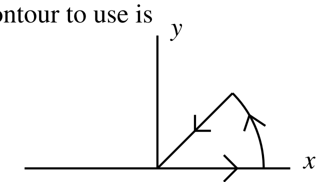
11. For $n = 1, 2, \ldots$ and a_n complex numbers, the Dirichlet series is $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$. Here s is a complex number. Thus $n^s \equiv \exp(s \log(n))$ where this refers to the principal branch of the logarithm. Show $\frac{1}{n^s} = \frac{1}{n^{\operatorname{Re} s}} e^{i \ln(n) \operatorname{Im}(s)}$. Then show that if the a_n are uniformly bounded, the Weierstrass M test applies and the Dirichlet series converges absolutely and uniformly for $\operatorname{Re} s \geq 1 + \varepsilon$ for any positive ε. Explain why the function of s is analytic for $\operatorname{Re}(s) > 1$. Obtain more delicate results using the Dirichlet partial summation formula in case $\sum_n |a_n| < \infty$. Note that if all $a_n = 1$, this function is $\xi(s)$ the Riemann zeta function whose zeros are a great mystery.

12. Suppose f' exists on $\mathbb{C}$ and $f(z_n) = 0$ with $z_n \to z$ where the z_n are distinct complex numbers. Show that then $f(z) = 0$ on $\mathbb{C}$. **Hint:** If not, then explain why $f(w) = (w-z)^m g(w)$ where $g(z) \neq 0$. Then $0 = f(z_n) = (z_n - z)^m g(z_n)$. Hence $g(z) = 0$, a contradiction. Thus argue the power series of f expanded about z is 0 and it converges to f on all of $\mathbb{C}$.

13. The power series for $\sin z, \cos z$ converge for all $z \in \mathbb{C}$. Using the above problem, explain why the usual trig. identities valid for z, w real continue to hold for all complex z, w. For example, $\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w)$. If you allow z

complex, explain why $\sin(z)$ cannot be bounded. **Hint:** Use the theorem about the zero set having a limit point on a connected open set.

14. The functions $z \to \sin\left(z^2\right), z \to \cos\left(z^2\right)$, and $z \to e^{iz^2}$ are all analytic functions since the chain rule continues to hold for functions of a complex variable. This problem is on the Fresnel integrals using contour integrals. In this case, there is no singular part of the function. The contour to use is



Then using this contour and the integral $\int_0^\infty e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$, explain why

$$
\begin{aligned}
0 &= \int_{\gamma_r} e^{iz^2} dz + \int_0^r e^{ix^2} dx - \int_0^r e^{i\left(t\left(\frac{1+i}{\sqrt{2}}\right)\right)^2} \left(\frac{1+i}{\sqrt{2}}\right) dt \\
&= \int_{\gamma_r} e^{iz^2} dz + \int_0^r e^{ix^2} dx - \int_0^r e^{-t^2} \left(\frac{1+i}{\sqrt{2}}\right) dt \\
&= \int_{\gamma_r} e^{iz^2} dz + \int_0^r e^{ix^2} dx - \frac{\sqrt{\pi}}{2} \left(\frac{1+i}{\sqrt{2}}\right) + e(r)
\end{aligned}
$$

Where $\lim_{r\to 0} e(r) = 0$. Now examine the first integral. Explain the following steps and why this integral converges to 0 as $r \to \infty$.

$$
\left| \int_{\gamma_r} e^{iz^2} dz \right| = \left| \int_0^{\frac{\pi}{4}} e^{i\left(re^{it}\right)^2} rie^{it} dt \right| \leq r \int_0^{\frac{\pi}{4}} e^{-r^2 \sin 2t} dt = \frac{r}{2} \int_0^1 \frac{e^{-r^2 u}}{\sqrt{1-u^2}} du
$$

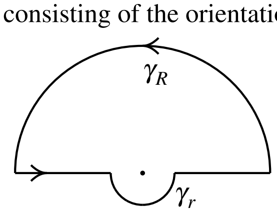
$$
= \frac{r}{2} \int_0^{r^{-(3/2)}} \frac{1}{\sqrt{1-u^2}} du + \frac{r}{2} \left(\int_0^1 \frac{1}{\sqrt{1-u^2}} \right) e^{-\left(r^{1/2}\right)}
$$

15. If $\gamma : C \to \mathbb{C}$ is a parametrization of a curve with γ being differentiable, one to one on $(a,b), a < b$ with continuous derivative, the length of C is defined as

$$
\sup \left\{ \sum_P |\gamma(t_i) - \gamma(t_{i-1})|, P \text{ a partition of } [a,b] \right\}
$$

Show this is independent of equivalent smooth parametrization and that in every case, it equals $\int_a^b |\gamma'(t)| \, dt$, the integral of the absolute value of the derivative.

16. Consider the following contour consisting of the orientation shown by the arrows.

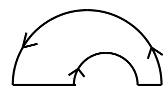


There is a large semicircle on the top of radius R and a small one of radius r. If γ refers to the piecewise smooth, oriented contour consisting of the two straight lines and two semicircles, find, using the method of residues $\int_\gamma \frac{e^{iz}}{z}dz$. The result should depend on r and R. Show the contour integral over γ_R converges to 0 as $R\to\infty$. Then find the limit of the contour integral over γ_r and show it is πi in the limit as $r\to 0$. Then obtain limits as $r\to 0$ and $R\to\infty$ and show that $2i\int_{-\infty}^{\infty}\frac{\sin x}{x}\equiv 2i\lim_{R\to\infty}\int_{-R}^{R}\frac{\sin x}{x}dx=\pi i$. This is another way to get $\int_0^\infty \frac{\sin x}{x}dx=\frac{\pi}{2}$.

17. Find the following improper integral. $\int_{-\infty}^{\infty}\frac{\cos x}{1+x^4}dx$ **Hint:** Use upper semicircle contour and consider instead $\int_{-\infty}^{\infty}\frac{e^{ix}}{1+x^4}dx$. This is because the integral over the semicircle will converge to 0 as $R\to\infty$ if you have e^{iz} but this won't happen if you use $\cos z$ because $\cos z$ will be unbounded. Just write down and check and you will see why this happens. Thus you should use $\frac{e^{iz}}{1+z^4}$ and take real part. I think the standard calculus techniques will not work for this horrible integral.

18. Find $\int_{-\infty}^{\infty}\frac{\cos(x)}{\left(1+x^2\right)^2}dx$. **Hint:** Do the same as above replacing $\cos x$ with e^{ix}.

19. Consider the following contour.



The small semicircle has radius r and is centered at $(1,0)$. The large semicircle has radius R and is centered at $(0,0)$. Use the method of residues to compute

$$\lim_{r\to 0}\left(\lim_{R\to\infty}\int_r^R \frac{x}{1-x^3}dx+\int_{-R}^{r}\frac{x}{1-x^3}dx\right)$$

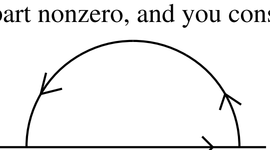
This is called the Cauchy principal value for $\int_{-\infty}^{\infty}\frac{x}{1-x^3}dx$. The integral makes no sense in terms of a real honest integral. The function has a pole on the x axis. However, you can define such a Cauchy principal value. Rather than belabor this issue, I will illustrate with this example. These principal value integrals occur because of cancelation. They depend on a particular way of taking a limit. They are not as mathematically respectable but are certainly important. They are in that general area of finding something by taking a certain kind of symmetric limit.

20. Find $\int_0^{2\pi}\frac{\cos(\theta)}{1+\sin^2(\theta)}d\theta$.

21. Find $\int_0^{2\pi}\frac{d\theta}{2-\sin\theta}$.

22. Find $\int_{-\pi/2}^{\pi/2}\frac{d\theta}{2-\sin\theta}$.

23. Suppose you have a function $f(z)$ which is the quotient of two polynomials in which the degree of the top is two less than the degree of the bottom and all zeros of the denominator have imaginary part nonzero, and you consider the contour.



Then define $\int_{\gamma_R} f(z)e^{isz}dz$ in which s is real and positive. Explain why the integral makes sense and why the part of it on the semicircle converges to 0 as $R\to\infty$. Use this to find $\int_{-\infty}^{\infty}\frac{e^{isx}}{k^2+x^2}dx,\ k>0$.

24. Show using methods from real analysis that for $b \geq 0$, $\int_0^\infty e^{-x^2} \cos(2bx)\, dx = \frac{\sqrt{\pi}}{2} e^{-b^2}$
Hint: Let $F(b) \equiv \int_0^\infty e^{-x^2} \cos(2bx)\, dx - \frac{\sqrt{\pi}}{2} e^{-b^2}$. Show $F(0) = 0$. Recall that

$$\int_0^\infty e^{-x^2} dx = \frac{1}{2}\sqrt{\pi}.$$

Explain using a modification of Problem 48 on Page 226 why

$$\begin{aligned} F'(b) &= \int_0^\infty -2xe^{-x^2} \sin(2bx)\, dx + 2b\frac{\sqrt{\pi}}{2} e^{-b^2} \\ F'(b) &= 2b\left(\int_0^\infty e^{-x^2} \cos(2bx)\, dx + \frac{\sqrt{\pi}}{2} e^{-b^2}\right) \end{aligned}$$

$$= 2b\left(F(b) + \frac{\sqrt{\pi}}{2} e^{-b^2} + \frac{\sqrt{\pi}}{2} e^{-b^2}\right) = 2bF(b) + \sqrt{\pi} 2be^{-b^2}$$

Now use the integrating factor method for solving linear differential equations from beginning differential equations to solve the ordinary differential equation. If you have not seen this method, it is just this. To solve $y' + f(x)y = g(x)$, multiply both sides by $e^{F(x)}$ where $F'(x) = f(x)$. This reduces the left side to $\frac{d}{dx}\left(e^{F(x)}y\right)$. Thus $\frac{d}{db}\left(e^{-b^2}F(b)\right) = \sqrt{\pi}2be^{-2b^2}$ Then do $\int db$ to both sides and use that $F(0) = \frac{1}{2}\sqrt{\pi}$.

25. You can do the same problem as above using contour integration. For $b > 0$, use the contour which goes from $-a$ to a to $a + ib$ to $-a + ib$ to $-a$. Then let $a \to \infty$ and show that the integral of e^{-z^2} over the vertical parts of this contour converge to 0.
Hint: For $z = x + ib$, $e^{-z^2} = e^{-\left(x^2 - b^2 + 2ixb\right)} = e^{b^2} e^{-x^2} \left(\cos(2xb) + i\sin(2xb)\right)$.

26. Consider the circle of radius 1 oriented counter clockwise. $\int_\gamma z^{-6} \cos(z)\, dz =$?

27. The circle of radius 1 is oriented counter clockwise. Evaluate $\int_\gamma z^{-7} \cos(z)\, dz$.

28. Find $\int_0^\infty \frac{2+x^2}{1+x^4} dx$.

29. Find $\int_0^\infty \frac{x^{1/3}}{1+x^2} dx$.

30. Suppose f is analytic on an open set U and $\alpha \in U$. Define

$$g(z) \equiv \begin{cases} \frac{f(z)-f(\alpha)}{z-\alpha} \text{ if } z \neq \alpha \\ f'(\alpha) \text{ if } z = \alpha \end{cases}$$

show that g is analytic on U.

31. Let γ^* be a C^1 oriented closed curve and let $\alpha \notin \gamma^*$ Then

$$n(\gamma, \alpha) \equiv \frac{1}{2\pi i} \int_{\gamma^*} \frac{1}{z - \alpha} dz = m \in \mathbb{N}$$

This is called the winding number. **Hint:** Now let $g(t) \equiv \int_{a_1}^t \frac{\gamma'(s)}{\gamma(s)-\alpha} dt$. Show that $\frac{d}{dt}\left(e^{-g(t)}(\gamma(t) - \alpha)\right) = 0$. Explain why this requires $e^{-g(t)}(\gamma(t) - \alpha)$ is a constant. Since γ parametrizes a closed curve, argue that $g(b_k) = 2m\pi i$ for an integer m.

32. Use Problem 30 to state a theorem whose conclusion is that

$$f(\alpha) n(\gamma^*, \alpha) = \frac{1}{2\pi i} \int_{\gamma^*} \frac{f(z)}{z-\alpha} dz.$$

Here γ^* is a closed curve. This is a case of a general Cauchy integral formula for cycles.

33. Using the open mapping theorem, show that if U is an open connected set and f is analytic and $|f|$ has a maximum on $\overline{U}$, then this maximum occurs on the boundary of U. This is called the maximum modulus theorem.

34. Use the counting zeros theorem, get an estimate for a ball centered at 0 which will contain all zeros of $p(z) = z^7 + 5z^5 - 3z^2 + z + 5$. **Hint:** You might compare with $q(z) = z^7$. You know all of its zeros. For $|z| = r$ large enough, consider $\lambda z^7 + (1-\lambda) p(z)$ and $\lambda \in [0,1]$.

35. Use the counting zeros theorem to give another proof of the fundamental theorem of algebra. This one is even easier than the earlier one based on Liouville's theorem, but it uses the harder theorem about counting zeros.

36. As in the Liouville proof of the fundamental theorem of algebra, if $p(z)$ is a non-constant polynomial with no zero, then $\lim_{|z|\to\infty} (1/|p(z)|) = 0$. Then $1/|p(z)|$ has a maximum on $\mathbb{C}$, say it has such a maximum at z_0. Now exploit the maximum modulus theorem of Problem 33 on balls containing z_0 at the center to obtain a contradiction. Provide details.

37. Rouche's theorem considers the case where there are no poles or zeros on $C(z_0, r)$, the counter clockwise oriented circle bounding $B(z_0, r)$ and finitely many zeros and poles in $B(z_0, r)$. Such a function is called meromorphic. Rouche's theorem says that in the situation just described and f such a meromorphic function $\frac{1}{2\pi i} \int_{C(z_0,r)} \frac{f'(z)}{f(z)} dz = N - P$ where N is the number of zeros counted according to multiplicity and P the number of poles, also counted according to something which will be apparent when this is proved. Use the proof given above for counting zeros to verify this more general theorem. It just involves finding the residues of f'/f at the zeros and poles.

38. If U is an open connected subset of $\mathbb{C}$ and $f : U \to \mathbb{R}$ is analytic, what can you say about f? **Hint:** You might consider the open mapping theorem.

39. Using the fundamental theorem of algebra and the partial fractions theorem of Problem 19 on Page 40, show that the result of Theorem 11.3.2 is always obtained for rational functions. Give an example where Theorem 11.3.2 is better.

40. Show that if U is an open connected set in $\mathbb{C}$ and $f : U \to \mathbb{C}$ is one to one and analytic, then $f^{-1} : f(U) \to U$ is also analytic. Really? What about $f(z) = z^3$? **Hint:** Do something like in Theorem 7.10.1. Here it may be easier because by the open mapping theorem, you know $f(U)$ is open so there are no endpoints to worry about.

Chapter 12

Series and Transforms

12.1 Fourier Series

A Fourier series is an expression of the form $\sum_{k=-\infty}^{\infty} c_k e^{ikx}$ where this infinite sum is understood 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 Napoleon'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.[1] From the very beginning, the fundamental question has been related to the nature of convergence of these series. Dirichlet was the first to prove significant theorems on this in 1829, but questions lingered till the mid 1960's when a problem involving convergence of Fourier series was solved for the first time and the solution of this problem was a big surprise.[2] This chapter is on the classical theory of convergence of Fourier series studied by Dirichlet, Riemann, and Fejer.

If you can approximate a function f with an expression of the form $\sum_{k=-\infty}^{\infty} c_k e^{ikx}$ then the function must have the property $f(x+2\pi) = f(x)$ because this is true of every term in the above series. More generally, here is a definition.

Definition 12.1.1 *A function f defined on $\mathbb{R}$ is a periodic function of period T if $f(x+T) = f(x)$ for all x.*

As just explained, Fourier series are useful for representing periodic functions and no other kind of function.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, you can study this function in terms of the function $g(x) \equiv f\left(\frac{Tx}{2\pi}\right)$ where g is periodic of period 2π.

Definition 12.1.2 *For $f \in R([-\pi,\pi])$ and f periodic on $\mathbb{R}$, define the Fourier series of f as*

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{12.1}$$

where

$$c_k \equiv \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y) e^{-iky} dy. \tag{12.2}$$

Also define the n^{th} partial sum of the Fourier series of f by

$$S_n(f)(x) \equiv \sum_{k=-n}^{n} c_k e^{ikx}. \tag{12.3}$$

[1] Fourier was with Napoleon in Egypt when the Rosetta Stone was discovered and wrote about Egypt in Description de l'Égypte. He was a teacher of Champollion who eventually made it possible to read Egyptian by using the Rosetta Stone, discovered at this time. This expedition of Napoleon caused great interest in all things Egyptian in the first part of the nineteenth century.

[2] The question was whether the Fourier series of a function in L^2 converged almost everywhere to the function where the term "almost everywhere" has a precise meaning. (L^2 means the square of the function is integrable.) It turned out that it did, to the surprise of many because it was known that the Fourier series of a function in L^1 does not necessarily converge to the function almost everywhere, this from an example given by Kolmogorov in the 1920's. The problem was solved by Carleson in 1965.

It may be interesting to see where this formula came from. Suppose then that $f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$, 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 switch the sum and the integral on the right side even though there is absolutely no reason to believe this makes any sense. Then $\int_{-\pi}^{\pi} f(x) e^{-imx} dx = \sum_{k=-\infty}^{\infty} c_k \int_{-\pi}^{\pi} e^{ikx} e^{-imx} dx = c_m \int_{-\pi}^{\pi} 1 dx = 2\pi c_m$ 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 12.1.2 might be interesting.

In case f is real valued, $\overline{c_k} = c_{-k}$ 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 $c_k \equiv \alpha_k + i\beta_k$

$$S_n f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) dy + \sum_{k=1}^{n} 2 [\alpha_k \cos kx - \beta_k \sin kx]$$

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 $b_k = -2\beta_k$,

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ky) dy, \ b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ky) dy$$

and

$$S_n f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx \tag{12.4}$$

This is often the way Fourier series are presented in elementary courses where it is only real functions which are to be approximated. However it is easier to stick with Definition 12.1.2.

The partial sums of a Fourier series can be written in a particularly simple form which is presented next.

$$\begin{aligned} S_n f(x) &= \sum_{k=-n}^{n} c_k e^{ikx} = \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iky} dy\right) e^{ikx} \\ &= \int_{-\pi}^{\pi} \overbrace{\frac{1}{2\pi} \sum_{k=-n}^{n} \left(e^{ik(x-y)}\right)}^{D_n(x-y)} f(y) dy \equiv \int_{-\pi}^{\pi} D_n(x-y) f(y) dy. \end{aligned}$$

The function $D_n(t) \equiv \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikt}$ is called the Dirichlet Kernel.

Theorem 12.1.3 *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 since

$$e^{ik(t+2\pi)} = \cos(kt+2\pi k) + i\sin(kt+2\pi k) = \cos(kt) + i\sin(kt) = e^{ikt}$$

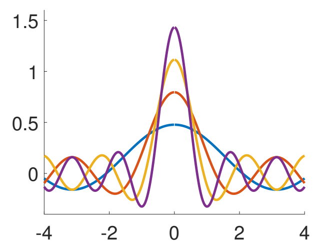
It remains to verify Part 3. Note $2\pi D_n(t) = \sum_{k=-n}^{n} e^{ikt} = 1 + 2\sum_{k=1}^{n} \cos(kt)$. Therefore,

$$2\pi D_n(t)\sin\left(\frac{t}{2}\right) = \sin\left(\frac{t}{2}\right) + 2\sum_{k=1}^{n} \sin\left(\frac{t}{2}\right)\cos(kt)$$

$$= \sin\left(\frac{t}{2}\right) + \sum_{k=1}^{n} \sin\left(\left(k+\frac{1}{2}\right)t\right) - \sin\left(\left(k-\frac{1}{2}\right)t\right) = \sin\left(\left(n+\frac{1}{2}\right)t\right)$$

where the easily verified trig. identity $\cos(a)\sin(b) = \frac{1}{2}(\sin(a+b) - \sin(a-b))$ is used to get to the second line. This proves 3 and proves the theorem. ■

Here is a picture of the Dirichlet kernels for $n = 1,2,3$ and 4



Note they are not nonnegative but there is a large central positive bump which gets larger as n gets larger.

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.

12.2 Criteria for Convergence

Fourier series like to converge to the midpoint of the jump of a function under suitable conditions. This was first shown by Dirichlet in 1829 after others had tried unsuccessfully to prove such a result. The condition given for convergence in the following theorem is due to Dini. [3] It is a generalization of the usual theorem presented in elementary books on Fourier series methods. Fourier did not appreciate the difficulty of this question and was happy to believe that the series did converge to the function in some useable sense despite the doubts of people like Lagrange and Laplace. For over 150 years they studied this question and major results appeared as recently as the mid 1960's.

It may be that the study of Fourier series and their convergence drove the development of real analysis during the nineteenth century as much as any other topic. Dirichlet did his work before Riemann gave the best description of the integral and it was Riemann who gave a general theory of integration including piecewise continuous functions. Cauchy's integral for continuous functions was the current state of the art at the time of Dirichlet. Thus the theorem about to be presented uses a more sophisticated theory of integration than that which was available then.

Recall $\lim_{t\to x+} f(t) \equiv f(x+)$, and $\lim_{t\to x-} f(t) \equiv f(x-)$.

Theorem 12.2.1 *Let f be periodic of period 2π which is in $R([-\pi,\pi])$. Suppose at some x, $f(x+)$ and $f(x-)$ both exist and that the Dini conditions hold which are that for small positive y,*

$$|f(x-y)-f(x-)| \leq Ky^{\theta}, 0<\theta\leq 1,\ |f(x+y)-f(x+)| \leq Ky^{\theta}, 0<\theta\leq 1$$

for $0<y\leq\delta,\ \delta>0$. Then

$$\lim_{n\to\infty} S_n f(x) = \frac{f(x+)+f(x-)}{2}. \tag{12.5}$$

Proof: $S_n f(x) = \int_{-\pi}^{\pi} D_n(x-y) f(y)\, dy$. Change variables $x-y\to y$ and use the periodicity of f and D_n along with the formula for $D_n(y)$ to write this as

$$S_n f(x) = \int_{-\pi}^{\pi} D_n(y) f(x-y)\, dy = \int_0^{\pi} D_n(y) f(x-y)\, dy + \int_{-\pi}^{0} D_n(y) f(x-y)\, dy$$

$$= \int_0^{\pi} D_n(y)\left[f(x-y)+f(x+y)\right] dy$$

$$= \int_0^{\pi} \frac{1}{\pi} \frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{\sin\left(\frac{y}{2}\right)} \left[\frac{f(x-y)+f(x+y)}{2}\right] dy. \tag{12.6}$$

since $\int_{-\pi}^{\pi} D_n(y)\, dy = 1, S_n f(x) - \frac{f(x+)+f(x-)}{2} =$

$$\int_0^{\pi} \frac{1}{\pi} \frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{\sin\left(\frac{y}{2}\right)} \left[\frac{f(x-y)+f(x+y)}{2} - \frac{f(x+)+f(x-)}{2}\right] dy$$

$$= \int_{\delta}^{\pi} \frac{1}{\pi} \sin\left(\left(n+\frac{1}{2}\right)y\right) \left[\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{2\sin\left(\frac{y}{2}\right)}\right] dy$$

$$+ \int_0^{\delta} \frac{2}{\pi} \sin\left(\left(n+\frac{1}{2}\right)y\right) \frac{y/2}{\sin\left(\frac{y}{2}\right)} \left[\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{2y}\right] dy \tag{12.7}$$

In the first integral of 12.7, the function in $[\ \ ]$ is in $L^1([\delta,\pi])$ because $\sin(y/2)$ is bounded away from 0. Therefore, the first of these integrals converges to 0 by the Riemann Lebesgue theorem. In the second integral, $|[\ \ ]| \leq Ky^{1-\theta}$ and $\frac{y/2}{\sin\left(\frac{y}{2}\right)}$ is bounded on $[0,\delta]$ so this function multiplying $\sin\left(\left(n+\frac{1}{2}\right)y\right)$ is in $L^1(0,\delta)$. Therefore, the second integral also converges to 0. ■

The following corollary is obtained immediately from the above proof with minor modifications.

Corollary 12.2.2 *Let f be a periodic function of period 2π which is an element of $R([-\pi,\pi])$. Suppose at some x, the function*

$$y \to \left|\frac{f(x-y)+f(x+y)-2s}{y}\right| \tag{12.8}$$

is in $R((0,\pi])$. Then $\lim_{n\to\infty} S_n f(x) = s$.

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. There are whole books based on Fourier series such as Trigonometric Series by Zygmund [27] which appeared first in 1935 and there is a lot more in this book than the short introduction to the topic presented here.

There is another easy to check condition which implies convergence to the midpoint of the jump. It was shown above that

$$\left|S_n f(x) - \frac{f(x+)+f(x-)}{2}\right| =$$

$$\left|\int_0^\pi \frac{1}{\pi}\frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{\sin\left(\frac{y}{2}\right)}\left[\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{2}\right]dy\right|$$

$$= \int_0^\delta \frac{2}{\pi}\frac{\sin\left(\left(n+\frac{1}{2}\right)y\right)}{y}\frac{y/2}{\sin(y/2)}\left[\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{2}\right]dy +$$

$$\int_\delta^\pi \frac{1}{\pi}\sin\left(\left(n+\frac{1}{2}\right)y\right)\left[\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{2\sin\left(\frac{y}{2}\right)}\right]dy$$

In the second integral, the expression in [ ] is in $L^1([\delta,\pi])$ and so, by the Riemann Lebesgue lemma, this integral converges to 0.

If you know that f has finite total variation in $[x-\delta, x+\delta]$, then you could use Lemma 10.2.7 to conclude that the first integral converges, as $n\to\infty$, to $g(0+)$ where $g(y) = \frac{y/2}{\sin(y/2)}\frac{f(x-y)-f(x-)+f(x+y)-f(x+)}{2}$ so that $g(0+) = 0$. Thus, there is another corollary.

Corollary 12.2.3 *Let f be a periodic function of period 2π which is an element of $R([-\pi,\pi])$. Suppose at some x, $f(x+)$ and $f(x-)$ both exist and f is of bounded variation on $[x-\delta, x+\delta]$ for some $\delta > 0$. Then*

$$\lim_{n\to\infty} S_n f(x) = \frac{f(x+)+f(x-)}{2}. \tag{12.9}$$

There are essentially two conditions which yield convergence to the mid point of the jump, the Dini conditions and the Jordan condition which is on finite total variation. You have to have something of this sort. If you only know you have continuity from the right and the left, you don't necessarily get convergence to the midpoint of the jump. At least this is so for the Fourier series. If you use the Ceasaro means, this kind of convergence takes place without either of these two technical conditions. The study of Ceasaro means is associated with Fejer whose work came much later in early 1900's. This is considered later.

It might be of interest to note that in the argument for convergence given earlier in Lemma 10.2.7, h was determined by $|g(t) - g(0+)|$ small enough. If you have the case of a continuous function defined on a closed and bounded interval, $|g(t) - g(s)|$ would be small enough whenever $|t - s|$ is suitably small independent of the choice of s thanks to uniform continuity. When the above corollary is applied to convergence of Fourier series, one massages things as above to reduce to the kind of thing given in Lemma 10.2.7 as shown above and a single h will then suffice for all the points at once. Once h has been determined, the convergence of the other terms in Lemma 10.2.7 for such a continuous periodic function of bounded variation will not depend on the point and so this argument ends up showing that one has uniform convergence of the Fourier series to the function if the periodic function is of bounded variation and continuous on every interval.

12.3 Integrating and Differentiating Fourier Series

First here is a review of what it means for a function to be piecewise continuous.

Definition 12.3.1 *Let f be a function defined on $[a,b]$. It is called piecewise continuous if there is a partition of $[a,b]$, $\{x_0, \cdots, x_n\}$ such that on $[x_{k-1}, x_k]$ there is a continuous function g_k such that $f(x) = g_k(x)$ for all $x \in (x_{k-1}, x_k)$.*

You can typically integrate Fourier series term by term and things will work out according to your expectations. More precisely, if the Fourier series of f is $\sum_{k=-\infty}^{\infty} a_k e^{ikx}$ then it will be true that

$$F(x) \equiv \int_{-\pi}^{x} f(t)\,dt = \lim_{n\to\infty} \sum_{k=-n}^{n} a_k \int_{-\pi}^{x} e^{ikt}\,dt$$

$$= a_0(x+\pi) + \lim_{n\to\infty} \sum_{k=-n,k\neq 0}^{n} a_k \left(\frac{e^{ikx}}{ik} - \frac{(-1)^k}{ik} \right).$$

I shall show this is true for the case where f is an arbitrary 2π periodic function which is piecewise continuous according to the above definition. However, with a better theory of integration, it all works for much more general functions than these. It is limited here to a simpler case because we don't have a very sophisticated theory of integration. With the Lebesgue theory of integration, all restrictions vanish. It suffices to consider very general functions with no assumptions of continuity. This is still a remarkable result however, even with the restriction to piecewise continuous functions.

Note that it is not necessary to assume anything about the function f being the limit of its Fourier series. Let

$$G(x) \equiv F(x) - a_0(x+\pi) = \int_{-\pi}^{x} (f(t) - a_0)\,dt$$

Then G equals 0 at $-\pi$ and π because $2\pi a_0 = \int_{-\pi}^{\pi} f(t)\,dt$. Therefore, the periodic extension of G, still denoted as G, is continuous. Also

$$|G(x) - G(x_1)| \leq \left| \int_{x_1}^{x} M\,dt \right| \leq M|x - x_1|$$

where M is an upper bound for $|f(t) - a_0|$. Thus the Dini condition of Corollary 12.2.1 holds. Therefore for all $x \in \mathbb{R}$,

$$G(x) = \sum_{k=-\infty}^{\infty} A_k e^{ikx} \tag{12.10}$$

where $A_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} G(y) e^{iky} dy$. Now from 12.10 and the definition of the Fourier coefficients for f,

$$G(\pi) = F(\pi) - a_0 2\pi = 0 = A_0 + \lim_{n\to\infty} \sum_{k=-n,k\neq 0}^{n} A_k (-1)^k \tag{12.11}$$

and so

$$A_0 = -\lim_{n\to\infty} \sum_{k=-n,k\neq 0}^{n} A_k (-1)^k \equiv -\sum_{k=-\infty,k\neq 0}^{\infty} A_k (-1)^k \tag{12.12}$$

Next consider A_k for $k \neq 0$.

$$\begin{aligned} A_k &\equiv \frac{1}{2\pi}\int_{-\pi}^{\pi} G(x) e^{-ikx} dx \equiv \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{x} (f(t) - a_0) dt e^{-ikx} dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx} \int_{-\pi}^{x} (f(t) - a_0) dt dx \end{aligned}$$

Now let $\psi(x) \equiv \int_{-\pi}^{x} (f(t) - a_0) dt$ and $\phi_k(x) = \frac{e^{-ikx}}{-ik}$. Then the above integral is of the form $\frac{1}{2\pi}\int_{-\pi}^{\pi} \psi(x) d\phi_k(x)$. Since $\psi(\pi)\phi_k(\pi) = 0 = \psi(-\pi)\phi_k(-\pi)$, integration by parts, Theorem 9.4.1, says that the above equals

$$-\frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_k(x) d\psi(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{-ikx}}{ik} d\psi(x)$$

Use Corollary 9.3.18 to write as a sum of finitely many integrals

$$\sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \frac{e^{-ikx}}{ik} d\psi(x).$$

It follows from Proposition 9.4.3 that this is

$$A_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{-ikx}}{ik} (f(x) - a_0) dx = \frac{1}{ik}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-ikx} dx \equiv \frac{a_k}{ik}$$

Thus this shows from 12.12, that for all x,

$$G(x) = \sum_{k=-\infty,k\neq 0}^{\infty} \frac{a_k}{ik} e^{ikx} + A_0,\ A_0 = -\sum_{k=-\infty,k\neq 0}^{\infty} A_k (-1)^k$$

and so,

$$\int_{-\pi}^{x} f(t) dt - \int_{-\pi}^{x} a_0 = \sum_{k=-\infty,k\neq 0}^{\infty} \frac{a_k}{ik} e^{ikx} - \frac{a_k}{ik} e^{ik(-\pi)}$$

which shows that

$$\int_{-\pi}^{x} f(t) dt = \int_{-\pi}^{x} a_0 + \sum_{k=-\infty,k\neq 0}^{\infty} a_k \int_{-\pi}^{x} e^{ikt} dt$$

This proves the following theorem.

Theorem 12.3.2 *Let f be 2π periodic and piecewise continuous. Then*

$$\int_{-\pi}^{x} f(t) dt = \int_{-\pi}^{x} a_0 dt + \lim_{n\to\infty} \sum_{k=-n,k\neq 0}^{n} a_k \int_{-\pi}^{x} e^{ikt} dt$$

where a_k are the Fourier coefficients of f. This holds for all $x \in \mathbb{R}$.

Example 12.3.3 *Let $f(x) = x$ for $x \in [-\pi, \pi)$ and extend f to make it 2π periodic. Then the Fourier coefficients of f are $a_0 = 0$, $a_k = \frac{(-1)^k i}{k}$. Therefore,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} te^{-ikt} = \frac{i}{k}\cos \pi k, \int_{-\pi}^{x} tdt = \frac{1}{2}x^2 - \frac{1}{2}\pi^2 = \lim_{n\to\infty}\sum_{k=-n,k\neq 0}^{n}\frac{(-1)^k i}{k}\int_{-\pi}^{x} e^{ikt}dt$$

$$= \lim_{n\to\infty}\sum_{k=-n,k\neq 0}^{n}\frac{(-1)^k i}{k}\left(\frac{\sin xk}{k} + i\frac{-\cos xk + (-1)^k}{k}\right)$$

For fun, let $x = 0$ and conclude $-\frac{1}{2}\pi^2 =$

$$\lim_{n\to\infty}\sum_{k=-n,k\neq 0}^{n}\frac{(-1)^k i}{k}\left(i\frac{-1+(-1)^k}{k}\right) = \lim_{n\to\infty}\sum_{k=-n,k\neq 0}^{n}\frac{(-1)^{k+1}}{k}\left(\frac{-1+(-1)^k}{k}\right)$$

$$= \lim_{n\to\infty} 2\sum_{k=1}^{n}\frac{(-1)^k + (-1)}{k^2} = \sum_{k=1}^{\infty}\frac{-4}{(2k-1)^2}$$

and so $\frac{\pi^2}{8} = \sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}$

Of course it is not reasonable to suppose that you can differentiate a Fourier series term by term and get good results.

Consider the series for $f(x) = 1$ if $x \in (0, \pi]$ and $f(x) = -1$ on $(-\pi, 0)$ with $f(0) = 0$. In this case $a_0 = 0$. $a_k = \frac{1}{2\pi}\left(\int_0^{\pi} e^{-ikt}dt - \int_{-\pi}^{0} e^{-ikt}dt\right) = \frac{i}{\pi}\frac{\cos \pi k - 1}{k}$ so the Fourier series is $\sum_{k\neq 0}\left(\frac{(-1)^k - 1}{\pi k}\right) ie^{ikx}$.What happens if you differentiate it term by term? It gives $\sum_{k\neq 0} -\frac{(-1)^k - 1}{\pi} e^{ikx}$ which fails to converge anywhere because the k^{th} term fails to converge to 0. This is in spite of the fact that f has a derivative away from 0.

However, it is possible to prove some theorems which let you differentiate a Fourier series term by term. Here is one such theorem.

Theorem 12.3.4 *Suppose for $x \in [-\pi, \pi]$ $f(x) = \int_{-\pi}^{x} f'(t)\,dt + f(-\pi)$ and $f'(t)$ is piecewise continuous. Denoting by f the periodic extension of the above, then if $f(x) = \sum_{k=-\infty}^{\infty} a_k e^{ikx}$ it follows the Fourier series of f' is $\sum_{k=-\infty}^{\infty} a_k ike^{ikx}$.*

Proof: Since f' is piecewise continuous, 2π periodic it follows from Theorem 12.3.2

$$f(x) - f(-\pi) = \sum_{k=-\infty}^{\infty} b_k\left(\int_{-\pi}^{x} e^{ikt}dt\right)$$

where b_k is the k^{th} Fourier coefficient of f'. Thus $b_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f'(t)\,e^{-ikt}dt$. Breaking the integral into pieces if necessary, and integrating these by parts yields finally

$$b_k = \frac{1}{2\pi}\left[f(t)\,e^{-ikt}|_{-\pi}^{\pi} + ik\int_{-\pi}^{\pi} f(t)\,e^{-ikt}dt\right] = ik\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,e^{-ikt}dt = ika_k$$

where a_k is the Fourier coefficient of f. Since f is periodic of period 2π, the boundary term vanishes. It follows the Fourier series for f' is $\sum_{k=-\infty}^{\infty} ika_k e^{ikx}$ as claimed. ■

Note the conclusion of this theorem is only about the Fourier series of f'. It does not say the Fourier series of f' converges pointwise to f'. However, if f' satisfies a Dini condition, then this will also occur. For example, if f' has a bounded derivative at every point, then by the mean value theorem $|f'(x) - f'(y)| \leq K|x-y|$ and this is enough to show the Fourier series converges to $f'(x)$ thanks to Corollary 12.2.1.

12.4 Ways of Approximating Functions

Given above is a theorem about Fourier series converging pointwise to a periodic function or more generally to the mid point of the jump of the function. Notice that some sort of smoothness of the function approximated was required, the Dini condition. It can be shown that if this sort of thing is not present, the Fourier series of a continuous periodic function may fail to converge to it in a very spectacular manner. In fact, Fourier series don't do very well at converging pointwise. However, there is another way of converging at which Fourier series cannot be beat. It is mean square convergence.

Definition 12.4.1 *Let f be a function defined on an interval, $[a,b]$. Then a sequence, $\{g_n\}$ of functions is said to converge uniformly to f on $[a,b]$ if*

$$\lim_{n\to\infty} \sup \{|f(x) - g_n(x)| : x \in [a,b]\} = 0.$$

The expression $\sup \{|f(x) - g_n(x)| : x \in [a,b]\}$ is sometimes written[3] as $||f - g_n||_0$. More generally, if f is a function,

$$\|f\|_0 \equiv \sup \{|f(x)| : x \in [a,b]\}$$

The sequence is said to converge mean square to f if

$$\lim_{n\to\infty} \|f - g_n\|_2 \equiv \lim_{n\to\infty} \left(\int_a^b |f - g_n|^2 \, dx \right)^{1/2} = 0$$

12.5 Uniform Approximation with Trig. Polynomials

It turns out that if you don't insist the a_k be the Fourier coefficients, then every continuous 2π periodic function $\theta \to f(\theta)$ can be approximated uniformly with a Trig. polynomial of the form $p_n(\theta) \equiv \sum_{k=-n}^{n} a_k e^{ik\theta}$. This means that for all $\varepsilon > 0$ there exists a $p_n(\theta)$ such that $\|f - p_n\|_0 < \varepsilon$. Here $\|f\|_0 \equiv \max \{|f(x)| : x \in \mathbb{R}\}$.

Definition 12.5.1 *Recall the n^{th} partial sum of the Fourier series $S_n f(x)$ is given by*

$$S_n f(x) = \int_{-\pi}^{\pi} D_n(x-y) f(y) \, dy = \int_{-\pi}^{\pi} D_n(t) f(x-t) \, dt$$

where $D_n(t)$ is the Dirichlet kernel, $D_n(t) = (2\pi)^{-1} \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\left(\frac{t}{2}\right)}$ The n^{th} Fejer mean, $\sigma_n f(x)$ is the average of the first n of the $S_n f(x)$. Thus

$$\sigma_{n+1} f(x) \equiv \frac{1}{n+1} \sum_{k=0}^{n} S_k f(x) = \int_{-\pi}^{\pi} \left(\frac{1}{n+1} \sum_{k=0}^{n} D_k(t) \right) f(x-t) \, dt$$

The Fejer kernel is $F_{n+1}(t) \equiv \frac{1}{n+1} \sum_{k=0}^{n} D_k(t)$.

As was the case with the Dirichlet kernel, the Fejer kernel has some properties.

[3] There is absolutely no consistency in this notation. It is often the case that $||\cdot||_0$ is what is referred to in this definition as $||\cdot||_2$. Also $||\cdot||_0$ here is sometimes referred to as $||\cdot||_\infty$. Sometimes $||\cdot||_2$ referrs to a norm which involves derivatives of the function.

Lemma 12.5.2 *The Fejer kernel has the following properties.*

1. $F_{n+1}(t) = F_{n+1}(t+2\pi)$
2. $\int_{-\pi}^{\pi} F_{n+1}(t)\,dt = 1$
3. $\int_{-\pi}^{\pi} F_{n+1}(t) f(x-t)\,dt = \sum_{k=-n}^{n} b_k e^{ik\theta}$ *for a suitable choice of* b_k.
4. $F_{n+1}(t) = \frac{1-\cos((n+1)t)}{4\pi(n+1)\sin^2\left(\frac{t}{2}\right)}$, $F_{n+1}(t) \geq 0, F_n(t) = F_n(-t)$.
5. *For every* $\delta > 0, \lim_{n\to\infty} \sup\{F_{n+1}(t) : \pi \geq |t| \geq \delta\} = 0$. *In fact, for*
$$|t| \geq \delta, F_{n+1}(t) \leq \frac{2}{(n+1)\sin^2\left(\frac{\delta}{2}\right)4\pi}.$$

Proof: Part 1.) is obvious because F_{n+1} is the average of functions for which this is true.

Part 2.) is also obvious for the same reason as Part 1.). Part 3.) is obvious because it is true for D_n in place of F_{n+1} and then taking the average yields the same sort of sum.

The last statements in 4.) are obvious from the formula which is the only hard part of 4.).

$$F_{n+1}(t) = \frac{1}{(n+1)\sin\left(\frac{t}{2}\right)2\pi} \sum_{k=0}^{n} \sin\left(\left(k+\frac{1}{2}\right)t\right)$$
$$= \frac{1}{(n+1)\sin^2\left(\frac{t}{2}\right)2\pi} \sum_{k=0}^{n} \sin\left(\left(k+\frac{1}{2}\right)t\right)\sin\left(\frac{t}{2}\right)$$

Using the identity $\sin(a)\sin(b) = \cos(a-b) - \cos(a+b)$ with $a = \left(k+\frac{1}{2}\right)t$ and $b = \frac{t}{2}$, it follows

$$F_{n+1}(t) = \frac{1}{(n+1)\sin^2\left(\frac{t}{2}\right)4\pi} \sum_{k=0}^{n} (\cos(kt) - \cos(k+1)t) = \frac{1-\cos((n+1)t)}{(n+1)\sin^2\left(\frac{t}{2}\right)4\pi}$$

which completes the demonstration of 4.).

Next consider 5.). Since F_{n+1} is even it suffices to show

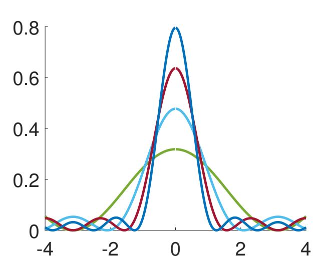
$$\lim_{n\to\infty} \sup\{F_{n+1}(t) : \pi \geq t \geq \delta\} = 0.$$

For the given t,

$$F_{n+1}(t) \leq \frac{1-\cos((n+1)t)}{(n+1)\sin^2\left(\frac{\delta}{2}\right)4\pi} \leq \frac{2}{(n+1)\sin^2\left(\frac{\delta}{2}\right)4\pi}$$

which shows 5.). This proves the lemma.

Here is a picture of the Fejer kernels $F_{n+1}(t)$ for $n = 1,2,3,4$.



Note how these kernels are nonnegative, unlike the Dirichlet kernels. Also there is a large bump in the center which gets increasingly large as n gets larger. The fact these kernels are nonnegative is what is responsible for the superior ability of the Fejer means to approximate a continuous function.

Theorem 12.5.3 *Let f be a continuous and 2π periodic function. Then*

$$\lim_{n\to\infty} ||f-\sigma_{n+1}f||_0 = 0.$$

Proof: Let $\varepsilon > 0$ be given. Then by part 2. of Lemma 12.5.2,

$$|f(x)-\sigma_{n+1}f(x)| = \left|\int_{-\pi}^{\pi} f(x) F_{n+1}(y)\,dy - \int_{-\pi}^{\pi} F_{n+1}(y) f(x-y)\,dy\right|$$

$$= \left|\int_{-\pi}^{\pi} (f(x)-f(x-y)) F_{n+1}(y)\,dy\right| \le \int_{-\pi}^{\pi} |f(x)-f(x-y)| F_{n+1}(y)\,dy$$

$$= \int_{-\delta}^{\delta} |f(x)-f(x-y)| F_{n+1}(y)\,dy + \int_{\delta}^{\pi} |f(x)-f(x-y)| F_{n+1}(y)\,dy$$
$$+ \int_{-\pi}^{-\delta} |f(x)-f(x-y)| F_{n+1}(y)\,dy$$

Since F_{n+1} is even and $|f|$ is continuous and periodic, hence bounded by some constant M the above is dominated by

$$\le \int_{-\delta}^{\delta} |f(x)-f(x-y)| F_{n+1}(y)\,dy + 4M\int_{\delta}^{\pi} F_{n+1}(y)\,dy$$

Now choose δ such that for all x, it follows that if $|y| < \delta$ then $|f(x)-f(x-y)| < \varepsilon/2$. This can be done because f is uniformly continuous on $[-\pi,\pi]$ by Theorem 6.7.2 on Page 114. Since it is periodic, it must also be uniformly continuous on $\mathbb{R}$. (why?) Therefore, for this δ, this has shown that for all x, $|f(x)-\sigma_{n+1}f(x)| \le \varepsilon/2 + 4M\int_{\delta}^{\pi} F_{n+1}(y)\,dy$ and now by Lemma 12.5.2 it follows

$$||f-\sigma_{n+1}f||_0 \le \varepsilon/2 + \frac{8M\pi}{(n+1)\sin^2\left(\frac{\delta}{2}\right)4\pi} < \varepsilon$$

provided n is large enough. ■

12.6 Mean Square Approximation

The partial sums of the Fourier series of f do a better job approximating f in the mean square sense than any other linear combination of the functions, $e^{ik\theta}$ for $|k| \le n$. This will be shown next. It is nothing but a simple computation. Recall the Fourier coefficients are

$$a_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-ik\theta}\,d\theta$$

Also recall that

$$\int_{-\pi}^{\pi} e^{ik\theta}e^{-il\theta}\,d\theta = \begin{cases} 2\pi \text{ if } l = k \\ 0 \text{ if } l \ne k \end{cases}$$

Then using this fact as needed, consider the following computation in which I will try to choose b_k to make

$$\int_{-\pi}^{\pi}\left|f(\theta)-\sum_{k=-n}^{n}b_k e^{ik\theta}\right|^2 d\theta \tag{12.13}$$

as small as possible. Remember that $|z|^2 = z\bar{z}$ whenever z is a complex number. Using this and doing routine computations,

$$\begin{aligned}
&\int_{-\pi}^{\pi}\left|f(\theta)-\sum_{k=-n}^{n}b_k e^{ik\theta}\right|^2 d\theta \\
=\ &\int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2\,\mathrm{Re}\int_{-\pi}^{\pi}\sum_{k=-n}^{n}\overline{f(\theta)}b_k e^{i\theta}d\theta + 2\pi\sum_{k=-n}^{n}|b_k|^2 \\
=\ &\int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2(2\pi)\,\mathrm{Re}\sum_{k=-n}^{n}\overline{a_k}b_k + 2\pi\sum_{k=-n}^{n}|b_k|^2
\end{aligned}$$

Note that if $b_k = a_k$, this would equal

$$\int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2(2\pi)\sum_{k=-n}^{n}|a_k|^2 + 2\pi\sum_{k=-n}^{n}|a_k|^2 = \int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2\pi\sum_{k=-n}^{n}|a_k|^2$$

In the general case, it follows from the Cauchy Schwarz inequality,

$$\geq \int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2(2\pi)\left(\sum_{k=-n}^{n}|a_k|^2\right)^{1/2}\left(\sum_{k=-n}^{n}|b_k|^2\right)^{1/2} + 2\pi\sum_{k=-n}^{n}|b_k|^2$$

$$\begin{aligned}
\geq\ &\int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2\pi\left(\sum_{k=-n}^{n}|a_k|^2 + \sum_{k=-n}^{n}|b_k|^2\right) + 2\pi\sum_{k=-n}^{n}|b_k|^2 \\
=\ &\int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2\pi\sum_{k=-n}^{n}|a_k|^2
\end{aligned}$$

Therefore, the expression in 12.13 is minimized when $b_k = a_k$. We also observe the following fundamental inequality. For a_k the Fourier coefficients,

$$\begin{aligned}
\int_{-\pi}^{\pi}\left|f(\theta)-\sum_{k=-n}^{n}a_k e^{ik\theta}\right|^2 d\theta &\equiv \int_{-\pi}^{\pi}|f(\theta)-S_n f(\theta)|^2 d\theta \\
&= \int_{-\pi}^{\pi}|f(\theta)|^2 d\theta - 2\pi\sum_{k=-n}^{n}|a_k|^2 \geq 0
\end{aligned}$$

so this yields Parseval's inequality, an important inequality involving the Fourier coefficients, $\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(\theta)|^2 d\theta \geq \sum_{k=-n}^{n}|a_k|^2$. This has proved most of the following approximation theorem.

Theorem 12.6.1 *Let $\alpha_n f(x)$ denote any linear combination of the functions $\theta \to e^{ik\theta}$ for $-n \leq k \leq n$. Then*

$$\int_{-\pi}^{\pi}|f-\alpha_n f|^2 dx \geq \int_{-\pi}^{\pi}|f-S_n f|^2 dx$$

Also, $\int_{-\pi}^{\pi}|S_n f|^2 dx \leq \int_{-\pi}^{\pi}|f|^2 dx$.

Proof: It only remains to verify the last inequality. However, a short computation shows that $\int_{-\pi}^{\pi} |S_n f|^2 \, dx = 2\pi \sum_{k=-n}^{n} |a_k|^2 \leq \int_{-\pi}^{\pi} |f(\theta)|^2 \, d\theta$. ■

Now it is easy to prove the following fundamental theorem.

Theorem 12.6.2 *Let $f \in R([-\pi,\pi])$ and it is periodic of period 2π. Then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |f - S_n f|^2 \, dx = 0.$$

Proof: First assume f is continuous and 2π periodic. Then by Theorem 12.6.1,

$$\begin{aligned}
\int_{-\pi}^{\pi} |f - S_n f|^2 \, dx &\leq \int_{-\pi}^{\pi} |f - \sigma_{n+1} f|^2 \, dx \\
&\leq \int_{-\pi}^{\pi} \|f - \sigma_{n+1} f\|_0^2 \, dx = 2\pi \, \|f - \sigma_{n+1} f\|_0^2
\end{aligned}$$

and the last expression converges to 0 by Theorem 12.5.3. Here $\sigma_{n+1} f$ is the Ceasaro mean of f.

Next suppose $f \in R([-\pi,\pi])$. By Lemma 10.1.2, there is h with $|h| \leq |f|$, h continuous and h equal to 0 at the $\pm\pi$ and $\int_{-\pi}^{\pi} |f - h|^2 \, dx < \varepsilon$. Since h vanishes at the endpoints, if h is extended off $[-\pi,\pi]$ to be 2π periodic, it follows the resulting function, still denoted by h, is continuous. Then using the inequality (For a better inequality, see Problem 2.) $(a+b+c)^2 \leq 4\left(a^2+b^2+c^2\right)$,

$$\int_{-\pi}^{\pi} |f - S_n f|^2 \, dx = \int_{-\pi}^{\pi} (|f-h| + |h - S_n h| + |S_n h - S_n f|)^2 \, dx$$

$$\begin{aligned}
&\leq 4 \int_{-\pi}^{\pi} \left(|f-h|^2 + |h - S_n h|^2 + |S_n h - S_n f|^2\right) dx \\
&\leq 4\varepsilon + 4 \int_{-\pi}^{\pi} |h - S_n h|^2 \, dx + 4 \int_{-\pi}^{\pi} |S_n (h - f)|^2 \, dx
\end{aligned}$$

By Theorem 12.6.1, this is dominated by $\leq 8\varepsilon + 4\int_{-\pi}^{\pi} |h - S_n h|^2 \, dx$ and by the first part, the last term converges to 0. Since ε is arbitrary, this proves the theorem. ■

12.7 Exercises

1. Suppose f has infinitely many derivatives and is also periodic with period 2π. Let the Fourier series of f be $\sum_{k=-\infty}^{\infty} a_k e^{ik\theta}$. Show that
$$\lim_{k\to\infty} k^m a_k = \lim_{k\to\infty} k^m a_{-k} = 0$$
for every $m \in \mathbb{N}$.

2. The proof of Theorem 12.6.2 used the inequality $(a+b+c)^2 \leq 4\left(a^2+b^2+c^2\right)$ whenever a, b and c are nonnegative numbers. In fact the 4 can be replaced with 3. Show this is true.

3. 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 another approach to 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 12.4. Now consider the power series of the trig functions making use of the error estimate for the remainder after m terms.

4. The inequality established above,

$$2\pi \sum_{k=-n}^{n} |a_k|^2 \leq \int_{-\pi}^{\pi} |S_n f(\theta)|^2 \, d\theta \leq \int_{-\pi}^{\pi} |f(\theta)|^2 \, d\theta$$

is called Bessel's inequality. Use this inequality to give an easy proof that for all $f \in R([-\pi,\pi]), \lim_{n\to\infty} \int_{-\pi}^{\pi} f(x) e^{inx} dx = 0$. Recall that in the Riemann Lebesgue lemma $|f| \in R((a,b])$ so while this exercise is easier, it lacks the generality of the earlier proof.

5. 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 $x = \frac{\pi}{2}$.

6. 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 $x = \pi$.

7. 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.

8. Suppose $f,g \in R([-\pi,\pi])$. Show $\frac{1}{2\pi} \int_{-\pi}^{\pi} f\overline{g} dx = \sum_{k=-\infty}^{\infty} \alpha_k \overline{\beta_k}$, where α_k are the Fourier coefficients of f and β_k are the Fourier coefficients of g.

9. Suppose $f(x) = \sum_{k=1}^{\infty} a_k \sin kx$ and that the convergence is uniform. Recall something like this holds for power series. Is it reasonable to suppose that $f'(x) = \sum_{k=1}^{\infty} a_k k \cos kx$? Explain.

10. 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. The earlier version only dealt with sums in one direction.

11. Suppose f is a differentiable function of period 2π and suppose that both f and f' are in $R([-\pi,\pi])$ 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 $\sum_{k=-\infty}^{\infty} a_k e^{ikx}$ be the Fourier series for f. Then from the definition of a_k, show that for $k \neq 0, a_k = \frac{1}{ik} a'_k$ where a'_k is the Fourier coefficient of f'. Now use the Bessel's inequality to argue that $\sum_{k=-\infty}^{\infty} \left|a'_k\right|^2 < \infty$ and then show this implies $\sum |a_k| < \infty$. You might want to use the Cauchy Schwarz inequality in Theorem 2.15.1 to do this part. Then using the version of the Weierstrass M test given in Problem 10 obtain uniform convergence of the Fourier series to f.

12. Let f be a function defined on $\mathbb{R}$. Then f is even if $f(\theta) = f(-\theta)$ for all $\theta \in \mathbb{R}$. Also f is called odd if for all $\theta \in \mathbb{R}$, $-f(\theta) = f(-\theta)$. Now using the Weierstrass approximation theorem show directly that if h is a continuous even 2π periodic function, then for every $\varepsilon > 0$ there exists an m and constants, $a_0, \cdots, a_m$ such that $\left|h(\theta) - \sum_{k=0}^{m} a_k \cos^k(\theta)\right| < \varepsilon$ for all $\theta \in \mathbb{R}$. **Hint:** Note the function arccos is continuous and maps $[-1,1]$ onto $[0,\pi]$. Using this show you can define g a continuous function on $[-1,1]$ by $g(\cos\theta) = h(\theta)$ for θ on $[0,\pi]$. Now use the Weierstrass approximation theorem on $[-1,1]$.

13. Show that if f is any odd 2π periodic function, then its Fourier series can be simplified to an expression of the form $\sum_{n=1}^{\infty} b_n \sin(nx)$ and also $f(m\pi) = 0$ for all $m \in \mathbb{N}$.

14. Consider the symbol $\sum_{k=1}^{\infty} a_n$. The infinite sum might not converge. Summability methods are systematic ways of assigning a number to such a symbol. The n^{th} Ceasaro mean σ_n is defined as the average of the first n partial sums of the series. Thus $\sigma_n \equiv \frac{1}{n}\sum_{k=1}^{n} S_k$ where $S_k \equiv \sum_{j=1}^{k} a_j$. Show that if $\sum_{k=1}^{\infty} a_n$ converges then $\lim_{n\to\infty} \sigma_n$ also exists and equals the same thing. Next find an example where, although $\sum_{k=1}^{\infty} a_n$ fails to converge, $\lim_{n\to\infty} \sigma_n$ does exist. This summability method is called Ceasaro summability. Recall the Fejer means were obtained in just this way.

15. Modify Theorem 12.5.3 to consider the case of a piecewise continuous function f. Show that at every x, $\sigma_{n+1}(f)(x) \to \frac{f(x+)-f(x-)}{2}$. This requires no extra conditions. Piecewise continuous is enough.

16. Let $0 < r < 1$ and for f a periodic function of period 2π where $f \in R([-\pi,\pi])$, consider $A_r f(\theta) \equiv \sum_{k=-\infty}^{\infty} r^{|k|} a_k e^{ik\theta}$ where the a_k are the Fourier coefficients of f. Show that if f is continuous, then $\lim_{r\to 1-} A_r f(\theta) = f(\theta)$. **Hint:** You need to find a kernel and write as the integral of the kernel convolved with f. Then consider properties of this kernel as was done with the Fejer kernel. In carrying out the details, you need to verify the convergence of the series is uniform in some sense in order to switch the sum with an integral.

17. In the above problem, if f is piecewise continuous, can you show that convergence happens to the midpoint of the jump?

18. In the formula for the Bernstein polynomials, suppose $f(0) = f(1) = 0$. Show that $p_n(0) = p_n(1) = 0$. Now if f is continuous on $\mathbb{R}$, 2π periodic, and $f(-\pi) = f(\pi)$, show there is a sequence of periodic continuous functions f_n such that f_n is a polynomial on $(-\pi, \pi)$ and $f_n(-\pi) = f(\pi)$ such that $\lim_{n\to\infty} \|f - f_n\|_\infty = 0$. Explain why f_n satisfies an appropriate Dini condition at every point and hence $\lim_{m\to\infty} S_m f_n(x) = f_n(x)$ where $S_m f_n$ is the m^{th} partial sum for the Fourier series of f_n.

12.8 The Fourier Transform

The Fourier transform is very useful in applications. It is essentially a characteristic function in probability for example. These completely characterize probability measures. It is used in many other places also. To do it right, you really ought to be using the Lebesgue integral, but this has not been discussed yet so the presentation ends up being a little fussier than it would be if it were based on a better integral.

Definition 12.8.1 *For f Riemann integrable on finite intervals, the Fourier transform is defined by*

$$Ff(t) \equiv \lim_{R\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{-itx} f(x)\, dx$$

whenever this limit exists. Of course this happens if $f \in L^1(\mathbb{R})$ thanks to Lemma 10.0.4. The inverse Fourier transform is defined the same way except you delete the minus sign in the complex exponential.

$$F^{-1}f(t) \equiv \lim_{R\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{itx} f(x)\, dx$$

Does it deserve to be called the "inverse" Fourier transform? This question will be explored somewhat below.

The next theorem justifies the terminology above which defines F^{-1} and calls it the inverse Fourier transform. Roughly it says that the inverse Fourier transform of the Fourier transform equals the mid point of the jump. Thus if the original function is continuous, it restores the original value of this function. Surely this is what you would want by calling something the inverse Fourier transform.

Now for certain special kinds of functions, the Fourier transform is indeed in L^1 and one can show that it maps this special kind of function to another function of the same sort. This can be used as the basis for a general theory of Fourier transforms. However, the following does indeed give adequate justification for the terminology that F^{-1} is called the inverse Fourier transform.

Theorem 12.8.2 *Let $g \in L^1(\mathbb{R})$ and suppose g is locally Holder continuous from the right and from the left at x as in 10.8 and 10.9, or the Jordan condition which says that g is of finite total variation on $[x-\delta, x+\delta]$ for some $\delta > 0$. Then*

$$\lim_{R\to\infty} \frac{1}{2\pi} \int_{-R}^{R} e^{ixt} \int_{-\infty}^{\infty} e^{-ity} g(y)\, dy dt = \frac{g(x+) + g(x-)}{2}.$$

Proof: Note that

$$\begin{aligned} \int_{-R}^{R} e^{ixt} \int_{-\infty}^{\infty} e^{-ity} g(y)\, dy dt &= \int_{-\infty}^{\infty} e^{-ity} g(y)\, dy \int_{-R}^{R} e^{ixt} dt \\ &= \int_{-\infty}^{\infty} e^{-ity} g(y) \int_{-R}^{R} e^{ixt} dy dt \end{aligned}$$

One merely takes a constant outside the integral and then moves a constant inside an integral. Consider the following manipulations.

$$\frac{1}{2\pi}\int_{-R}^{R} e^{ixt}\int_{-\infty}^{\infty} e^{-ity}g(y)\,dydt =$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-R}^{R} e^{ixt}e^{-ity}g(y)\,dtdy = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-R}^{R} e^{i(x-y)t}g(y)\,dtdy$$

$$\begin{aligned}
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} g(y)\left(\int_0^R e^{i(x-y)t}dt + \int_0^R e^{-i(x-y)t}dt\right)dy \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} g(y)\left(\int_0^R 2\cos((x-y)t)\,dt\right)dy
\end{aligned}$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} g(y)\frac{\sin R(x-y)}{x-y}dy = \frac{1}{\pi}\int_{-\infty}^{\infty} g(x-y)\frac{\sin Ry}{y}dy$$

$$\begin{aligned}
&= \frac{1}{\pi}\int_0^{\infty}(g(x-y)+g(x+y))\frac{\sin Ry}{y}dy \\
&= \frac{2}{\pi}\int_0^{\infty}\left(\frac{g(x-y)+g(x+y)}{2}\right)\frac{\sin Ry}{y}dy
\end{aligned}$$

From Theorem 10.2.5 or Corollary 10.2.8,

$$\begin{aligned}
&\lim_{R\to\infty}\frac{1}{2\pi}\int_{-R}^{R} e^{ixt}\int_{-\infty}^{\infty} e^{-ity}g(y)\,dydt \\
= &\lim_{R\to\infty}\frac{2}{\pi}\int_0^{\infty}\left(\frac{g(x-y)+g(x+y)}{2}\right)\frac{\sin Ry}{y}dy = \frac{g(x+)+g(x-)}{2}.\ \blacksquare
\end{aligned}$$

Also we have the Fourier cosine formula. This is interesting because you might have a function which is not periodic so there would be no hope of representing the function as a Fourier series but this next theorem says you can represent it in terms of a Fourier integral. It is the Fourier integral theorem.

Theorem 12.8.3 *Let f be piecewise continuous on every finite interval and*

$$\int_{-\infty}^{\infty}|f(y)|\,dy < \infty$$

and let $x\in(-\infty,\infty)$ satisfy the conditions 10.8 and 10.9 or the Jordan condition that f is of finite total variation on $[x-\delta, x+\delta]$ for some $\delta>0$. Then

$$\lim_{R\to\infty}\frac{1}{\pi}\int_0^R\int_{-\infty}^{\infty}\cos(t(x-y))\,f(y)\,dydt = \frac{f(x+)+f(x-)}{2} \tag{12.14}$$

Proof: Consider the following:

$$\frac{1}{\pi}\int_0^R\int_{-\infty}^{\infty}\cos(t(x-y))\,f(y)\,dydt = \frac{1}{\pi}\int_{-\infty}^{\infty}\int_0^R\cos(t(x-y))\,f(y)\,dtdy$$

The justification for this interchange of integration follows as earlier. You can discount $\int_{|y|>M} |f(y)| \, dy$ for M large enough and then use Fubini's theorem from Theorem 9.9.3 on Page 216 to interchange the order. It is the same argument in Lemma 10.1.1.

$$\begin{aligned}
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\int_{0}^{R}\cos(t(x-y)) f(y)\,dt\,dy = \frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\int_{0}^{R}\cos(t(x-y))\,dt\,dy \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\frac{\sin(R(x-y))}{x-y}dy = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x-u)\frac{\sin(Ru)}{u}dy \\
&= \frac{1}{\pi}\int_{0}^{\infty} f(x-u)\frac{\sin(Ru)}{u}dy + \frac{1}{\pi}\int_{-\infty}^{0} f(x-u)\frac{\sin(Ru)}{u}dy \\
&= \frac{1}{\pi}\int_{0}^{\infty} f(x-u)\frac{\sin(Ru)}{u}dy + \frac{1}{\pi}\int_{0}^{\infty} f(x+u)\frac{\sin(Ru)}{u}dy \\
&= \frac{1}{2\pi}\int_{0}^{\infty}\frac{f(x-u)+f(x+u)}{2u}\sin(Ru)\,du
\end{aligned}$$

Also, from Theorem 10.2.5 or Corollary 10.2.8, the limit of this as $R \to \infty$ is

$$\frac{f(x-)+f(x+)}{2}$$

This verifies 12.14. ■

You can come up with lots of enigmatic integration formulas using this representation theorem. You start with a function you want to represent and then use this formula to represent it. The representation in terms of the formula yields strange and wonderful integration facts. Two cases which are nice to note are the case that f is even and the case that f is odd. If you are only interested in the function on the nonnegative real numbers, you could consider it either way as part of an odd function or part of an even function. This will change what happens at 0. First suppose it is even. Then

$$\begin{aligned}
&\lim_{R\to\infty}\frac{1}{\pi}\int_{0}^{R}\int_{-\infty}^{\infty}\cos(t(x-y)) f(y)\,dy\,dt \\
= &\lim_{R\to\infty}\frac{1}{\pi}\int_{0}^{R}\int_{-\infty}^{\infty}(\cos tx\cos ty + \sin tx\sin ty) f(y)\,dy\,dt \\
= &\lim_{R\to\infty}\frac{1}{\pi}\int_{0}^{R}\cos tx\int_{-\infty}^{\infty}\cos ty f(y)\,dy\,dt
\end{aligned}$$

because $y \to \sin(ty)$ is odd. Thus you get

$$\begin{aligned}
\frac{f(x+)+f(x-)}{2} &= \lim_{R\to\infty}\frac{1}{\pi}\int_{0}^{R}\cos(tx)\int_{-\infty}^{\infty}\cos(ty) f(y)\,dy\,dt \\
&= \lim_{R\to\infty}\frac{2}{\pi}\int_{0}^{R}\cos(tx)\int_{0}^{\infty}\cos(ty) f(y)\,dy\,dt \qquad (12.15)
\end{aligned}$$

In case f is odd, you get

$$\begin{aligned}
\frac{f(x+)+f(x-)}{2} &= \lim_{R\to\infty}\frac{1}{\pi}\int_{0}^{R}\sin(tx)\int_{-\infty}^{\infty}\sin(ty) f(y)\,dy\,dt \\
&= \lim_{R\to\infty}\frac{2}{\pi}\int_{0}^{R}\sin(tx)\int_{0}^{\infty}\sin(ty) f(y)\,dy\,dt \qquad (12.16)
\end{aligned}$$

Lets see what this formula says for certain choices of f. We need to have f be in $L^1(\mathbb{R})$ but this is easy to arrange. Just let it vanish off some interval. First suppose $f(y) = y$ for $y \in [-1,1]$ and let it be 0 if $|y| > 1$. Of course this function has a jump at -1 and 1. Then from 12.16,

$$\begin{aligned} & \lim_{R\to\infty} \frac{2}{\pi} \int_0^R \sin(tx) \int_0^1 \sin(ty)\, y\,dy\,dt \\ = & \lim_{R\to\infty} \frac{2}{\pi} \int_0^R \sin(tx) \left(\frac{1}{t^2}(\sin t - t\cos t)\right) dt \end{aligned}$$

Thus

$$\lim_{R\to\infty} \frac{2}{\pi} \int_0^R \sin(tx) \left(\frac{1}{t^2}(\sin t - t\cos t)\right) dt = \begin{cases} x \text{ if } |x| < 1 \\ 1/2 \text{ if } |x| = 1 \\ 0 \text{ if } |x| > 1 \end{cases}$$

It might not be the first thing you would think of. I am not sure whether the integrand is even in $L^1(\mathbb{R})$ although the above limit does exist.

Now suppose that $f(y) = e^{-|y|}$. This is an even function. From 12.15

$$e^{-|x|} = \lim_{R\to\infty} \frac{2}{\pi} \int_0^R \cos(tx) \int_0^\infty \cos(ty)\, e^{-y}\,dy\,dt = \lim_{R\to\infty} \frac{2}{\pi} \int_0^R \frac{\cos(tx)}{t^2+1} dt$$

I think this is a pretty amazing formula. Obviously you can make these up all day, amazing formulas which none of the usual tools will allow you to compute.

Definition 12.8.4 *Let $f \in L^1(\mathbb{R})$. The Fourier cosine and sine transforms are defined respectively as $g(x) \equiv$*

$$\sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\cos(xt)\,dt, \quad \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\sin(xt)\,dt$$

Note that 12.15, 12.16 sort of say that if you take the cosine transform of the cosine transform, you get the function back, a similar assertion holding for the Fourier sine transform.

12.9 The Inversion of Laplace Transforms

How does the Fourier transform relate to the Laplace transform? This is considered next. Recall that from Theorem 10.3.5 if g has exponential growth $|g(t)| \leq Ce^{\lambda t}$, then if $\mathrm{Re}(s) > \lambda$, one can define $\mathscr{L}g(s)$ as

$$\mathscr{L}g(s) \equiv \int_0^\infty e^{-su} g(u)\,du$$

and also $s \to \mathscr{L}g(s)$ is differentiable on $\mathrm{Re}(s) > \lambda$ in the sense that if $h \in \mathbb{C}$ and $G(s) \equiv \mathscr{L}g(s)$, then

$$\lim_{h\to 0} \frac{G(s+h) - G(s)}{h} = G'(s) = -\int_0^\infty u e^{-su} g(u)\,du$$

Thus G is analytic and has all derivatives. Then the next theorem shows how to invert the Laplace transform. It is another one of those results which says that you get the mid point of the jump when you do a certain process. It is like what happens in Fourier series where the Fourier series converges to the midpoint of the jump under suitable conditions and like what was just shown for the inverse Laplace transform.

The next theorem gives a more specific version of what is contained in Theorem 10.4.8. However, this theorem does assume a Holder continuity condition which is not needed for Theorem 10.4.8. I think that it is usually the case that the needed Holder condition will be available.

Theorem 12.9.1 *Let g be a piecewise continuous function defined on $(0,\infty)$ which has exponential growth $|g(t)| \leq Ce^{\lambda t}$ for some real λ and $g(t)$ is Holder continuous from the right and left as in 10.8 and 10.9. For* $\operatorname{Re}(s) > \lambda$

$$\mathscr{L}g(s) \equiv \int_0^\infty e^{-su} g(u)\,du$$

Then for any $\gamma > \lambda$,

$$\lim_{R\to\infty} \frac{1}{2\pi} \int_{-R}^{R} e^{(\gamma+iy)t} \mathscr{L}g(\gamma+iy)\,dy = \frac{g(t+)+g(t-)}{2} \tag{12.17}$$

Proof: This follows from plugging in the formula for the Laplace transform of g and then using the above. Thus

$$\begin{aligned}
\frac{1}{2\pi}\int_{-R}^{R} e^{(\gamma+iy)t}\mathscr{L}g(\gamma+iy)\,dy &= \frac{1}{2\pi}\int_{-R}^{R} e^{(\gamma+iy)t}\int_{-\infty}^{\infty} e^{-(\gamma+iy)u} g(u)\,dudy \\
&= \frac{1}{2\pi}\int_{-R}^{R} e^{\gamma t}e^{iyt}\int_{-\infty}^{\infty} e^{-(\gamma+iy)u} g(u)\,dudy \\
&= e^{\gamma t}\frac{1}{2\pi}\int_{-R}^{R} e^{iyt}\int_{-\infty}^{\infty} e^{-iyu}e^{-\gamma u} g(u)\,dudy
\end{aligned}$$

Now apply Theorem 12.8.2 to conclude that

$$\begin{aligned}
&\lim_{R\to\infty}\frac{1}{2\pi}\int_{-R}^{R} e^{(\gamma+iy)t}\mathscr{L}g(\gamma+iy)\,dy = e^{\gamma t}\lim_{R\to\infty}\frac{1}{2\pi}\int_{-R}^{R} e^{iyt}\int_{-\infty}^{\infty} e^{-iyu}e^{-\gamma u} g(u)\,dudy \\
= \quad & e^{\gamma t}\frac{g(t+)e^{-\gamma t+}+g(t-)e^{-\gamma t-}}{2} = \frac{g(t+)+g(t-)}{2}. \blacksquare
\end{aligned}$$

In particular, this shows that if $\mathscr{L}g(s) = \mathscr{L}h(s)$ for all s large enough, both g,h having exponential growth, then f,g must be equal except for jumps and in fact, at any point where they are both Holder continuous from right and left, the mid point of their jumps is the same. That integral is called the Bromwich integral.

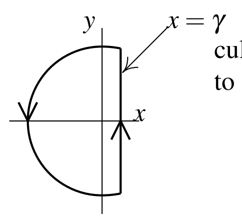
This answers the question raised earlier about whether the Laplace transform method even makes sense to use because it shows that if two functions have the same Laplace transform, then they are the same function except at jumps where the midpoint of the jumps coincide.

Using the method of residues, one can actually compute the inverse Laplace transform using this Bromwich integral. I will give an easy example, leaving out the technical details relative to estimates. These are in my book on my web site Calculus of Real and Complex Variables.

Example 12.9.2 *Find the inverse Laplace transform of $\frac{1}{1+s^2}$.*

The idea is that this is (hopefully) of the form $\mathscr{L}(f(t))$ and your task is to find $f(t)$. To do this, you use the following contour and write the above Bromwich integral as a contour integral. In this example, let $\gamma > 0$. You should verify that

$$\int_{-R}^{R} e^{(\gamma+iy)t} \mathscr{L}g(\gamma+iy)\,dy = -i\int_{\gamma_R^*} e^{zt}\frac{1}{1+z^2}dz - \left(-i\int_{C_R} e^{zt}\frac{1}{1+z^2}dz\right)$$



It turns out that in this example, and many other examples, the circular part of the contour integral, denoted here with $\int_{C_R}$, will converge to 0 as $R \to \infty$ and so

$$\begin{aligned} &\lim_{R\to\infty}\frac{1}{2\pi}\int_{-R}^{R} e^{(\gamma+iy)t}\mathscr{L}g(\gamma+iy)\,dy \\ =\ &\frac{1}{2\pi}2\pi i(-i)(\text{sum of residues}) = \text{sum of residues}\end{aligned}$$

In this example,

$$\lim_{z\to i}\frac{e^{zt}}{z+i} = \frac{1}{2i}e^{it},\ \lim_{z\to -i}\frac{e^{zt}}{(z-i)} = e^{-it}\frac{1}{-2i}$$

Thus, these add to

$$\frac{1}{2i}e^{ti} + \left(e^{-it}\frac{1}{-2i}\right) = \frac{1}{2}ie^{-it} - \frac{1}{2}ie^{it} = \sin t$$

Thus the Bromwich improper integral is $\frac{1}{2\pi}2\pi\sin t = \sin t$. We just found an inverse Laplace transform. In general, this illustrates the following procedure which works under fairly general conditions.

Procedure 12.9.3 *Let $F(s)$ satisfy $|F(z)| < C/|z|^{\alpha}$ for some $\alpha > 0$ for all $|z|$ large enough with $F(z)$ analytic except for finitely many poles. To find $f(t)$ such that $\mathscr{L}f = F$,*

$$f(t) = \left(\textit{sum of residues at the poles of } F(z)e^{tz}\right)$$

12.10 Exercises

1. Generalize the Riemann Lebesgue lemma to show that if f is piecewise continuous on finite intervals and in $L^1(\mathbb{R})$, then $\lim_{n\to\infty}\int_{-\infty}^{\infty} f(t)\sin(nt+k)\,dt = 0$. Here k is any constant.

2. Show that if f is in $L^1(\mathbb{R})$, then $\lim_{n\to\infty}\int_{-\infty}^{\infty} f(t)\cos(nt+k)\,dt = 0$.

3. In fact, all you need in the above problems is to assume that f is Riemann integrable on finite intervals and that $\lim_{R\to\infty}\int_{-R}^{R} f(t)\,dt$ exists. Show this. **Hint:** This mainly requires fussing over whether you end up with something Riemann integrable given that f is when you do certain things. I have presented it for piecewise continuous functions because that is where the interesting application resides and it is obvious that everything stays piecewise continuous. If you want to do these things right, you need to be using the Lebesgue integral anyway.

4. $\lim_{R\to\infty}\int_{-R}^{R}\frac{t}{1+t^2}dt = 0$ but $\lim_{n\to\infty}\int_{-R}^{R^2}\frac{t}{1+t^2}dt = \infty$. Show this.

5. Here are some functions which are either even or odd. Use 12.15, 12.16 to represent these functions and thereby obtain amazing integration formulas. Be sure to deal with the situation at jumps.

   (a) $f(x) = \mathscr{X}_{[-1,1]}(x)$. Thus $f(x) = 1$ on $[-1,1]$ and is 0 elsewhere. Show that

$$\lim_{R\to\infty} \frac{2}{\pi}\int_0^R \cos(tx)\left(\frac{1}{t}\sin t\right)dt = \begin{cases} 1 \text{ if } x \in (-1,1) \\ 1/2 \text{ if } |x| = 1 \\ 0 \text{ if } |x| > 1 \end{cases}$$

   Isn't this an amazing result?

   (b) $f(x) = \cos(x)$ on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and 0 if $|x| > \frac{\pi}{2}$.

   (c) $f(x) = \sin(x)$ on $(-2\pi, 2\pi)$ and 0 if $|x| > 2\pi$.

   (d) $f(x) = \cos(x)$ on $(-\pi, \pi)$ and 0 if $|x| > \pi$.

   (e) $f(x) = e^{-x}$ if $x > 0, -e^{-|x|}$ if $x < 0$.

6. Find $\lim_{R\to\infty}\int_{-1}^1 \frac{\sin(\pi t)\sin(Rt)}{t^2}dt$. **Hint:** Consider the following steps. Using the version of Fubini's theorem from Theorem 9.9.3 on Page 216

$$\int_{-1}^1 \frac{\sin(\pi t)\sin(Rt)}{t^2} = 2\int_0^1 \frac{\sin(\pi t)\sin(Rt)}{t^2}dt$$

$$= 2\int_0^1 \frac{\sin(\pi t)}{t}\int_0^R \cos(yt)\,dydt = 2\int_0^R\int_0^1 \frac{\sin(\pi t)}{t}\cos(yt)\,dtdy$$

   Now the Fourier cosine transform for the even function $x \to \frac{\sin(\pi x)}{x}\mathscr{X}_{[-1,1]} \equiv f(x)$ is

$$\lim_{R\to\infty}\frac{2}{\pi}\int_0^R \cos(xy)\int_0^1 \frac{\sin(\pi t)}{t}\cos(yt)\,dtdy = f(x)$$

   This equals the value of this function at x. It looks like what this problem is asking is $\pi f(0)$. What is $f(0)$?

7. Do something similar to the above problem in order to compute

$$\lim_{R\to\infty}\int_0^1 \frac{\sin(Rx)}{\sqrt{x}}dx$$

   Hint: $\int_0^1 \frac{\sin(Ry)}{\sqrt{y}}dy = \int_0^1 \sqrt{y}\int_0^R \cos(ty)\,dtdy = \int_0^R\int_0^1 \sqrt{y}\cos(ty)\,dydt$. Isn't this a case of the Fourier cosine transform for the even function $x \to \sqrt{|x|}$?

8. In example 12.9.2, verify the details. In particular, show that the contour integral over the circular part of the contour converges to 0 as $R \to \infty$. Explain why γ could be any positive number in this case. In general, you just need γ to be large enough so that the contour encloses all the poles and the function for which you are trying to find the inverse Laplace transform, needs to be dominated by $C/|z|^\alpha$ for some $\alpha > 0$ for all $|z|$ large enough. Try to show this. It is a little technical but only involves elementary considerations. Thus verify Procedure 12.9.3 is valid by showing that the part of the contour integral over the circular part converges to 0 if the assumed estimate holds.

9. Find inverse Laplace transforms using the method of residues in Procedure 12.9.3 for $\frac{b}{(s-a)^2+b^2}, \frac{s-a}{(s-a)^2+b^2}$, and the other entries of the table in Problem 8 on Page 248.

10. Let $f(x)$ be the odd 2π periodic extension of $f(x) = \mathscr{X}_{[0,\pi]}(x)$. Explain why its Fourier series is of the form $\sum_{k=1}^{\infty} a_n \sin(nx)$. Doing minimal computations, why can you say that na_n cannot converge to 0 as $n \to \infty$. **Hint:** See Problem 55 on Page 228.

11. Find the Fourier series expansion for the above function and use it to find interesting summation formulas, for example $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} = \frac{\pi}{4}$.

12. This and the remaining problems will require a beginning course in linear algebra. All that is needed is in any of my linear algebra books. Fill in the details. A fundamental matrix for the $n \times n$ matrix A is an $n \times n$ matrix $\Phi(t)$ having functions as entries such that $\Phi'(t) = A\Phi(t), \Phi(0) = I$. That is, you have a differential equation involving a dependent variable which is a matrix along with an initial condition $\Phi(0) = I$, the identity matrix. Using the properties of the Laplace transform, we can take the Laplace transform of both sides and get $sF(s) - I = AF(s)$ so $(sI - A)F(s) = I$ and so the Laplace transform of $\Phi(t)$ denoted here as $F(s)$ is $F(s) = (sI - A)^{-1}$. Now from linear algebra, there is a formula for this inverse valid for all s large enough which comes as $\frac{1}{\det(sI-A)}(\text{cofactor matrix })^T$. If $|s|$ is large enough, the inverse does indeed exist because there are only finitely many eigenvalues. Now each term in $(sI - A)^{-1}$ is a rational function for which the degree of the numerator is at least one more than the degree of the denominator. Thus it satisfies the necessary conditions for the Bromwich integral and thus there exists a unique such $\Phi(t)$.

13. Next show that $A\Phi(t) = \Phi(t)A$. Fill in the details. To do this, let $\Psi(t) \equiv A\Phi(t) - \Phi(t)A$. Thus $\Psi(0) = 0$. Also

$$\Psi'(t) = A\Phi'(t) - \Phi'(t)A = A^2\Phi(t) - A\Phi(t)A = A(A\Phi(t) - \Phi(t)A) = A\Psi(t)$$

In short, $\Psi'(t) = A\Psi(t), \Psi(0) = 0$. Use the Laplace transform method to show that $\Psi(t) = 0$. Also show that $\Phi(-t)\Phi(t) = I$. To do this last one, define the function $\Psi(t) \equiv \Phi(-t)\Phi(t)$,

$$\begin{aligned}
\Psi'(t) &= -\Phi'(-t)\Phi(t) + \Phi(-t)\Phi'(t) = -A\Phi(-t)\Phi(t) + \Phi(-t)A\Phi(t) \\
&= -\Phi(-t)A\Phi(t) + \Phi(-t)A\Phi(t) = 0
\end{aligned}$$

thus $\Psi(0) = I$ and $\Psi'(0) = 0$ so each entry of $\Psi(t)$ is a constant from the mean value theorem. Thus $\Psi(t) = I$ for all t. Also show that $\Phi(t)\Phi(s) = \Phi(t+s)$ using similar considerations using Laplace transforms.

14. All linear equations in an undergraduate differential equations course, which includes the vast majority of what is discussed in these courses, can be written as $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{f}(t), \mathbf{x}(0) = \mathbf{x}_0$. Fill in details. $\mathbf{x}' - A\mathbf{x} = \mathbf{f}(t)$. Multiply by $\Phi(-t)$ and you get $\frac{d}{dt}(\Phi(-t)\mathbf{x}(t)) = \Phi(-t)\mathbf{f}(t)$. Note that the left side equals the following by the product and chain rule.

$$-\Phi'(-t)\mathbf{x} + \Phi(-t)\mathbf{x}' = \Phi(-t)\mathbf{x}' - A\Phi(-t)\mathbf{x} = \Phi(-t)\left(\mathbf{x}' - A\mathbf{x}\right)$$

Now it follows from considering the individual entries of the matrices and vectors that

$$
\begin{aligned}
\Phi(-t)\mathbf{x}(t)-\mathbf{x}_0 &= \int_0^t \Phi(-s)\mathbf{f}(s)\,ds, \\
\mathbf{x}(t) &= \Phi(t)\mathbf{x}_0+\Phi(t)\int_0^t \Phi(-s)\mathbf{f}(s)\,ds \\
\mathbf{x}(t) &= \Phi(t)\mathbf{x}_0+\int_0^t \Phi(t-s)\mathbf{f}(s)\,ds
\end{aligned}
$$

This is the variation of constants formula for the unique solution to the initial value problem. This has shown that if there is a solution, then it is the above. Next verify that the above does solve the initial value problem applying fundamental theorem of calculus to the entries of the matrices. This completes somewhat more than what is accomplished in an entire undergraduate differential equations course. Furthermore, unlike what is done in these wretched busy work courses, this leads somewhere.

Chapter 13

The Generalized Riemann Integral

The preceding part of the book is essentially devoted to nineteenth century analysis. The generalized Riemann integral is a relatively recent development from around 1957. However, it is very close to the Riemann integral. One replaces the norm of the partition which is a single number with a gauge function.

13.1 Definitions and Basic Properties

This chapter is on the generalized Riemann integral. The Riemann Darboux integral presented earlier has been obsolete for over 100 years. The integral of this chapter is certainly not obsolete and is in certain important ways the very best integral currently known. This integral is called the generalized Riemann integral, also the Henstock Kurzweil integral after the two people who invented it and sometimes the gauge integral. Other books which discuss this integral are the books by Bartle [7], Bartle and Sherbert, [8], Henstock [16], or McLeod [22]. Considerably more is presented in some of these references. In what follows, F will be an increasing function, the most important example being $F(x) = x$. In the Stieltjes integral, we featured $\|P\| < \delta$. One does the same thing here except here δ is not a positive number but a positive function.

Definition 13.1.1 *Let $[a,b]$ be a closed and bounded interval. A tagged division*[1] *of $[a,b] = I$ is a set of the form $P \equiv \{(I_i, t_i)\}_{i=1}^n$ where $t_i \in I_i = [x_{i-1}, x_i]$, and $a = x_{i-1} < \cdots < x_n = b$. Let the t_i be referred to as the tags. A function $\delta : \mathbb{R} \to (0,\infty)$ is called a gauge function or simply gauge for short. A tagged division P is called δ fine if*

$$I_i \subseteq (t_i - \delta(t_i), t_i + \delta(t_i)).$$

A δ fine division, is understood to be tagged. More generally, a collection, $\{(I_i, t_i)\}_{i=1}^p$ is δ fine if the above inclusion holds for each of these intervals and their interiors are disjoint even if their union is not equal to the whole interval, $[a,b]$. In this definition, one often requires that $I_i \subseteq [t_i - \delta(t_i), t_i + \delta(t_i)]$ rather than the open interval above. It appears to not matter much in what is presented here.

The following fundamental result is essential.

Proposition 13.1.2 *Let $[a,b]$ be an interval and let δ be a gauge function on $[a,b]$. Then there exists a δ fine tagged division of $[a,b]$.*

Proof: Suppose not. Then one of $\left[a, \frac{a+b}{2}\right]$ or $\left[\frac{a+b}{2}, b\right]$ must fail to have a δ fine tagged division because if they both had such a δ fine division, the union of the two δ fine divisions would be a δ fine division of $[a,b]$. Denote by I_1 the interval which does not have a δ fine division. Then repeat the above argument, dividing I_1 into two equal intervals and pick the one, I_2 which fails to have a δ fine division. Continue this way to get a nested sequence of closed intervals, $\{I_i\}$ having the property that each interval in the set fails to have a δ fine division and $\operatorname{diam}(I_i) \to 0$. Therefore, $\cap_{i=1}^{\infty} I_i = \{x\}$ where $x \in [a,b]$. Now $(x - \delta(x), x + \delta(x))$ must contain some I_k because the diameters of these intervals converge to zero. It follows that $\{(I_k, x)\}$ is a δ fine division of I_k, contrary to the construction which required that none of these intervals had a δ fine division. ■

[1] In beginning calculus books, this is often called a partition and I followed this convention earlier. The word, division is a much better word to use. What the x_i do is to "divide" the interval into little subintervals.

With this proposition and definition, it is time to define the generalized Riemann integral. The functions being integrated typically have values in $\mathbb{R}$ or $\mathbb{C}$ but there is no reason to restrict to this situation and so in the following definition, X will denote the space in which f has its values. For example, X could be $\mathbb{R}^p$ which becomes important in multivariable calculus. For now, just think $\mathbb{C}$. It will be assumed Cauchy sequences converge and there is a norm although it is likely possible to generalize even further.

Definition 13.1.3 *For $a = x_{i-1} < \cdots < x_n = b$, and F an increasing function, ΔF_i will be defined as $F(x_i) - F(x_{i-1})$. Let X be a complete normed vector space. (For example, $X = \mathbb{R}$ or $X = \mathbb{C}$ or $X = \mathbb{R}^p$.) Then $f : [a,b] \to X$ is generalized Riemann integrable, written as $f \in R^*[a,b]$ if there exists $R \in X$ such that for all $\varepsilon > 0$, there exists a gauge δ, such that if $P \equiv \{(I_i,t_i)\}_{i=1}^n$ is δ fine, then defining $S(P,f)$ by*

$$S(P,f) \equiv \sum_{i=1}^{n} f(t_i)\Delta F_i,$$

it follows $|S(P,f) - R| < \varepsilon$. If such an R exists, then the integral is defined as follows.

$$\int_I f\,dF \equiv \int_a^b f\,dF \equiv R.$$

Here $|\cdot|$ refers to the norm on X. For $\mathbb{R}$, this is just the absolute value.

Note that if P is δ_1 fine and $\delta_1 \leq \delta$ then it follows P is also δ fine. Because of this, it follows that the generalized integral is unique if it exists.

Proposition 13.1.4 *If $R, \hat{R}$ both work in the above definition of the generalized Riemann integral, then $\hat{R} = R$.*

Proof: Let $\varepsilon > 0$ and let δ correspond to R and $\hat{\delta}$ correspond to $\hat{R}$ in the above definition. Then let $\delta_0 = \min\left(\delta, \hat{\delta}\right)$. Let P be δ_0 fine. Then P is both δ and $\hat{\delta}$ fine. Hence,

$$\left|R - \hat{R}\right| \leq |R - S(P,f)| + \left|S(P,f) - \hat{R}\right| < 2\varepsilon$$

Since ε is arbitrary, it follows that $R = \hat{R}$. ■

Is there a simple way to tell whether a given function is in $R^*[a,b]$? The following Cauchy criterion is useful to make this determination. It looks just like a similar condition for Riemann Stieltjes integration.

Proposition 13.1.5 *A function $f : [a,b] \to X$ is in $R^*[a,b]$ if and only if for every $\varepsilon > 0$, there exists a gauge function δ_ε such that if P and Q are any two divisions which are δ_ε fine, then $|S(P,f) - S(Q,f)| < \varepsilon$.*

Proof: Suppose first that $f \in R^*[a,b]$. Then there exists a gauge, δ_ε, and an element of X, R, such that if P is δ_ε fine, then $|R - S(P,f)| < \varepsilon/2$. Now let P, Q be two such δ_ε fine divisions. Then

$$|S(P,f) - S(Q,f)| \leq |S(P,f) - R| + |R - S(Q,f)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Conversely, suppose the condition of the proposition holds. Let $\varepsilon_n \to 0+$ as $n \to \infty$ and let δ_{ε_n} denote the gauge which goes with ε_n. Without loss of generality, assume that δ_{ε_n}

is decreasing in n. (If not, replace it with the minimum of itself and earlier gauges.) Let R_{ε_n} denote the closure of all the sums, $S(P,f)$ where P is δ_{ε_n} fine. From the condition, it follows $\operatorname{diam}(R_{\varepsilon_n}) \leq \varepsilon_n$ and that these closed sets are nested in the sense that $R_{\varepsilon_n} \supseteq R_{\varepsilon_{n+1}}$ because δ_{ε_n} is decreasing in n. Therefore, there exists a unique, $R \in \cap_{n=1}^{\infty} R_{\varepsilon_n}$. To see this, let $r_n \in R_{\varepsilon_n}$. Then since the diameters of the R_{ε_n} are converging to 0, $\{r_n\}$ is a Cauchy sequence which must converge to some $R \in X$. Since R_{ε_n} is closed, it follows $R \in R_{\varepsilon_n}$ for each n. Letting $\varepsilon > 0$ be given, there exists $\varepsilon_n < \varepsilon$ and for P a δ_{ε_n} fine division, $|S(P,f) - R| \leq \varepsilon_n < \varepsilon$.Therefore, $R = \int_I f$. ■

Are there examples of functions which are in $R^*[a,b]$? Are there examples of functions which are not? It turns out the second question is harder than the first although it is very easy to answer this question in the case of the obsolete Riemann integral. The generalized Riemann integral is a vastly superior integral which can integrate a very impressive collection of functions. Consider the first question. It turns out that $R[a,b] \subseteq R^*[a,b]$. Recall the definition of the Riemann integral given above which is listed here for convenience.

Definition 13.1.6 *A bounded function f defined on $[a,b]$ is said to be Riemann Stieltjes integrable if there exists a number I with the property that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if*

$$P \equiv \{x_0, x_1, \cdots, x_n\}$$

is any partition having $\|P\| < \delta$, and $z_i \in [x_{i-1}, x_i]$,

$$\left| I - \sum_{i=1}^{n} f(z_i)(F(x_i) - F(x_{i-1})) \right| < \varepsilon.$$

The number $\int_a^b f(x)\, dF$ is defined as I.

First note that if $\delta > 0$ is a number and if every interval in a division has length less than δ then the division is δ fine. In fact, you could pick the tags as any point in the intervals. Then the following theorem follows immediately.

Theorem 13.1.7 *Suppose that f is Riemann Stieltjes integrable according to Definition 13.1.6. Then f is generalized Riemann integrable and the integrals are the same.*

Proof: Just let the gauge functions be constant functions. ■

In particular, the following important theorem follows from Theorem 9.3.7.

Theorem 13.1.8 *Let f be continuous on $[a,b]$ and let F be any increasing integrator. Then $f \in R^*[a,b]$.*

This integral can integrate almost anything you can imagine, including the function which equals 1 on the rationals and 0 on the irrationals which is not Riemann integrable. This will be shown later.

The integral is linear. This will be shown next.

Theorem 13.1.9 *Suppose α and β are constants and that f and g are in $R^*[a,b]$. Then $\alpha f + \beta g \in R^*[a,b]$ and*

$$\int_I (\alpha f + \beta g)\, dF = \alpha \int_I f dF + \beta \int_I g dF.$$

Proof: Let $\eta = \frac{\varepsilon}{|\beta|+|\alpha|+1}$ and choose gauges, δ_g and δ_f such that if P is δ_g fine,

$$\left| S(P,g) - \int_I g dF \right| < \eta$$

and that if P is δ_f fine,

$$\left| S(P,f) - \int_I f dF \right| < \eta.$$

Now let $\delta = \min(\delta_g, \delta_f)$. Then if P is δ fine the above inequalities both hold. Therefore, from the definition of $S(P,f)$,

$$S(P, \alpha f + \beta g) = \alpha S(P,f) + \beta S(P,g)$$

and so

$$\left| S(P, \alpha f + \beta g) - \left(\beta \int_I g dF + \alpha \int_I f dF \right) \right| \leq \left| \beta S(P,g) - \beta \int_I g dF \right|$$

$$+ \left| \alpha S(P,f) - \alpha \int_I f dF \right| \leq |\beta| \eta + |\alpha| \eta < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows the number $\beta \int_I g dF + \alpha \int_I f dF$ qualifies in the definition of the generalized Riemann integral and so $\alpha f + \beta g \in R^* [a,b]$ and

$$\int_I (\alpha f + \beta g) dF = \beta \int_I g dF + \alpha \int_I f dF. \blacksquare$$

The following lemma is also very easy to establish from the definition.

Lemma 13.1.10 *If $f \geq 0$ and $f \in R^* [a,b]$, then $\int_I f dF \geq 0$. Also, if f has complex values and is in $R^* [I]$, then both $\operatorname{Re} f$ and $\operatorname{Im} f$ are in $R^* [I]$.*

Proof: To show the first part, let $\varepsilon > 0$ be given and let δ be a gauge function such that if P is δ fine then $|S(f,P) - \int_I f dF| \leq \varepsilon$. Since F is increasing, it is clear that $S(f,P) \geq 0$. Therefore, $\int_I f dF \geq S(f,P) - \varepsilon \geq -\varepsilon$ and since ε is arbitrary, it follows $\int_I f dF \geq 0$ as claimed.

To verify the second part, note that by Proposition 13.1.5 there exists a gauge, δ such that if P,Q are δ fine then $|S(f,P) - S(f,Q)| < \varepsilon$. But

$$\begin{aligned} |S(\operatorname{Re} f, P) - S(\operatorname{Re} f, Q)| &= |\operatorname{Re}(S(f,P)) - \operatorname{Re}(S(f,Q))| \\ &\leq |S(f,P) - S(f,Q)| \end{aligned}$$

and so the conditions of Proposition 13.1.5 are satisfied and you can conclude $\operatorname{Re} f \in R^* [I]$. Similar reasoning applies to $\operatorname{Im} f$. $\blacksquare$

Corollary 13.1.11 *If $|f|, f \in R^* [a,b]$, where f has values in $\mathbb{C}$, then*

$$\left| \int_I f dF \right| \leq \int_I |f| dF.$$

Proof: Let $|\alpha| = 1$ and $\alpha \int_I f dF = \left|\int_I f dF\right|$. Then by Theorem 13.1.9 and Lemma 13.1.10,

$$\begin{aligned} \left|\int_I f dF\right| &= \int_I \alpha f dF = \int_I (\operatorname{Re}(\alpha f) + i \operatorname{Im}(\alpha f)) dF \\ &= \int_I \operatorname{Re}(\alpha f) dF + i \int_I \operatorname{Im}(\alpha f) dF \\ &= \int_I \operatorname{Re}(\alpha f) dF \leq \int_I |f| dF \blacksquare \end{aligned}$$

Note the assumption that $|f| \in R^*[a,b]$. I will point out later that you can't assume $|f|$ is also generalized Riemann integrable. This is just like the case with series. A series may converge without converging absolutely.

The following lemma is also fundamental. It is about restricting f to a smaller interval and concluding that the function is still generalized Riemann integrable on this smaller interval.

Lemma 13.1.12 *If $f \in R^*[a,b]$ and $[c,d] \subseteq [a,b]$, then $f \in R^*[c,d]$.*

Proof: Let $\varepsilon > 0$ and choose a gauge δ such that if P is a division of $[a,b]$ which is δ fine, then $|S(P,f) - R| < \varepsilon/2$. Now pick a δ fine division of $[c,d]$, $\{(I_i,t_i)\}_{i=r}^{l}$ and let $\{(I_i,t_i)\}_{i=1}^{r-1}$, $\{(I_i,t_i)\}_{i=l+1}^{n}$ be $\boxed{\textbf{fixed}}$ δ fine divisions on $[a,c]$ and $[d,b]$ respectively.

Now let P_1 and Q_1 be δ fine divisions of $[c,d]$ and let P and Q be the respective δ fine divisions of $[a,b]$ just described which are obtained from P_1 and Q_1 by adding in $\{(I_i,t_i)\}_{i=1}^{r-1}$ and $\{(I_i,t_i)\}_{i=l+1}^{n}$. Then

$$\varepsilon > |S(P,f) - R| + |S(Q,f) - R| \geq |S(Q,f) - S(P,f)| = |S(Q_1,f) - S(P_1,f)|.$$

By the above Cauchy criterion, Proposition 13.1.5, $f \in R^*[c,d]$ as claimed. $\blacksquare$

Corollary 13.1.13 *Suppose $c \in [a,b]$ and that $f \in R^*[a,b]$. Then $f \in R^*[a,c]$ and $f \in R^*[c,b]$. Furthermore,*

$$\int_I f dF = \int_a^c f dF + \int_c^b f dF.$$

Here $\int_a^c f dF$ means $\int_{[a,c]} f dF$.

Proof: Let $\varepsilon > 0$. Let δ_1 be a gauge function on $[a,c]$ such that whenever P_1 is a δ_1 fine division of $[a,c]$,

$$\left|\int_a^c f dF - S(P_1,f)\right| < \varepsilon/3.$$

Let δ_2 be a gauge function on $[c,b]$ such that whenever P_2 is a δ_2 fine division of $[c,b]$,

$$\left|\int_c^b f dF - S(P_2,f)\right| < \varepsilon/3.$$

Let δ_3 be a gauge function on $[a,b]$ such that if P is a δ_3 fine division of $[a,b]$,

$$\left|\int_a^b f dF - S(P,f)\right| < \varepsilon/3.$$

Now define a gauge function

$$\delta(x) \equiv \begin{cases} \min(\delta_1, \delta_3) \text{ on } [a,c] \\ \min(\delta_2, \delta_3) \text{ on } [c,b] \end{cases}$$

Then letting P_1 be a δ fine division on $[a,c]$ and P_2 be a δ fine division on $[c,b]$, it follows that $P = P_1 \cup P_2$ is a δ_3 fine division on $[a,b]$ and all the above inequalities hold. Thus noting that $S(P,f) = S(P_1,f) + S(P_2,f)$,

$$\begin{aligned}
\left| \int_I f dF - \left(\int_a^c f dF + \int_c^b f dF \right) \right| &\leq \left| \int_I f dF - (S(P_1,f) + S(P_2,f)) \right| \\
&+ \left| S(P_1,f) + S(P_2,f) - \left(\int_a^c f dF + \int_c^b f dF \right) \right| \\
\leq \quad & \left| \int_I f dF - S(P,f) \right| + \left| S(P_1,f) - \int_a^c f dF \right| + \left| S(P_2,f) - \int_c^b f dF \right| \\
< \quad & \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.
\end{aligned}$$

Since ε is arbitrary, the conclusion of the corollary follows. ■

The following lemma, sometimes called Henstock's lemma is of great significance. When you have a tagged division of $[a,b]$ denoted as $\{(I_i,t_i)\}_{i=1}^q \equiv P$ with

$$I_i \subseteq (t_i - \delta(t_i), t_i + \delta(t_i))$$

so the division is δ fine, it would make perfect sense to write $\sum_{i \in \mathscr{I}} f(t_i) \Delta F_i$ for $\mathscr{I} \subseteq \{1,2,...,q\}$. If you know that $f \in R^*[a,b]$ and that

$$\left| \sum_{i=1}^q f(t_i) \Delta F_i - \int_a^b f dF \right| \leq \varepsilon$$

whenever P is δ fine, then it follows from Corollary 13.1.13 and induction that

$$\left| \sum_{i=1}^q f(t_i) \Delta F_i - \int_a^b f dF \right| = \left| \sum_{i=1}^q f(t_i) \Delta F_i - \sum_{i=1}^q \int_{I_i} f dF \right| \leq \varepsilon$$

Henstock's lemma says that you can replace the sum over all $i \in \{1,2,...,q\}$ with the sum over $i \in \mathscr{I} \subseteq \{1,2,...,q\}$ and preserve the same inequality. That is, you can take the sum over any subset of the tags. The reason such a remarkable result holds is that δ is a function, not a single number. Thus you can change it some places and not others.

Lemma 13.1.14 *Suppose that $f \in R^*[a,b]$ and that whenever Q is a δ fine division of $I = [a,b]$,*

$$\left| S(Q,f) - \int_I f dF \right| < \varepsilon.$$

Then if $P = \{(I_i,t_i)\}_{i=1}^n$ is any δ fine division of I, meaning that

$$I_i \subseteq (t_i - \delta(t_i), t_i + \delta(t_i)),$$

and $P' = \left\{ \left(I_{i_j}, t_{i_j} \right) \right\}_{j=1}^r$ is a subset of P, then

$$\left| \sum_{j=1}^r f\left(t_{i_j}\right) \Delta F_i - \sum_{j=1}^r \int_{I_{i_j}} f dF \right| \leq \varepsilon.$$

Proof: Let $(J_k, t_k) \in P \setminus P'$. From Lemma 13.1.12, $f \in R^*[J_k]$. Therefore, letting Q_k be a suitable δ fine division of J_k, using δ as a generic gauge as small as the original δ,

$$\left| \int_{J_k} f dF - S(Q_k, f) \right| < \frac{\eta}{|P \setminus P'| + 1}$$

where $\eta > 0$ and $|P \setminus P'|$ denotes the number of intervals from P which are not in P'. There are $|P \setminus P'|$ different values of k. Let $\tilde{P}$ be the division which results from all the Q_k along with P'. We modify the original δ only on the intervals of $P \setminus P'$ always making it smaller. Then

$$\varepsilon > \left| S\left(\widetilde{P}, f\right) - \int_I f dF \right|$$

$$= \left| \left(\sum_{j=1}^{r} f\left(t_{i_j}\right) \Delta F_i - \sum_{j=1}^{r} \int_{I_{i_j}} f dF \right) + \left(\sum_{k=1}^{|P \setminus P'|} S(Q_k, f) - \int_{J_k} f dF \right) \right|$$

$$\begin{aligned} &\geq \left| \sum_{j=1}^{r} f\left(t_{i_j}\right) \Delta F_i - \sum_{j=1}^{r} \int_{I_{i_j}} f dF \right| - \frac{\eta |P \setminus P'|}{|P \setminus P'| + 1} \\ &> \left| \sum_{j=1}^{r} f\left(t_{i_j}\right) \Delta F_i - \sum_{j=1}^{r} \int_{I_{i_j}} f dF \right| - \eta \end{aligned}$$

Then $\left| \sum_{j=1}^{r} f\left(t_{i_j}\right) \Delta F_i - \sum_{j=1}^{r} \int_{I_{i_j}} f dF \right| < \varepsilon + \eta$ and since η is arbitrary,

$$\left| \sum_{j=1}^{r} f\left(t_{i_j}\right) \Delta F_i - \sum_{j=1}^{r} \int_{I_{i_j}} f dF \right| \leq \varepsilon \ \blacksquare$$

Consider $\left\{(I_j, t_j)\right\}_{j=1}^{p}$ a subset of a division of $[a,b]$. If δ is a gauge and $\left\{(I_j, t_j)\right\}_{j=1}^{p}$ is δ fine, meaning that $I_j \subseteq (t_j - \delta(t_j), t_j + \delta(t_j))$, this can always be considered as a subset of a δ fine division of the whole interval and so the following corollary is just a rewording of the above.

Lemma 13.1.15 *Suppose that $f \in R^*[a,b]$ and that whenever Q is a δ fine division of $I = [a,b]$,*

$$\left| S(Q, f) - \int_I f dF \right| \leq \varepsilon.$$

Then if $\{(I_i, t_i)\}_{i=1}^{p}$ is δ fine, maybe not dividing all of I, it follows that

$$\left| \sum_{j=1}^{p} f(t_j) \Delta F(I_j) - \sum_{j=1}^{p} \int_{I_j} f dF \right| \leq \varepsilon.$$

Here is another corollary in the special case where f has real values. In this lemma, one splits the indices into those for which $f(t_i) \Delta F_i \geq \int_{I_i} f dF$ and the ones in which the inequality is turned around. You can do this because of the above corollary.

Corollary 13.1.16 *Suppose $f \in R^*[a,b]$ has values in $\mathbb{R}$ and that*

$$\left|S(P,f) - \int_I f dF\right| \leq \varepsilon$$

for all P which is δ fine. Then if $P = \{(I_i,t_i)\}_{i=1}^n$ is δ fine,

$$\sum_{i=1}^n \left|f(t_i)\Delta F_i - \int_{I_i} f dF\right| \leq 2\varepsilon. \tag{13.1}$$

Proof: Let $\mathscr{I} \equiv \left\{i : f(t_i)\Delta F_i \geq \int_{I_i} f dF\right\}$ and let $\mathscr{I}^C \equiv \{1,\cdots,n\} \setminus \mathscr{I}$. Then by Henstock's lemma

$$\left|\sum_{i\in\mathscr{I}} f(t_i)\Delta F_i - \sum_{i\in\mathscr{I}} \int_{I_i} f dF\right| = \sum_{i\in\mathscr{I}} \left|f(t_i)\Delta F_i - \int_{I_i} f dF\right| \leq \varepsilon$$

and

$$\left|\sum_{i\in\mathscr{I}^C} f(t_i)\Delta F_i - \sum_{i\in\mathscr{I}^C} \int_{I_i} f dF\right| = \sum_{i\in\mathscr{I}^C} \left|f(t_i)\Delta F_i - \int_{I_i} f dF\right| \leq \varepsilon$$

so adding these together yields 13.1. ■

This generalizes immediately to the following.

Corollary 13.1.17 *Suppose $f \in R^*[a,b]$ has values in $\mathbb{C}$ and that*

$$\left|S(P,f) - \int_I f dF\right| \leq \varepsilon \tag{13.2}$$

for all P which is δ fine. Then if $P = \{(I_i,t_i)\}_{i=1}^n$ is δ fine,

$$\sum_{i=1}^n \left|f(t_i)\Delta F_i - \int_{I_i} f dF\right| \leq 4\varepsilon. \tag{13.3}$$

Proof: It is clear that if 13.2 holds, then $|S(P,\operatorname{Re} f) - \operatorname{Re}\int_I f dF| \leq \varepsilon$, which shows that $\operatorname{Re}\int_I f dF = \int_I \operatorname{Re} f dF$. Similarly $\operatorname{Im}\int_I f dF = \int_I \operatorname{Im} f dF$. Therefore, using Corollary 13.1.16, $\sum_{i=1}^n \left|\operatorname{Re} f(t_i)\Delta F_i - \int_{I_i} \operatorname{Re} f dF\right| \leq 2\varepsilon$ and $\sum_{i=1}^n \left|i\operatorname{Im} f(t_i)\Delta F_i - \int_{I_i} i\operatorname{Im} f dF\right| \leq 2\varepsilon$. Adding and using the triangle inequality, yields 13.3. ■

13.2 Monotone Convergence Theorem

There is nothing like the following theorem in the context of Riemann integration.

Example 13.2.1 *Let $\{r_n\}_{m=1}^\infty$ be the rational numbers in $[0,1]$. Let $f_n(x)$ be 1 if $x \in \{r_1,\cdots,r_n\}$ and 0 elsewhere and $F(x) = x$. Then f_n is Riemann integrable and converges pointwise to the function $f(x)$ which is 1 on all rationals in $[0,1]$ and zero elsewhere. However, f is not Riemann integrable. Indeed, there is a gap between the upper and lower sums. See Theorem 9.3.10 on Page 200.*

In contrast to this example, here is the monotone convergence theorem.

Theorem 13.2.2 *Let $f_n(x) \geq 0$ and suppose $f_n \in R^*[a,b]$,$\cdots f_n(x) \leq f_{n+1}(x) \cdots$ and that $f(x) = \lim_{n\to\infty} f_n(x)$ for all x and that $f(x)$ has real values. Also suppose that $\left\{\int_a^b f_n dF\right\}_{n=1}^{\infty}$ is a bounded sequence. Then $f \in R^*[a,b]$ and*

$$\int_a^b f dF = \lim_{n\to\infty} \int_a^b f_n dF \tag{13.4}$$

Proof: Let $\varepsilon > 0$ be given. Let η be small enough that $3\eta + (F(b) - F(a))\eta < \varepsilon$. Since $\left\{\int_a^b f_n dF\right\}_{n=1}^{\infty}$ is increasing and bounded, there exists $I = \lim_{n\to\infty} \int_a^b f_n dF$. Therefore, there exists N such that $\left|\int_a^b f_N dF - I\right| < \eta$ so for all $n \geq N$, $\left|\int_a^b f_n dF - I\right| < \eta$. By assumption $f(x) = \lim_{j\to\infty} f_j(x)$. Now define for $j \geq N$, $F_j \equiv \left\{x : \left|f_j(x) - f(x)\right| < \eta\ \right\}$. Since $f_j(x)$ is increasing, these F_j are also increasing and since $\lim_{n\to\infty} f_n(x) = f(x)$, the union of these F_j is all of $[a,b]$. Let $E_{j+1} \equiv F_{j+1} \setminus F_j$, and $E_N \equiv F_N$ so these E_j are disjoint and they partition $[a,b]$. On E_j, E_j, $\left|f_j(x) - f(x)\right| < \eta$.

Next is a choice of fineness so that the sum will be close to the integral. Let δ_n be such that when $P = \{(I_i,t_i)\}_{i=1}^{q_n}$ is δ_n fine, then

$$\left|\sum_{i=1}^{q_n} f_n(t_i)\Delta F_i - \int_a^b f_n dF\right| < \eta 2^{-(n+1)} \tag{13.5}$$

I will choose a smaller δ in order to use just one instead of one for each $n \geq N$. Let $\delta(x)$ be defined by

$$\delta(x) \equiv \min(\delta_N(x), \cdots, \delta_j(x)) \text{ for } x \in E_j \tag{13.6}$$

Since the union of the E_j is $[a,b]$, this defines $\delta(x) > 0$ on $[a,b]$, $\delta(x)$ being positive because it is a minimum of finitely many positive numbers on each E_j, $j \geq N$. Also for any $n \geq N, \delta(x) \leq \delta_n(x)$ on $E_n \cup E_{n+1} \cup \cdots$ because δ_n is in the list so $\delta(x) \leq \delta_n(x)$ on $E_n \cup E_{n+1} \cup \cdots$. If $x \in E_j$ for some $j < n$, then $\delta(x) \leq \delta_j(x)$ so from Henstock's lemma, Lemma 13.1.15,

$$\left|\sum_{i\in\mathscr{I}_j} \int_{I_i} f_j dF - \sum_{i\in\mathscr{I}_j} f_j(t_i)\Delta F_i\right| \leq \eta 2^{-(j+1)} \tag{13.7}$$

Let P be δ fine. Modify δ to make $\delta = \hat{\delta}_n \leq \delta_n$ on the rest of $[a,b]$,other than the finitely many intervals containing the tags from $E_n \cup E_{n+1} \cup \cdots$ and consider a $\hat{\delta}_n$ fine division of $[a,b]$ which retains the intervals corresponding to the tags which are in $E_n \cup E_{n+1} \cup \cdots$. Then Henstock's lemma, Lemma 13.1.15 implies the following for $\mathscr{I}_j$ those i where the tag t_i is in E_j, $j \geq n$.

$$\left|\sum_{j=n}^{\infty}\sum_{i\in\mathscr{I}_j} f_n(t_i)\Delta F_i - \sum_{j=n}^{\infty}\sum_{i\in\mathscr{I}_j}\int_{I_i} f_n dF\right| \leq \eta 2^{-(n+1)} \leq \eta \tag{13.8}$$

This is because the double sum $\sum_{j=n}^{\infty}\sum_{i\in\mathscr{I}_j}$ selects intervals which have tags in $E_n \cup E_{n+1} \cup \cdots$ on which $\delta \leq \delta_n$. Also, for δ given as above and $N \leq j \leq n, \delta \leq \delta_j$ on $[a,b]$ from the construction. Note that P does not depend on the choice of $n \geq N$.

I want to estimate $\left|S(P,f_n) - \int_a^b f_n dF\right|$ for $n \geq N$ and this choice of a gauge function δ. The idea is to estimate in terms of two sums according to which E_j contains the tags. Then split again according to whether $j \geq n$ or $j < n$. Thus

$$\left|S(P, f_n) - \int_a^b f_n dF\right| = \left|\sum_{j=N}^{\infty} \sum_{i \in \mathscr{I}_j} f_n(t_i) \Delta F_i - \sum_{j=N}^{\infty} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF\right|$$

Next, split further according to whether $j \geq n$.

$$\left|S(P, f_n) - \int_a^b f_n dF\right| \leq \left|\sum_{j=n}^{\infty} \sum_{i \in \mathscr{I}_j} f_n(t_i) \Delta F_i - \sum_{j=n}^{\infty} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF\right| \tag{13.9}$$

$$+\left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_n(t_i) \Delta F_i - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF\right| \tag{13.10}$$

By 13.8,

$$\leq \eta + \left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_n(t_i) \Delta F_i\right| \tag{13.11}$$

Next split up the last term in 13.11.

$$\left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_n(t_i) \Delta F_i\right|$$

$$\leq \left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_j dF\right| \tag{13.12}$$

$$+\left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_j dF - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_j(t_i) \Delta F_i\right| \tag{13.13}$$

$$+\left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_j(t_i) \Delta F_i - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_n(t_i) \Delta F_i\right| \tag{13.14}$$

Then 13.12 is $\left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_n dF - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_j dF\right| =$

$$\sum_{j=N}^{n-1} \sum_{i=\mathscr{I}_j} \int_{I_i} (f_n - f_j) dF \leq \int_a^b f_n dF - \int_a^b f_j dF \leq I - \int_a^b f_N dF < \eta$$

Next consider 13.13. This term satisfies

$$\left|\sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} \int_{I_i} f_j dF - \sum_{j=N}^{n-1} \sum_{i \in \mathscr{I}_j} f_j(t_i) \Delta F_i\right| \leq$$

$$\sum_{j=N}^{\infty} \left|\sum_{i \in \mathscr{I}_j} \int_{I_i} f_j dF - \sum_{i \in \mathscr{I}_j} f_j(t_i) \Delta F_i\right| \leq \sum_{j=N}^{\infty} \eta 2^{-(j+1)} \leq \eta$$

which comes from 13.7. Next consider 13.14. Since $t_i \in E_j$, this term satisfies

$$\left|\sum_{j=N}^{n-1}\sum_{i\in\mathscr{I}_j} f_j(t_i)\Delta F_i - \sum_{j=N}^{n-1}\sum_{i\in\mathscr{I}_j} f_n(t_i)\Delta F_i\right|$$

$$\leq \sum_{j=N}^{n-1}\sum_{i\in\mathscr{I}_j}\left(f_n(t_i)-f_j(t_i)\right)\Delta F_i \leq \sum_{j=N}^{n-1}\sum_{i\in\mathscr{I}_j}\eta\Delta F_i \leq \eta\left(F(b)-F(a)\right)$$

Since $f_n(t_i)-f_j(t_i)\leq f(t_i)-f_j(t_i)<\eta$. It follows that

$$\begin{aligned}\left|S(P,f_n)-\int_a^b f_n dF\right| &\leq \eta+\eta+\eta+\eta\left(F(b)-F(a)\right)\\ &< 3\eta+\eta\left(F(b)-F(a)\right)<\varepsilon.\end{aligned}$$

Now let $n\to\infty$ and this yields $|S(P,f)-I| = \left|S(P,f)-\lim_{n\to\infty}\int_a^b f_n dF\right| < \varepsilon$. Since ε is arbitrary, it follows that $f\in R^*[a,b]$ and $I=\int_a^b f dF$. ■

13.3 Computing Generalized Integrals

Here we give some idea of how to compute these generalized Riemann integrals. It turns out they can integrate many things which the earlier Riemann Stieltjes integral cannot. In particular, you can have jumps in the integrator function at points of discontinuity of the function. You can also integrate functions which are continuous nowhere.

Example 13.3.1 *Let*

$$f(x)=\begin{cases}1 \text{ if } x\in\mathbb{Q}\\ 0 \text{ if } x\notin\mathbb{Q}\end{cases}$$

Then $f\in R^*[0,1]$ and for $F(x)=x$,

$$\int_0^1 f dF = 0.$$

This is obvious. Let $f_n(x)$ equal 1 on the first n rational numbers in an enumeration of the rationals and zero everywhere else. Clearly $f_n(x)\uparrow f(x)$ for every x and also f_n is Riemann integrable and has integral 0. Now apply the monotone convergence theorem. Note this example is one which has no Riemann or Darboux integral.

Example 13.3.2 *Let $F(x)$ be an integrator function given by*

$$F(x)\equiv\begin{cases}0 \text{ if } x\leq 1\\ 1 \text{ if } x\in(1,2)\\ 2 \text{ if } x\geq 2\end{cases}$$

Thus this has a jump at 1 and 2. Let $f(x)=\mathscr{X}_{(1,2)}(x)$. We assume $x\in\left[0,3\sqrt{2}\right]$ where the fact $\sqrt{2}$ is irrational is convenient. Note that here the integrator function and the function being integrated are both discontinuous at the two points 1,2.

Suppose $\int_0^{3\sqrt{2}} f dF$ exists, the ordinary Stieltjes integral, and let P_n consist of points $\left\{k\frac{3\sqrt{2}}{n}, k = 0, 1, \cdots, n\right\}$. Thus all division points are irrational. If this integral exists, then we would need to have a number I such that

$$|S(P_n, f) - I| < \varepsilon$$

for all n large enough. However, we can choose the tags in such a way that $S(P_n, f) = 0$ and another way such that $S(P_n, f) = 2$. Therefore, this Riemann Stieltjes integral does not exist.

However, consider the following piecewise linear continuous functions. f_m is piecewise linear, equal to 0 on $(-\infty, 1]$, 1 on $[1 + 1/m, 2 - 1/m]$, and then 0 on $[2, \infty)$. Then $\int_0^{3\sqrt{2}} f_m dF$ does exist and it is clear that the integrals are increasing in m. Using the Procedure 9.7.1, it follows that the Riemann Stieltjes integral $\int_0^{3\sqrt{2}} f_m dF = 0$. By the monotone convergence theorem, since $\lim_{m\to\infty} f_m(x) = \mathscr{X}_{(1,2)}(x)$, it follows that in terms of the generalized Riemann integral,

$$\int_0^{3\sqrt{2}} \mathscr{X}_{(1,2)}(x) \, dF$$

is defined and equals 0.

Why do we consider gauges instead of the norm of the partition? The next example will illustrate this question. Compare with the above in which the Riemann Stieltjes integral does not exist because of the freedom to pick tags on either side of the jump. A suitable gauge function can prevent this.

Example 13.3.3 *Consider the same example. This time, let*

$$\delta(x) = \begin{cases} \min\left(\frac{1}{2}|x-1|, \frac{1}{2}|x-2|\right) \text{ if } x \notin \{1,2\} \\ 1/2 \text{ if } x \in \{1,2\} \end{cases}$$

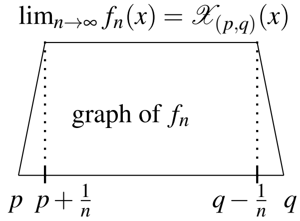
This gauge function forces 1,2 to be tags because if you have a tagged interval which contains either 1 or 2, then if t is the tag, and if it is not either 1 or 2, then the interval is not contained in $(t - \delta(t), t + \delta(t))$. Also 1,2 are both interior points of the interval containing them. Consequently if the division is δ fine, then the sum equals 0. From the definition, $\int_0^{3\sqrt{2}} f_m dF = 0$.

As pointed out in Theorem 13.1.7, if the Riemann Stieltjes integral exists, then so does the generalized Riemann Stieltjes integral and they are the same. Therefore, if you have $F(x) = x$, and I is an interval contained in $[a,b]$, then $\int_a^b \mathscr{X}_I(x) \, dF$ is the length of the interval. In general, we have the following result about the indicator function of intervals.

Proposition 13.3.4 *Suppose I is an interval and $F(x) = x$. Then $\mathscr{X}_{I\cap[a,b]} \in R^*[a,b]$.*

Proof: First suppose $I = (p,q)$. Then consider the sequence of continuous functions which are increasing in n and converge pointwise to $\mathscr{X}_{(p,q)}(x)$.

$\lim_{n\to\infty} f_n(x) = \mathscr{X}_{(p,q)}(x)$



Then $f_n \in R^* [a,b]$ and by the monotone convergence theorem, $\mathscr{X}_{(p,q)\cap[a,b]} \in R^* [a,b]$.

Next suppose $I = [p,q]$. In any case, $[p,q] \cap [a,b]$ is a closed interval which is contained in $[a,b]$. There is nothing to show if this intersection is $\emptyset$ because then $\mathscr{X}_{[p,q]\cap[a,b]}$ is 0. Consider the complement, $[p,q]^C \cap [a,b]$. This equals either the intersection of one or two open intervals with $[a,b]$. As shown above, each of these open intervals is the increasing limit of continuous functions. Hence, by the monotone convergence theorem, $\mathscr{X}_{[p,q]^C\cap[a,b]} \in R^* [a,b]$ and so $\mathscr{X}_{[p,q]\cap[a,b]} = 1- \mathscr{X}_{[p,q]^C\cap[a,b]} \in R^* [a,b]$ because the sum of generalized Riemann integrable functions is generalized Riemann integrable. Next suppose $I = (p,q]$. Then $I \cap [a,b] = (p,b], (p,q]$, or $[a,q]$. The first and last case were just considered, the first by expressing the indicator function of the interval as the increasing limit of continuous functions. It is only the second one which maybe is not clear. However, the complement of this set is $[a,p] \cup (q,b] \equiv U$ and $\mathscr{X}_{(p,q]} = 1 - \mathscr{X}_U$ where $\mathscr{X}_U \in R^* [a,b]$ and so, even in this case, $\mathscr{X}_{(p,q]} \in R^* [a,b]$. In case $I = [p,q)$, the situation is exactly similar. ■

Suppose the following conditions.

1. F is continuous and increasing on (a,b) but may have jumps at a and b.

2. Let $\hat{F}(a) \equiv F(a+), \hat{F}(b) \equiv F(b-)$

3. Let $f \in R\left([a,b], \hat{F}\right)$

Is $f \in R^* [a,b]$? What is $\int_a^b f dF$?

Let $\eta_n > 0$ be such that if $\|P\| < \eta_n$, then

$$\left|S\left(f,P,\hat{F}\right) - \int_a^b f d\hat{F}\right| < \varepsilon_n,\ \varepsilon_n \to 0$$

Now let $\delta_n(x) \equiv \min\left(\frac{1}{2}|x-a|, \frac{1}{2}|x-b|, \eta_n, \frac{1}{2}|b-a| \text{ if } x \notin \{a,b\}\right), \delta(a) = \delta(b) = \varepsilon_n$. Then let $P = \{(I_i,t_i)\}_{i=1}^{m_n}$ be a δ fine division of $[a,b]$. If a is in some I_i, then the tag for I_i must be a since if it is $t \neq a$, $(t-\delta(t), t+\delta(t))$ cannot contain a. Also, the interval containing a is not just a so in fact, a is the left end point of this interval. Similar considerations apply to b. Hence P is of the form

$$a = x_0^n < x_1^n \cdots < x_{m_n}^n = b$$

and the tags for $[a,x_1], [x_{m_n-1},b]$ are a,b. Thus $S(f,P,F)$ is of the form

$$(F(x_1^n) - F(a)) f(a) + \sum_{k=2}^{m_n-1} f(t_k)(F(x_k) - F(x_{k-1})) + f(b)\left(F(b) - F\left(x_{m_n-1}^n\right)\right)$$

Since $F = \hat{F}$ on (a,b), we can add in some terms on the top and bottom of the sum and see that this equals

$$\begin{aligned}
&(F(x_1^n) - F(a)) f(a) + \sum_{k=1}^{m_n} f(t_k)\left(\hat{F}(x_k) - \hat{F}(x_{k-1})\right) \\
&+ f(b)\left(F(b) - F\left(x_{m_n-1}^n\right)\right) \\
&- \left(f(b)\left(F(b-) - F\left(x_{m_n-1}^n\right)\right) + f(a)(F(x_1^n) - F(a+))\right)
\end{aligned}$$

As $\eta_n \to 0$, that last term $f(b)\left(F(b-) - F\left(x^n_{m_n-1}\right)\right) + f(a)(F(x^n_1) - F(a+))$ converges to 0 and so the end result is

$$(F(a+) - F(a))f(a) + \int_a^b f d\hat{F} + f(b)(F(b) - F(b-))$$

Thus this must be the generalized Riemann integral.

Procedure 13.3.5 *Suppose*

1. Suppose F is increasing and may have jumps at a and b but no jumps on (a,b).
2. Let $\hat{F}(a) \equiv F(a+), \hat{F}(b) \equiv F(b-)$
3. Let $f \in R\left([a,b], \hat{F}\right)$
   Then $f \in R^*[a,b]$ and

$$\int_a^b f dF = (F(a+) - F(a))f(a) + \int_a^b f d\hat{F} + f(b)(F(b) - F(b-))$$

This says you add the weighted jumps at the end points to the ordinary Riemann integral taken with respect to $\hat{F}$ which is continuous on $[a,b]$. What if you had two intervals $[a,b], [b,c]$ and similar conditions holding on each. Then you would obtain

$$\begin{aligned}\int_a^c f dF &= (F(a+) - F(a))f(a) + \int_a^b f d\hat{F} + f(b)(F(b+) - F(b-)) \\ &\quad + \int_b^c f d\hat{F} + f(c)(F(c) - F(c-))\end{aligned}$$

The details are left to you. Note how you pick up the whole jump at b. You could of course string together as many of these as you want. In each of the two integrals, $\hat{F}$ is adjusted to be continuous on the corresponding closed interval. What is new here? The thing which is new is that f does not need to be continuous at points where F has a jump. It is not even clear that the ordinary Rieman Stieltjes integral exists.

13.4 Integrals of Derivatives

Consider the case where $F(t) = t$. Here I will write dt for dF. The generalized Riemann integral does something very significant which is far superior to what can be achieved with other integrals. It can always integrate derivatives. Suppose f is defined on an interval, $[a,b]$ and that $f'(x)$ exists for all $x \in [a,b]$, taking the derivative from the right or left at the endpoints. What about the formula

$$\int_a^b f'(t)\, dt = f(b) - f(a)? \tag{13.15}$$

Can one take the integral of f'? If f' is continuous there is no problem of course. However, sometimes the derivative may exist and yet not be continuous. Here is a simple example.

Example 13.4.1 *Let*

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) \text{ if } x \in (0,1] \\ 0 \text{ if } x = 0 \end{cases}.$$

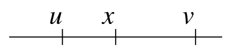
You can verify that f has a derivative on $[0,1]$ but that this derivative is not continuous.

The fact that derivatives are generalized Riemann integrable depends on the following simple lemma called the straddle lemma by McLeod [22].

Lemma 13.4.2 *Suppose $f : [a,b] \to \mathbb{R}$ is differentiable. Then there exist $\delta(x) > 0$ such that if $u \leq x \leq v$ and $u, v \in (x - \delta(x), x + \delta(x))$, then*

$$\left| f(v) - f(u) - f'(x)(v-u) \right| < \varepsilon |v-u| .$$

Proof: Consider the following picture.



From the definition of the derivative, there exists $\delta(x) > 0$ such that if $|v-x|$, $|x-u| < \delta(x)$, then

$$\left| f(u) - f(x) - f'(x)(u-x) \right| < \frac{\varepsilon}{2} |u-x|$$

and

$$\left| f'(x)(v-x) - f(v) + f(x) \right| < \frac{\varepsilon}{2} |v-x|$$

Now add these and use the triangle inequality along with the above picture to write

$$\left| f'(x)(v-u) - (f(v) - f(u)) \right| < \varepsilon |v-u| . \blacksquare$$

The next proposition says 13.15 makes sense for the generalized Riemann integral.

Proposition 13.4.3 *Suppose $f : [a,b] \to \mathbb{R}$ is differentiable. Then $f' \in R^*[a,b]$ and*

$$f(b) - f(a) = \int_a^b f' dx$$

where the integrator function is $F(x) = x$.

Proof: Let $\varepsilon > 0$ be given and let $\delta(x)$ be such that the conclusion of the above lemma holds for ε replaced with $\varepsilon/(b-a)$. Then let $P = \{(I_i, t_i)\}_{i=1}^n$ be δ fine. Then using the triangle inequality and the result of the above lemma with $\Delta x_i = x_i - x_{i-1}$,

$$\begin{aligned} \left| f(b) - f(a) - \sum_{i=1}^n f'(t_i) \Delta x_i \right| &= \left| \sum_{i=1}^n f(x_i) - f(x_{i-1}) - f'(t_i) \Delta x_i \right| \\ &\leq \sum_{i=1}^n \varepsilon / (b-a) \Delta x_i = \varepsilon . \blacksquare \end{aligned}$$

With this proposition there is a very simple statement of the integration by parts formula, the product rule gives a very simple version of integration by parts.

Corollary 13.4.4 *Suppose f, g are differentiable on $[a,b]$. Then $f'g \in R^*[a,b]$ if and only if $g'f \in R^*[a,b]$ and in this case,*

$$fg|_a^b - \int_a^b fg' dx = \int_a^b f' g dx$$

The following example, is very significant. It exposes an unpleasant property of the generalized Riemann integral. You can't multiply two generalized Riemann integrable functions together and expect to get one which is generalized Riemann integrable. This is like the case of summation. You know that if you multiply the terms of two conditionally convergent series, the resulting series is not necessarily convergent. Also, just because f is generalized Riemann integrable, you cannot conclude $|f|$ is. Again, this is like the situation with summation. This is very different than the case of the Riemann integral. It is unpleasant from the point of view of pushing symbols. The reason for this unpleasantness is that there are so many functions which can be integrated by the generalized Riemann integral.

Example 13.4.5 *Consider the function*

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$$

Then $f'(x)$ exists for all $x \in \mathbb{R}$ and equals

$$f'(x) = \begin{cases} 2x \sin\left(\frac{1}{x^2}\right) - \frac{2}{x}\cos\left(\frac{1}{x^2}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$$

Then f' is generalized Riemann integrable on $[0,1]$ because it is a derivative. Now let $\psi(x)$ denote the sign of $f'(x)$. Thus

$$\psi(x) \equiv \begin{cases} 1 \text{ if } f'(x) > 0 \\ -1 \text{ if } f'(x) < 0 \\ 0 \text{ if } f'(x) = 0 \end{cases}$$

Then ψ is a bounded function and you can argue it is Riemann integrable on $[0,1]$. However, $\psi(x) f'(x) = |f'(x)|$ and this is not generalized Riemann integrable .

Although you can't in general multiply two generalized integrable functions and get one which is also a generalized integrable function as shown in the example, sometimes you can. The following computation will show what the conditions should be.

Let $h \in R^*[a,b]$ with F an increasing integrator function which we assume is continuous here. This last assumption could be generalized. What are conditions for g such that hg is also $R^*[a,b]$? Let $\varepsilon_n \to 0$ and let δ_n be a gauge function such that if P_n is a tagged division which is δ_n fine, then

$$\left| \int_a^b h dF - S(P,h) \right| < \varepsilon_n$$

Without loss of generality, $\delta_n(x) \leq \varepsilon_n$. Say P_n consists of the division points $a = x_0^n < \cdots < x_{m_n}^n = b$. By Henstock's lemma, for t_i^m the tags,

$$\left| \sum_{i=1}^{k} h(t_i^n)\left(F(x_i^n) - F\left(x_{i-1}^n\right)\right) - \int_a^{x_k^n} h dF \right| < \varepsilon_n$$

Let $H_k^n \equiv \sum_{i=1}^k h(t_i^n)\left(F(x_i^n) - F\left(x_{i-1}^n\right)\right), H_0^n \equiv 0$. Thus, from the above, if $k \geq 1$

$$\left| H_k^n - \int_a^{x_k^n} h dF \right| = \left| e_{n_k} \right| < \varepsilon_n \tag{13.16}$$

Then

$$\sum_{k=1}^{m_n} h\left(t_k^n\right) g\left(t_k^n\right)\left(F\left(x_k^n\right)-F\left(x_{k-1}^n\right)\right)=\sum_{k=1}^{m_n}\left(H_k^n-H_{k-1}^n\right) g\left(t_k^n\right)$$

$$=\sum_{k=1}^{m_n} H_k^n g\left(t_k^n\right)-\sum_{k=0}^{m_n-1} H_k^n g\left(t_{k+1}^n\right)=\sum_{k=1}^{m_n-1} H_k^n\left(g\left(t_k^n\right)-g\left(t_{k+1}^n\right)\right)+H_{m_n}^n g\left(t_{m_n}^n\right) \tag{13.17}$$

Applying 13.16, this equals

$$\sum_{k=1}^{m_n-1}\left(\int_a^{x_k^n} h d F\right)\left(g\left(t_k^n\right)-g\left(t_{k+1}^n\right)\right)+H_{m_n}^n g\left(t_{m_n}^n\right) \tag{13.18}$$

$$+\sum_{k=1}^{m_n-1}\left(e_{n_k}\right)\left(g\left(t_k^n\right)-g\left(t_{k+1}^n\right)\right) \tag{13.19}$$

I want to estimate the last term. To do this, suppose g is either increasing or decreasing. Then this last term is dominated by

$$\varepsilon_n|g(b)-g(a)|$$

and so it converges to 0 as $n \rightarrow \infty$. The function $x \rightarrow \int_a^x h d F$ is continuous. See Problem 4 on Page 320 and the following problem. Therefore, since $\delta_n \leq \varepsilon_n$ the last term in 13.18 converges to

$$\left(\int_a^b h d F\right) g(b-)$$

The remaining term in 13.18 is just a Riemann sum for a continuous function having an integrator function given by an increasing function. Therefore, since $\delta_n \rightarrow 0$ the norm of the partitions consisting of the division points converges to 0 and so this Riemann sum, added to the other terms in 13.18 and 13.19 converges to

$$\int_a^b \int_a^x h d F d g+\left(\int_a^b h d F\right) g(b-)$$

thanks to Theorem 9.3.7 about the existence of the Riemann Stieltjes integral.

Of course you can also use Proposition 9.3.2 about functions of bounded variation being the difference of two increasing functions to conclude g could be a real valued bounded variation function. This proves the following theorem.

Theorem 13.4.6 *Let $f \in R^*[a,b]$ where the increasing integrator function F is continuous and suppose g is of bounded variation. Then $fg \in R^*[a,b]$ also.*

The proof of this theorem, patterned after the proof of the Dirichlet test for convergence of series, shows why you have to assume something more on g. This requirement, along with the fact that $f \in R^*[a,b]$ does not imply $|f| \in R^*[a,b]$ is really OBNOXIOUS. The reason for this is that the generalized Riemann integral can be like conditional convergence. Recall how strange things could take place. In the next chapter I will present a general abstract framework for Lebesgue integration. This is like absolutely convergent series and so many of the strange things will disappear and the resulting integral is much easier to use in applications. It also is the integral for the study of probability.

13.5 Exercises

1. Prove that if $f_n \in R^*[a,b]$ and $\{f_n\}$ converges uniformly to f, and $|f_n - f_m| \in R^*[a,b]$ for each m,n, then $f \in R^*[a,b]$ and $\lim_{n\to\infty} \int_I f_n = \int_I f$.

2. Suppose the integrator function is $F(x) = x$. Show that for I any interval, $\mathscr{X}_I$ is Riemann Stieltjes integrable and if $I \subseteq [a,b]$, then $\int_a^b \mathscr{X}_I dx$ is the length of I.

3. In Example 13.4.5 there is the function given
$$g(x) \equiv \begin{cases} 2x\sin\left(\frac{1}{x^2}\right) - \frac{2}{x}\cos\left(\frac{1}{x^2}\right) \text{ if } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$$
It equals the derivative of a function as explained in this example. Thus g is generalized Riemann integrable on $[0,1]$. Show that $h(x) = \max(0, g(x))$ and $h(x) = |g(x)|$ are not generalized Riemann integrable.

4. Let $f \in R^*[a,b]$ and consider the function $x \to \int_a^x f(t)\,dt$. Is this function continuous? Explain. **Hint:** Let $\varepsilon > 0$ be given and let a gauge δ be such that if P is δ fine then
$$\left|S(P,f) - \int_a^b f dx\right| < \varepsilon/2$$
Now pick $h < \delta(x)$ for some $x \in (a,b)$ such that $x+h < b$. Then consider the single tagged interval, $([x,x+h],x)$ where x is the tag. By Corollary 13.1.15
$$\left|f(x)h - \int_x^{x+h} f(t)\,dt\right| < \varepsilon/2.$$
Now you finish the argument and show f is continuous from the right. A similar argument will work for continuity from the left.

5. Generalize Problem 4 to the case where the integrator function is continuous. What if the integrator function is not continuous at x? Can you say that continuity holds at every point of continuity of F?

6. If F is a real valued increasing function, show that it has countably many points of discontinuity.

7. If $C \equiv \{r_i\}_{i=1}^{\infty}$ is a countable set in $[a,b]$, show that $\mathscr{X}_C$ is in $R^*[a,b]$. **Hint:** Let $C_n = \{r_1, \cdots, r_n\}$ and explain why $\mathscr{X}_{C_n}$ is generalized Riemann integrable. Then use the monotone convergence theorem.

8. Prove the first mean value theorem for integrals for the generalized Riemann integral in the case that $x \to \int_a^x f(t)\,dF$ is continuous.

9. Suppose $f, |f| \in R^*[a,b]$ and f is continuous at $x \in [a,b]$. Show $G(y) \equiv \int_a^y f(t)\,dt$ is differentiable at x and $G'(x) = f(x)$.

10. Suppose f has $n+1$ derivatives on an open interval containing c. Show using induction and integration by parts that
$$f(x) = f(c) + \sum_{k=1}^{n} \frac{f^k(c)}{k!}(x-c)^k + \frac{1}{n!}\int_c^x f^{(n+1)}(t)(x-t)^n\,dt.$$

Is this formula even valid for ordinary Riemann or Darboux integrals? **Hint:** It is fine if you assume f^{n+1} is continuous.

11. The ordinary Riemann integral is only applicable to bounded functions. However, the Generalized Riemann integral has no such restriction. Let $f(x) = x^{-1/2}$ for $x > 0$ and 0 for $x = 0$. Find $\int_0^1 x^{-1/2} dx$. **Hint:** Let $f_n(x) = 0$ for $x \in [0, 1/n]$ and $x^{-1/2}$ for $x > 1/n$. Now consider each of these functions and use the monotone convergence theorem.

12. Do the above problem directly from the definition without involving the monotone convergence theorem. This involves choosing an auspicious gauge function. Define the function to equal 0 at 0. It is undefined at this point so make it 0 there.

13. Can you establish a version of the monotone convergence theorem which has a decreasing sequence of functions, $\{f_k\}$ rather than an increasing sequence?

14. For E a subset of $\mathbb{R}$ and F an increasing integrator function, define E to be "measurable" if $\mathscr{X}_{E\cap[a,b]} \in R^*[a,b]$ for each interval $[a,b]$ and in this case, let
$$\mu(E) \equiv \sup\left\{\int_{-n}^{n} \mathscr{X}_E(t)\, dF : n \in \mathbb{N}\right\}$$
Show that if each E_k is measurable and the E_k are disjoint, then so is $\cup_{k=1}^{\infty} E_k$ and if E is measurable, then so is $\mathbb{R}\setminus E$. Show that intervals are all "measurable". **Hint:** This will involve the monotone convergence theorem. Thus the collection of measurable sets is closed with respect to countable disjoint unions and complements.

15. Nothing was said about the function being bounded in the presentation of the generalized Riemann integral. In Problem 51 on Page 227, it was shown that you need to have the function bounded if you are going to have the definition holding for a Riemann integral to exist. However, in the case of the generalized Riemann integral, this is not necessary. Suppose $F(x) = x$ so you have the usual Riemann type integral and let
$$f(x) = \begin{cases} 1/\sqrt{x} \text{ if } x \in (0,1] \\ 0 \text{ if } x = 0 \end{cases}$$
Letting $\varepsilon > 0$, consider $\delta(x) = \min(|x|, \varepsilon)$ for $x \neq 0$ and $\delta(0) = \varepsilon$. If you have $a > 0$ then $\int_a^b \frac{1}{\sqrt{x}} dx = \frac{1}{\sqrt{t}}(b-a)$ for some $t \in (a,b)$ thanks to the mean value theorem for integrals. This t could be a tag for the interval $[a,b]$ or it might be close enough to a tag γ that $\left|\frac{1}{\sqrt{t}} - \frac{1}{\sqrt{\gamma}}\right|$ is small. Modify δ on $[\varepsilon, 1]$ a compact set on which $1/\sqrt{x}$ is continuous. Compare the Riemann sums with the tags and $\int_{\varepsilon}^1 1/\sqrt{x} dx$. When you do this, it really looks a lot like the standard method of finding an improper Riemann integral.

16. Let F be an increasing integrator function. In Proposition 13.3.4 an argument was given which showed that if I is an interval contained on the interior of an interval (a,b), then $\mathscr{X}_I$ was in $R^*[a,b]$. However, this was not computed. In this problem we do this. Let $a < \alpha < \beta < b$. For $x \notin \{\alpha, \beta\}$, let
$$\delta(x) \equiv \min(|x-\alpha|, |x-\beta|)$$

and let $\delta(\alpha) = \delta(\beta) = \varepsilon > 0$. If you have any δ fine division P of $[a,b]$, note that if t is a tag such that $t \notin \{\alpha, \beta\}$, then $(t - \delta(t), t + \delta(t))$ cannot contain α and it cannot contain β. Therefore, both α, β are tags. Furthermore, explain why each of α and β is an interior point of the interval for which they are tags. Now explain why the division points are

$$a = x_0 < \cdots < x_k < \alpha < x_{k+2} < \cdots < x_m < \beta < x_{m+2} < \cdots < x_n = b$$

where $(x_{k+2} - x_k) < 2\varepsilon, (x_{m+2} - x_m) < 2\varepsilon$. Now explain why if $I = (\alpha, \beta)$, then $S(P, \mathscr{X}_I) = F(x_m) - F(x_{k+2})$ and if $I = (\alpha, \beta]$, $S(P, \mathscr{X}_I) = F(x_{m+2}) - F(x_{k+2})$ and if $I = [\alpha, \beta)$, then $S(P, f) = F(x_m) - F(x_k)$ and if $I = [\alpha, \beta]$, then $S(P, f) = F(x_{m+2}) - F(x_k)$. Now as $\varepsilon \to 0$, show these converge respectively to $F(\beta-) - F(\alpha+)$, $F(\beta+) - F(\alpha+)$, $F(\beta-) - F(\alpha-)$, and $F(\beta+) - F(\alpha-)$. Thus indicator functions of intervals are generalized Riemann integrable and we can even compute them. Consider a case where (α, β) is not contained in $[a,b]$ and use the same method to compute $\int_a^b \mathscr{X}_{(\alpha,\beta)} dF$. Also consider intervals of the form (α, ∞).

17. The gamma function is defined for $x > 0$ as

$$\Gamma(x) \equiv \int_0^\infty e^{-t} t^{x-1} dt \equiv \lim_{R \to \infty} \int_0^R e^{-t} t^{x-1} dt$$

Show the integral from 0 to R exists as a generalized Riemann integral directly from the definition. Also show that

$$\Gamma(x+1) = x\Gamma(x),\ \Gamma(1) = 1.$$

How does $\Gamma(n)$ for n an integer compare with $(n-1)!$? This was all done earlier for the improper Riemann integrals. Why is there no change with generalized Riemann integrals?

Chapter 14

The Lebesgue Integral

This short chapter is on the Lebesgue integral. The emphasis is on the abstract Lebesgue integral which is a general sort of construction depending on a measure space. It is an introduction to this topic. I will use the generalized Riemann integral to give non trivial examples in which the measure space is linked to the real line, thus tying it in to the topic of this book, advanced calculus for functions of one variable. Probably, these are the most important examples. However, the complete development of this topic is in other sources like [25] or [17]. This is also in my on line analysis books. As mentioned earlier, this integral can't do some of the things the generalized Riemann integral can, but it is a lot easier to use if you are interested in things like function spaces or probability which are typically built on this integral. Also, you can consider absolute values of integrable functions and get functions for which the integral at least makes sense since this integral is free of some of the pathology associated with the generalized Riemann integral. The right way to do all of this is by the use of functionals defined on continuous functions which vanish off some interval and to use the Riemann Stieltjes integrals, but to save trouble, I will emphasize the measure of sets directly because the machinery of the generalized integral has been developed. The example of measures on $\mathbb{R}$ is based on Dynkin's lemma, a very useful result in probability which is interesting for its own sake. This integral is like absolute convergent series whereas the generalized Riemann integral is more like the inclusion of conditionally convergent series.

14.1 Measures

The definition of a measure is given next. It is a very general notion so I am presenting it in this way. The case of main interest here is where $\Omega = \mathbb{R}$. However, if you want to study mathematical statistics or probability, it is very useful to understand this general formulation. Surely the study of the integral should lead somewhere. It turns out that the machinery developed makes it very easy to extend to Lebesgue measure on appropriate subsets of $\mathbb{R}^p$ also, but this will not be done in this book because this is a book on one variable ideas. See my on line analysis books to see this done.

Definition 14.1.1 *Let Ω be a nonempty set. A σ algebra $\mathscr{F}$ is a set whose elements are subsets of Ω which satisfies the following.*

1. *If $E_i \in \mathscr{F}$, for $i = 1,2,\cdots$, then $\cup_{i=1}^{\infty} E_i \in \mathscr{F}$.*
2. *If $E \in \mathscr{F}$, then $E^C \equiv \Omega \setminus E \in \mathscr{F}$*
3. *$\emptyset, \Omega$ are both in $\mathscr{F}$*

$\mu : \mathscr{F} \to [0,\infty]$ *is called a measure if whenever $E_i \in \mathscr{F}$ and $E_i \cap E_j = \emptyset$ for all $i \neq j$, then*

$$\mu\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu\left(E_i\right)$$

that sum is defined as $\sup_n \sum_{i=1}^{n} \mu\left(E_i\right)$. It could be a real number or $+\infty$. Such a pair $(\Omega, \mathscr{F})$ is called a measurable space. If you add in μ, written as $(\Omega, \mathscr{F}, \mu)$, it is called a measure space.

Example 14.1.2 *As a simple example, let $\Omega = \mathbb{N}$ and let $\mathscr{F} = \mathscr{P}(\mathbb{N})$ the set of all subsets, and let $\mu(E)$ be the number of elements of E. You might verify that this is a measure space.*

Observation 14.1.3 *If $(\Omega, \mathscr{F})$ is a measurable space and $E_i \in \mathscr{F}$, then $\cap_{i=1}^{\infty} E_i \in \mathscr{F}$. This is because $E_i \in \mathscr{F}$ and by DeMorgan's laws,*

$$\cap_{i=1}^{\infty} E_i = \left(\cup_{i=1}^{\infty} E_i^C\right)^C \in \mathscr{F} \text{ since each } E_i^C \in \mathscr{F}$$

Measures have the following fundamental property.

Lemma 14.1.4 *If μ is a measure and $F_i \in \mathscr{F}$, then $\mu(\cup_{i=1}^{\infty} F_i) \leq \sum_{i=1}^{\infty} \mu(F_i)$. Also if $F_n \in \mathscr{F}$ and $F_n \subseteq F_{n+1}$ for all n, then if $F = \cup_n F_n$,*

$$\mu(F) = \lim_{n \to \infty} \mu(F_n)$$

Symbolically, if $F_n \uparrow F$, then $\mu(F_n) \uparrow \mu(F)$. If $F_n \supseteq F_{n+1}$ for all n, then if $\mu(F_1) < \infty$ and $F = \cap_n F_n$, then

$$\mu(F) = \lim_{n \to \infty} \mu(F_n)$$

Symbolically, if $\mu(F_1) < \infty$ and $F_n \downarrow F$, then $\mu(F_n) \downarrow \mu(F)$.

Proof: Let $G_1 = F_1$ and if $G_1, \cdots, G_n$ have been chosen disjoint, let

$$G_{n+1} \equiv F_{n+1} \setminus \cup_{i=1}^{n} G_i$$

Thus the G_i are disjoint. In addition, these are all measurable sets. Now

$$\mu(G_{n+1}) + \mu(F_{n+1} \cap (\cup_{i=1}^{n} G_i)) = \mu(F_{n+1})$$

and so $\mu(G_n) \leq \mu(F_n)$. Therefore,

$$\mu(\cup_{i=1}^{\infty} G_i) = \mu(\cup_{i=1}^{\infty} F_i) = \sum_i \mu(G_i) \leq \sum_i \mu(F_i).$$

Now consider the increasing sequence of $F_n \in \mathscr{F}$. If $F \subseteq G$ and these are sets of $\mathscr{F}$

$$\mu(G) = \mu(F) + \mu(G \setminus F)$$

so $\mu(G) \geq \mu(F)$. Also

$$F = \cup_{i=1}^{\infty} (F_{i+1} \setminus F_i) + F_1$$

Then

$$\mu(F) = \sum_{i=1}^{\infty} \mu(F_{i+1} \setminus F_i) + \mu(F_1)$$

Now $\mu(F_{i+1} \setminus F_i) + \mu(F_i) = \mu(F_{i+1})$. If any $\mu(F_i) = \infty$, there is nothing to prove. Assume then that these are all finite. Then

$$\mu(F_{i+1} \setminus F_i) = \mu(F_{i+1}) - \mu(F_i)$$

and so $\mu(F) =$

$$\sum_{i=1}^{\infty} \mu(F_{i+1}) - \mu(F_i) + \mu(F_1) = \lim_{n \to \infty} \sum_{i=1}^{n} \mu(F_{i+1}) - \mu(F_i) + \mu(F_1) = \lim_{n \to \infty} \mu(F_{n+1})$$

Next suppose $\mu(F_1) < \infty$ and $\{F_n\}$ is a decreasing sequence. Then $F_1 \setminus F_n$ is increasing to $F_1 \setminus F$ and so by the first part,

$$\mu(F_1) - \mu(F) = \mu(F_1 \setminus F) = \lim_{n\to\infty} \mu(F_1 \setminus F_n) = \lim_{n\to\infty} (\mu(F_1) - \mu(F_n))$$

This is justified because $\mu(F_1 \setminus F_n) + \mu(F_n) = \mu(F_1)$ and all numbers are finite by assumption. Hence $\mu(F) = \lim_{n\to\infty} \mu(F_n)$. ■

Now I will specify a collection of subsets of $\mathbb{R}$ which I will refer to as measurable.

Definition 14.1.5 *Given an integrator function F let*

$$\mathscr{S} \equiv \{E \subseteq \mathbb{R} : \mathscr{X}_E \in R^* [p,q] \text{ for all closed intervals } [p,q]\}$$

I will call these sets "measurable".

Then these so called measurable sets satisfy the following:

Lemma 14.1.6 *The following hold.*

1. *All open intervals are in $\mathscr{S}$.*
2. *If $\{E_i\}_{i=1}^{\infty}$ are disjoint sets in $\mathscr{S}$ then $\cup_{i=1}^{\infty} E_i \in \mathscr{S}$ also.*
3. *If $E \in \mathscr{S}$, then $E^C \equiv \mathbb{R} \setminus E$ is also in $\mathscr{S}$.*

Proof: Consider 1. Let (a,b) be an open interval. There are several possibilities for $(a,b) \cap [p,q]$

1. $(a,b) \cap [p,q] = (\alpha,\beta) \subseteq [p,q]$,
2. $(a,b) \cap [p,q] = (\alpha,q] \subseteq [p,q]$,
3. $(a,b) \cap [p,q] = [p,\beta) \subseteq [p,q]$, or
4. $(a,b) \cap [p,q] = [p,q]$.

In case 1., let ψ_n vanish off (a,b), be continuous, and increase to $\mathscr{X}_{(\alpha,\beta)}$. Then $\psi_n \in R([p,q])$ so it is in $R^*([a,b])$ and by the monotone convergence $\mathscr{X}_{(\alpha,\beta)} \in R^*([a,b])$.

Next consider 2. Is $\mathscr{X}_{(\alpha,q]} \in R^*[p,q]$? Yes, and a similar argument to 1.) will hold. just get a sequence of functions continuous on $[p,q]$ and increasing to $\mathscr{X}_{(\alpha,q]}$ and use a similar argument.

Case 3.) is entirely similar.

The last case is obvious from Lemma 9.7.2. In fact, $\int_p^q \mathscr{X}_{[p,q]} dF = F(q+) - F(p-)$, this being the ordinary Riemann Stieltjes integral.

Letting $E \equiv \cup_{i=1}^{\infty} E_i$, the E_i being disjoint, why is $\mathscr{X}_E \in R^*[p,q]$? It follows from what was done above that $\mathscr{X}_{\cup_{i=1}^{m} E_i \cap [a,b]} \in R^*[p,q]$ because, since the sets are disjoint, $\mathscr{X}_{\cup_{i=1}^{m} E_i}$ equals $\sum_{i=1}^{m} \mathscr{X}_{E_i}$ and it was shown that the sum of functions in $R^*[p,q]$ is in $R^*[p,q]$. Now apply the monotone convergence theorem as $m \to \infty$ to conclude that $\mathscr{X}_{\cup_{i=1}^{\infty} E_i} \in R^*[p,q]$.

Next consider 3. If $E \in \mathscr{S}$, then $\mathscr{X}_{E^C} = 1 - \mathscr{X}_E$ and the two summands on the right are in $R^*[p,q]$ so the function on the left is also. ■

14.2 Dynkin's Lemma

Rather than attempt to show $\mathscr{S}$ is a σ algebra, I will show that $\mathscr{S}$ contains a σ algebra. This is fairly easy because of a very elegant lemma due to Dynkin which is part of the abstract theory of measures and integrals. This lemma is more interesting than the assertion that $\mathscr{S}$ contains a σ algebra.

Lemma 14.2.1 *Let $\mathscr{C}$ be a set whose elements are σ algebras each containing some subset $\mathscr{K}$ of the set of all subsets. Then $\cap\mathscr{C}$ is a σ algebra which contains $\mathscr{K}$.*

Proof: $\emptyset, \Omega$ are in $\cap\mathscr{C}$ because these are each in each σ algebra of $\mathscr{C}$. If $E_i \in \cap\mathscr{C}$, then if $\mathscr{F} \in \mathscr{C}$ it follows that $\cup_{i=1}^{\infty} E_i \in \mathscr{F}$ and so, since $\mathscr{F}$ is arbitrary, this shows this union is in $\cap\mathscr{C}$. If $E \in \cap\mathscr{C}$, then $E^C \in \mathscr{F}$ for each $\mathscr{F} \in \cap\mathscr{C}$ and so, as before, $E^C \in \cap\mathscr{C}$. Thus $\cap\mathscr{C}$ is a σ algebra. ■

Definition 14.2.2 *Let Ω be a set and let $\mathscr{K}$ be a collection of subsets of Ω. Then $\mathscr{K}$ is called a π system if $\emptyset, \Omega \in \mathscr{K}$ and whenever $A, B \in \mathscr{K}$, it follows $A \cap B \in \mathscr{K}$. $\sigma(\mathscr{K})$ will denote the intersection of all σ algebras containing $\mathscr{K}$. The set of all subsets of Ω is one such σ algebra which contains $\mathscr{K}$. Thus $\sigma(\mathscr{K})$ is the smallest σ algebra which contains $\mathscr{K}$.*

The following is the fundamental lemma which shows these π systems are useful. This is due to Dynkin. Note that the open intervals in $\mathbb{R}$ constitute a π system.

Lemma 14.2.3 *Let $\mathscr{K}$ be a π system of subsets of Ω, a set. Also let $\mathscr{G}$ be a collection of subsets of Ω which satisfies the following three properties.*

1. *$\mathscr{K} \subseteq \mathscr{G}$*
2. *If $A \in \mathscr{G}$, then $A^C \in \mathscr{G}$*
3. *If $\{A_i\}_{i=1}^{\infty}$ is a sequence of disjoint sets from $\mathscr{G}$ then $\cup_{i=1}^{\infty} A_i \in \mathscr{G}$.*

Then $\mathscr{G} \supseteq \sigma(\mathscr{K})$, where $\sigma(\mathscr{K})$ is the smallest σ algebra which contains $\mathscr{K}$.

Proof: First note that if

$$\mathscr{H} \equiv \{\mathscr{G} : 1 \text{ - } 3 \text{ all hold for } \mathscr{G}\}$$

then $\cap\mathscr{H}$ yields a collection of sets which also satisfies 1 - 3. Therefore, I will assume in the argument that $\mathscr{G}$ is the smallest collection satisfying 1 - 3. Let $A \in \mathscr{K}$ and define

$$\mathscr{G}_A \equiv \{B \in \mathscr{G} : A \cap B \in \mathscr{G}\}.$$

I want to show $\mathscr{G}_A$ satisfies 1 - 3 because then it must equal $\mathscr{G}$ since $\mathscr{G}$ is the smallest collection of subsets of Ω which satisfies 1 - 3. This will give the conclusion that for $A \in \mathscr{K}$ and $B \in \mathscr{G}$, $A \cap B \in \mathscr{G}$. This information will then be used to show that if $A, B \in \mathscr{G}$ then $A \cap B \in \mathscr{G}$. From this it will follow very easily that $\mathscr{G}$ is a σ algebra which will imply it contains $\sigma(\mathscr{K})$. Now here are the details of the argument.

Since $\mathscr{K}$ is given to be a π system, $\mathscr{K} \subseteq \mathscr{G}_A$. Property 3 is obvious because if $\{B_i\}$ is a sequence of disjoint sets in $\mathscr{G}_A$, then $A \cap \cup_{i=1}^{\infty} B_i = \cup_{i=1}^{\infty} A \cap B_i \in \mathscr{G}$ because $A \cap B_i \in \mathscr{G}$ and the property 3 of $\mathscr{G}$.

It remains to verify Property 2 so let $B \in \mathscr{G}_A$. I need to verify that $B^C \in \mathscr{G}_A$. In other words, I need to show that $A \cap B^C \in \mathscr{G}$. However, from De Morgan's laws,

$$A \cap B^C = \left(A^C \cup B\right)^C = \left(A^C \cup (A \cap B)\right)^C$$

Now $A^C \in \mathscr{G}$ because $A \in \mathscr{K} \subseteq \mathscr{G}$ and $\mathscr{G}$ is closed with respect to complements. Also, since $B \in \mathscr{G}_A, A \cap B \in \mathscr{G}$ and so $A^C \cup (A \cap B) \in \mathscr{G}$ because $\mathscr{G}$ is closed with respect to disjoint unions. Therefore, $\left(A^C \cup (A \cap B)\right)^C \in \mathscr{G}$ because $\mathscr{G}$ is closed with respect to complements. Thus $B^C \in \mathscr{G}_A$ as hoped. Thus $\mathscr{G}_A$ satisfies 1 - 3 and this implies, since $\mathscr{G}$ is the smallest such, that $\mathscr{G}_A \supseteq \mathscr{G}$. However, $\mathscr{G}_A$ is constructed as a subset of $\mathscr{G}$. This proves that for every $B \in \mathscr{G}$ and $A \in \mathscr{K}$, $A \cap B \in \mathscr{G}$. Now pick $B \in \mathscr{G}$ and consider

$$\mathscr{G}_B \equiv \{A \in \mathscr{G} : A \cap B \in \mathscr{G}\}.$$

I just proved $\mathscr{K} \subseteq \mathscr{G}_B$. The other arguments are identical to show $\mathscr{G}_B$ satisfies 1 - 3 and is therefore equal to $\mathscr{G}$. This shows that whenever $A, B \in \mathscr{G}$ it follows $A \cap B \in \mathscr{G}$.

This implies $\mathscr{G}$ is a σ algebra. To show this, all that is left is to verify $\mathscr{G}$ is closed under countable unions because then it follows $\mathscr{G}$ is a σ algebra. Let $\{A_i\} \subseteq \mathscr{G}$. Then let $A_1' = A_1$ and

$$A_{n+1}' \equiv A_{n+1} \setminus \left(\cup_{i=1}^n A_i\right) = A_{n+1} \cap \left(\cap_{i=1}^n A_i^C\right) = \cap_{i=1}^n \left(A_{n+1} \cap A_i^C\right) \in \mathscr{G}$$

because the above showed that finite intersections of sets of $\mathscr{G}$ are in $\mathscr{G}$. Since the A_i' are disjoint, it follows $\cup_{i=1}^\infty A_i = \cup_{i=1}^\infty A_i' \in \mathscr{G}$ Therefore, $\mathscr{G} \supseteq \sigma(\mathscr{K})$. ■

14.3 The Lebesgue Stieltjes Measures and Borel Sets

The σ algebra of interest here consists of $\mathscr{B}(\mathbb{R})$, the Borel sets of $\mathbb{R}$. $\mathscr{B}(\mathbb{R})$ is defined as the smallest σ algebra which contains the open sets.

Definition 14.3.1 *Let $\mathscr{B}(\mathbb{R})$ denote $\sigma(\mathscr{O})$ where $\mathscr{O}$ denotes the set of all open sets of $\mathbb{R}$.*

Then the following lemma is available.

Lemma 14.3.2 *Let $\mathscr{I}$ denote the set of open intervals. Then $\sigma(\mathscr{I}) = \mathscr{B}(\mathbb{R})$.*

Proof: By Theorem 6.5.9, every open set is a countable or finite union of open intervals. Therefore, each open set is contained in $\sigma(\mathscr{I})$. It follows that $\sigma(\mathscr{I}) \supseteq \mathscr{B}(\mathbb{R}) \equiv \sigma(\mathscr{O}) \supseteq \sigma(\mathscr{I})$. ■

Let $\mathscr{G}$ be those Borel sets E satisfy $\mathscr{X}_{E \cap [p,q]} \in R^*[p,q]$. Thus $\mathscr{G} \subseteq \mathscr{B}(\mathbb{R}) = \sigma(\mathscr{I})$. By Lemma 14.1.6 $\mathscr{G}$ contains the open intervals $\mathscr{I}$ and is closed with respect to countable disjoint unions and complements. Hence, from Dynkin's lemma, $\mathscr{G} \supseteq \sigma(\mathscr{I})$ and so $\mathscr{S} \supseteq \mathscr{B}(\mathbb{R})$, see Definition 14.1.5. This has proved the following.

Theorem 14.3.3 *Every set E in $\mathscr{B}(\mathbb{R})$ is measurable which means that the indicator function $\mathscr{X}_E$ is in $R^*[p,q]$ for any $[p,q]$.*

From this, it is easy to define the Lebesgue Stieltjes measures on the Borel sets. I will give the definition first and then show that it really is a measure.

Definition 14.3.4 *Let $\mu = \mu_F$ be defined on $\mathscr{B}(\mathbb{R})$ as follows.*

$$\mu(E) \equiv \sup_{n\in\mathbb{N}} \int_{-n}^{n} \mathscr{X}_{E\cap[-n,n]}(x)\,dF = \lim_{n\to\infty} \int_{-n}^{n} \mathscr{X}_{E\cap[-n,n]}(x)\,dF$$

Does this definition make sense? I need to verify the limit exists because

$$\left\{ \int_{-n}^{n} \mathscr{X}_{E\cap[-n,n]}(x)\,dF \right\}_{n=1}^{\infty}$$

is an increasing sequence. If unbounded, then the limit is ∞ and if bounded, it is the real limit of the sequence. Thus the definition does make sense.

Lemma 14.3.5 $\left\{\int_{-n}^{n} \mathscr{X}_{E\cap[-n,n]}(x)\,dF\right\}_{n=1}^{\infty}$ *is increasing in n. Also, if E is a bounded set contained in* $[-(n-1), n-1]$, *then* $\int_{-(n+1)}^{n+1} \mathscr{X}_{E\cap[-(n+1),n+1]}(x)\,dF = \int_{-n}^{n} \mathscr{X}_{E\cap[-n,n]}(x)\,dF$.

Proof: $\int_{-(n+1)}^{n+1} \mathscr{X}_{E\cap[-(n+1),n+1]}(x)\,dF =$

$$\int_{-(n+1)}^{-n} \mathscr{X}_{E\cap[-(n+1),n+1]}(x)\,dF + \int_{-n}^{n} \mathscr{X}_{E\cap[-(n+1),n+1]}(x)\,dF$$

$$+ \int_{n}^{n+1} \mathscr{X}_{E\cap[-(n+1),n+1]}(x)\,dF \geq \int_{-(n+1)}^{n+1} \mathscr{X}_{E\cap[-n,n]}(x)\,dF$$

If $E \subseteq [-(n-1), n-1]$, then

$$\int_{-(n+1)}^{n+1} \mathscr{X}_{E\cap[-(n+1),n+1]}(x)\,dF = \int_{-(n+1)}^{n+1} \mathscr{X}_{E\cap[-(n-1),n-1]}(x)\,dF$$

$$\begin{aligned} = \quad & \int_{-(n+1)}^{-n} \mathscr{X}_{E\cap[-(n-1),n-1]}(x)\,dF + \int_{-n}^{n} \mathscr{X}_{E\cap[-(n-1),n-1]}(x)\,dF \\ & + \int_{n}^{n+1} \mathscr{X}_{E\cap[-(n-1),n-1]}(x)\,dF \\ & = \int_{-n}^{n} \mathscr{X}_{E\cap[-(n-1),n-1]}(x)\,dF = \int_{-n}^{n} \mathscr{X}_{E\cap[-n,n]}(x)\,dF \end{aligned}$$

because on $[-(n+1), -n]$ and $[n, n+1]$ all the sums in defining the integrals are 0. Thus, if E is bounded, the integrals giving the measure as a limit are eventually constant and $\mu(E) < \infty$. ■

Next I need to verify that this is a measure. Let $\{E_i\}_{i=1}^{\infty}$ be disjoint sets in $\mathscr{B}(\mathbb{R})$. Then $\cup_{i=1}^{\infty} E_i \in \mathscr{B}(\mathbb{R})$ and so, from the monotone convergence theorem for generalized Riemann integrals,

$$\begin{aligned} \mu\left(\cup_{i=1}^{\infty} E_i\right) \quad & \equiv \quad \sup_n \int_{-n}^{n} \mathscr{X}_{\cup_{i=1}^{\infty} E_i \cap[-n,n]}(x)\,dF = \sup_n \lim_{m\to\infty} \int_{-n}^{n} \mathscr{X}_{\cup_{i=1}^{m} E_i\cap[-n,n]}(x)\,dF \\ & = \quad \sup_n \lim_{m\to\infty} \int_{-n}^{n} \sum_{i=1}^{m} \mathscr{X}_{E_i\cap[-n,n]} dF = \sup_n \sup_m \int_{-n}^{n} \sum_{i=1}^{m} \mathscr{X}_{E_i\cap[-n,n]} dF \end{aligned}$$

$$
\begin{aligned}
&= \sup_n \sup_m \sum_{i=1}^m \int_{-n}^n \mathscr{X}_{E_i\cap[-n,n]} dF = \sup_m \sup_n \sum_{i=1}^m \int_{-n}^n \mathscr{X}_{E_i\cap[-n,n]} dF \\
&= \sup_m \lim_{n\to\infty} \sum_{i=1}^m \int_{-n}^n \mathscr{X}_{E_i\cap[-n,n]} dF = \sup_m \sum_{i=1}^m \lim_{n\to\infty} \int_{-n}^n \mathscr{X}_{E_i\cap[-n,n]} dF \\
&= \sup_m \sum_{i=1}^m \mu(E_i) \equiv \sum_{i=1}^\infty \mu(E_i)
\end{aligned}
$$

In this computation, I have used the interchange of limits with supremums in case of an increasing sequence. Also, I have used the monotone convergence theorem. Therefore, this has only shown the desired result in case $\mu\left(\cup_{i=1}^\infty E_i\right)$ is finite because of the monotone convergence theorem we currently have. However, in case this is infinity, let l be a real number. Then by definition, there is n large enough that $\int_{-n}^n \mathscr{X}_{\cup_{i=1}^\infty E_i\cap[-n,n]}(x)\,dF > l$. If $\int_{-n}^n \mathscr{X}_{\cup_{i=1}^m E_i\cap[-n,n]}(x)\,dF \leq l$ for each m, then by the monotone convergence theorem, $\int_{-n}^n \mathscr{X}_{\cup_{i=1}^\infty E_i\cap[-n,n]}(x)\,dF \leq l$ also, from the monotone convergence theorem for generalized integrals, which would be a contradiction. Hence for large enough m,

$$
\sum_{i=1}^\infty \mu(E_i) \geq \int_{-n}^n \mathscr{X}_{\cup_{i=1}^m E_i\cap[-n,n]}(x)\,dF = \sum_{i=1}^m \int_{-n}^n \mathscr{X}_{E_i\cap[-n,n]} dF > l
$$

Since l is arbitrary, it follows that in this case, both $\mu\left(\cup_{i=1}^\infty E_i\right)$ and $\sum_{i=1}^\infty \mu(E_i)$ equal ∞. To understand measure of intervals here is a lemma.

For $[c,d] \subseteq (-n,n)$ it is not clear that $\int_{-n}^n \mathscr{X}_{[c,d]} dF$ the Riemann Stieltjes integral even exists. This is because $\mathscr{X}_{[c,d]}$ is not continuous and we do not assume F is continuous either. In particular, you could have F have a jump at c or at d. But with the generalized integral, one can get the appropriate result.

Theorem 14.3.6 *Let F be an increasing integrator function. Then*

1. $\mu_F([c,d]) = F(d+) - F(c-)$
2. $\mu_F((c,d)) = F(d-) - F(c+)$
3. $\mu_F((c,d]) = F(d+) - F(c+)$
4. $\mu_F([c,d)) = F(d-) - F(c-)$

Proof: For large $n, \mu_F([c,d]) = \int_{-n}^n \mathscr{X}_{[c,d]} dF, [c,d] \subseteq (-n,n)$. Define the gauge function

$$
\delta_\varepsilon(x) \equiv \min(|x-c|, |x-d|) \text{ if } x \notin \{c,d\} \text{ and } \delta_\varepsilon(c) = \delta_\varepsilon(d) = \varepsilon > 0
$$

Let P_ε be a δ_ε fine division of $[-n,n]$. Then both c,d are tags because, due to the definition of δ_ε, neither of these can be closer than $\delta_\varepsilon(t)$ for any $t \notin \{c,d\}$ because of the definition of δ_ε. For example, if c is not a tag, then there is some tag x such that $c \in (x-\delta_\varepsilon(x), x+\delta_\varepsilon(x))$ and so $\delta_\varepsilon(x) > |c-x|$ which does not happen. Similarly d must also be a tag. If I_c is the interval containing c then c is on the interior of I_c. Otherwise, the adjacent interval having t for a tag and c an endpoint, would have $c \in (t-\delta_\varepsilon(t), t+\delta_\varepsilon(t))$ which would say that $|c-t| < \delta_\varepsilon(t)$ which does not happen due to the definition of δ_ε. Similarly d is an interior point of I_d the closed interval containing d. Thus the division points x_j are

$$
-n = x_0 < \cdots < x_k < c < x_{k+1} < x_{k+2} < \cdots < x_m < d < x_{m+1} < x_{m+2} < \cdots < x_l = n
$$

where $x_{k+1}-x_k \leq 2\varepsilon$ and $x_{m+1}-x_m \leq 2\varepsilon$. In writing down a sum corresponding to this δ_ε division, it reduced to $F(x_{m+1})-F(x_k)$. Letting $\varepsilon \to 0$ yields the integral and it equals $F(d+)-F(c-)$. this shows 1.).

For 2.) $(c,d)=\cup_k\left[c+\frac{1}{k},d-\frac{1}{k}\right]$ for all j suitably large. Thus from Lemma 14.1.4,

$$\begin{aligned}
\mu_F((c,d)) &= \lim_{k\to\infty}\mu_F\left(\left[c+\frac{1}{k},d-\frac{1}{k}\right]\right) \\
&= \lim_{k\to\infty}\left(F\left(\left(d-\frac{1}{k}\right)+\right)-F\left(\left(c+\frac{1}{k}\right)-\right)\right) \\
&= \lim_{k\to\infty}\left(F\left(d-\frac{1}{k}\right)-F\left(c+\frac{1}{k}\right)\right)=F(d-)-F(c+)
\end{aligned}$$

For 3.) similar reasoning to the above using $(c,d]=\cup_k\left[c+\frac{1}{k},d\right]$,

$$\mu_F((c,d])=\lim_{k\to\infty}F(d+)-F\left(c+\frac{1}{k}\right)=F(d+)-F(c+).$$

Part 4.) is entirely similar. ■

14.4 Regularity

This has to do with approximating with certain special sets. This is of utmost importance if you want to use the Lebesgue integral in any significant way, especially for various function spaces. I will show that under reasonable conditions this needed regularity is automatic.

Definition 14.4.1 *A set is called F_σ if it is the countable union of closed sets. A set is called G_δ if it is the countable intersection of open sets.*

Lemma 14.4.2 *If A is an F_σ set, then if I is any interval, finite or infinite, $A\cap I$ is also an F_σ set. If A is a G_δ set, then if I is any interval, then $I\cap A$ is also a G_δ set.*

Proof: Consider the following example in which $I=[a,b)$. Say $A=\cup_{k=1}^\infty H_k$ where H_k is closed. $I=\cup_{j=1}^\infty\left[a,b-\frac{1}{j}\right]$ and so

$$\begin{aligned}
A\cap I &= \left(\cup_{k=1}^\infty H_k\right)\cap\cup_{j=1}^\infty\left[a,b-\frac{1}{j}\right]=\cup_{j=1}^\infty\left[a,b-\frac{1}{j}\right]\cap\cup_{k=1}^\infty H_k \\
&= \cup_{j=1}^\infty\cup_{k=1}^\infty\left(H_k\cap\left[a,b-\frac{1}{j}\right]\right)
\end{aligned}$$

which is still an F_σ set. It is still a countable union of closed sets. Other cases are similar.

Now consider the case where A is G_δ. Say $A=\cap_{k=1}^\infty V_k$ for V_k open. Consider the same half open interval. In this case, $[a,b)=\cap_{j=1}^\infty\left(a-\frac{1}{j},b\right)$. Then

$$A\cap I=\cap_{k=1}^\infty V_k\cap\cap_{j=1}^\infty\left(a-\frac{1}{j},b\right)=\cap_{j=1}^\infty\cap_{k=1}^\infty\left(V_k\cap\left(a-\frac{1}{j},b\right)\right)$$

which is a G_δ set, still being a countable intersection of open sets. Other cases are similar. ■

In particular, the above lemma shows that all intervals are both G_δ and F_σ as are finite unions of intervals. Indeed, $\mathbb{R}$ is both open and closed so in the lemma, you could take $A = \mathbb{R}$. Also note that $\emptyset$ is both open and closed. In what follows, μ is a measure which is defined on $\mathscr{B}(\mathbb{R})$ such that μ is finite on all bounded sets. We don't really need to worry about where it comes from. However, this is a good time to review the properties of measures in Lemma 14.1.4. I will use these as needed.

A measure satisfying the conclusions of the following theorem is called a regular measure.

Theorem 14.4.3 *Let μ be a measure defined on $\mathscr{B}(\mathbb{R})$ which is finite on bounded sets. Then for all E a Borel set, there is an F_σ set F and a G_δ set G such that*

$$F \subseteq E \subseteq G, \text{ and } \mu(G \setminus F) = 0 \tag{14.1}$$

Also for all E Borel,

$$\begin{aligned} \mu(E) &= \sup\{\mu(K), K \subseteq E \text{ and } K \text{ compact}\} \\ \mu(E) &= \inf\{\mu(V), V \supseteq E \text{ and } V \text{ open}\} \end{aligned} \tag{14.2}$$

Proof: Letting $\mathscr{I}$ denote the open intervals, recall that $\sigma(\mathscr{I}) = \mathscr{B}(\mathbb{R})$. Let

$$A_k = [-(k+1), -k) \cup [k, k+1),\ k = 0, 1, \ldots$$

These A_k are disjoint and partition $\mathbb{R}$, each being both G_δ and F_σ. The following is a definition of Borel sets such that the set intersected with each of these A_k is perfectly approximated from the inside and outside by a F_σ and G_δ set respectively. Thus

$$\mathscr{G} \equiv \left\{ \begin{array}{c} E \in \mathscr{B}(\mathbb{R}) : F_k \subseteq E \cap A_k \subseteq G_k, \\ \mu(G_k \setminus F_k) = 0, G_k \subseteq A_k \end{array} \right\},$$

for some G_k a G_δ set and F_k an F_σ set, this holding for all $k \in 0, 1, 2, \cdots$. From Lemma 14.4.2, $\mathscr{I} \subseteq \mathscr{G}$.

I want to show that $\mathscr{G}$ is closed with respect to countable disjoint unions and complements. Consider complements first. Say E is Borel and

$$F_k \subseteq E \cap A_k \subseteq G_k,\ \ \mu(G_k \setminus F_k) = 0,\ G_k \subseteq A_k$$

From Lemma 14.4.2, each A_k is both G_δ and F_σ. Thus

$$A_k \cap G_k^C \subseteq E^C \cap A_k \subseteq A_k \cap F_k^C$$

From Lemma 14.4.2, the left end is F_σ and the right end is G_δ. This is because the complement of a G_δ is an F_σ and the complement of an F_σ is a G_δ. Now

$$\begin{aligned} & \mu\left(A_k \cap F_k^C \setminus \left(A_k \cap G_k^C\right)\right) = \mu\left(\left(A_k \cap F_k^C\right) \cap \left(A_k^C \cup G_k\right)\right) \\ = \ & \mu\left(A_k \cap F_k^C \cap G_k\right) = \mu(G_k \setminus F_k) = 0 \end{aligned}$$

since $G_k \subseteq A_k$. Thus $\mathscr{G}$ is closed with respect to complements.

Next let $E_i \in \mathscr{G}$ and let the E_i be disjoint. Is $\cup_{i=1}^\infty E_i \equiv E \in \mathscr{G}$? Say for each k

$$F_k \subseteq E_i \cap A_k \subseteq G_k,\ \ \mu(G_k \setminus F_k) = 0,\ G_k \subseteq A_k$$

Then $F \equiv \cup_k F_k$ is F_σ because it is still a countable union of closed sets. Recall that the countable union of countable sets is countable. Also $\mu(E \setminus F) \leq \sum_k \mu(E_k \setminus F_k) = 0$. However, it may not be clear why $\cup_k G_k$ would be G_δ. However, the above implies that there exists an open set $V_i \supseteq E_i \cap A_k$ such that $\mu(E_i \cap A_k) + \frac{\varepsilon}{2^k 2^{i+1}} > \mu(V_i)$. Then if $V_k \equiv \cup_{i=1}^{\infty} V_i$,

$$\mu(V_k) \leq \sum_i \mu(V_i) < \sum_{i=0}^{\infty} \left(\mu(E_i \cap A_k) + \frac{\varepsilon}{2^k 2^{i+1}} \right) = \frac{\varepsilon}{2^k} + \mu(E \cap A_k),\ V_k \supseteq E \cap A_k$$

$$\mu(V_k \setminus (E \cap A_k)) < \frac{\varepsilon}{2^k}$$

Then $\mu\left(\cup_{k=0}^{\infty} V_k \setminus E\right) \leq \sum_{k=0}^{\infty} \mu(V_k \setminus (E \cap A_k)) < 2\varepsilon$. It follows that there exists open W_n containing E such that $\mu(W_n \setminus E) < 1/n$. These W_n can be assumed decreasing. Thus if $G \equiv \cap_n W_n, \mu(G \setminus E) = 0$. Hence $G \supseteq E \supseteq F$ and $\mu(G \setminus F) = \mu(G \setminus E) + \mu(E \setminus F) = 0$. Thus $\mathscr{G}$ is closed with respect to complements and countable disjoint unions so from Lemma 14.2.3 it contains $\sigma(\mathscr{I}) = \mathscr{B}(\mathbb{R})$ but $\mathscr{G}$ was defined to consist of sets of $\mathscr{B}(\mathbb{R})$ so $\mathscr{G} = \mathscr{B}(\mathbb{R})$.

The first claim 14.1 was just shown. Let $l < \mu(E)$ then $\mu(E \cap [-n,n]) > l$ for large enough n and so there is a closed set K contained in $E \cap [-n,n]$ such that $l < \mu(K)$ also. This shows the first of 14.2. There is nothing to show in the second if $\mu(E) = \infty$. So assume $\mu(E)$ is finite. Then letting G be from the first part, $G = \cap_n W_n$ where W_n is open and these are decreasing open sets. We can assume $\mu(W_1) < \mu(E) + 1$ from the argument given above to show 14.1. Thus $\mu(G) = \mu(E) = \lim_{n\to\infty} \mu(W_n)$ and so for large $n, \mu(E) + \varepsilon > \mu(W_n)$. This shows the second part of 14.1. ■

This shows that all those Lebesgue Stieltjes measures are regular.

Next is to define the kind of function which can be integrated. The measure space of this section dealing with the Lebesgue Stieltjes measures is specific to $\mathbb{R}$ but what comes next is the general notion in an abstract measure space.

14.5 Measurable Functions

You can integrate nonnegative measurable functions. All this will be presented in general. Thus the functions are defined on a measure space. I am going to present this in the general setting but you can apply it to the measure space just developed consisting of the Lebesgue Stieltjes measure on $\mathbb{R}$ or on the counting measure of Example 14.1.2.

Notation 14.5.1 *In whatever context $f^{-1}(S) \equiv \{\omega : f(\omega) \in S\}$. It is called the inverse image of S and everything in the theory of the Lebesgue integral is formulated in terms of inverse images. For a real valued f, $f^{-1}(\lambda,\infty)$ may sometimes be written as $[f > \lambda]$.*

Lemma 14.5.2 *Let $f : \Omega \to (-\infty,\infty]$ where $\mathscr{F}$ is a σ algebra of subsets of Ω. The following are equivalent.*

$$f^{-1}((d,\infty]) \in \mathscr{F} \text{ for all finite } d,$$

$$f^{-1}((-\infty,d)) \in \mathscr{F} \text{ for all finite } d,$$

$$f^{-1}([d,\infty]) \in \mathscr{F} \text{ for all finite } d,$$

$$f^{-1}((-\infty,d]) \in \mathscr{F} \text{ for all finite } d,$$

$$f^{-1}((a,b)) \in \mathscr{F} \text{ for all } a < b, -\infty < a < b < \infty.$$

Definition 14.5.3 *Any of these equivalent conditions in the above lemma is what is meant the statement that f is measurable.*

Proof of the lemma: First note that the first and the third are equivalent. To see this, observe
$$f^{-1}([d,\infty]) = \cap_{n=1}^{\infty} f^{-1}((d-1/n,\infty]),$$
and so if the first condition holds, then so does the third.
$$f^{-1}((d,\infty]) = \cup_{n=1}^{\infty} f^{-1}([d+1/n,\infty]),$$
and so if the third condition holds, so does the first.

Similarly, the second and fourth conditions are equivalent. Now
$$f^{-1}((-\infty,d]) = (f^{-1}((d,\infty]))^C$$
so the first and fourth conditions are equivalent. Thus the first four conditions are equivalent and if any of them hold, then for $-\infty < a < b < \infty$,
$$f^{-1}((a,b)) = f^{-1}((-\infty,b)) \cap f^{-1}((a,\infty]) \in \mathscr{F}.$$
Finally, if the last condition holds,
$$f^{-1}([d,\infty]) = \left(\cup_{k=1}^{\infty} f^{-1}((-k+d,d))\right)^C \in \mathscr{F}$$
and so the third condition holds. Therefore, all five conditions are equivalent. ■

From this, it is easy to verify that pointwise limits of a sequence of measurable functions are measurable.

Corollary 14.5.4 *If $f_n(\omega) \to f(\omega)$ where all functions have values in $(-\infty,\infty]$, then if each f_n is measurable, so is f.*

Proof: Note the following:
$$f^{-1}\left((b+\frac{1}{l},\infty]\right) = \cup_{k=1}^{\infty} \cap_{n\geq k} f_n^{-1}\left((b+\frac{1}{l},\infty]\right) \subseteq f^{-1}\left(\left[b+\frac{1}{l},\infty\right]\right)$$
This follows from the definition of the limit. Therefore,
$$\begin{aligned} f^{-1}((b,\infty]) &= \cup_{l=1}^{\infty} f^{-1}\left((b+\frac{1}{l},\infty]\right) = \cup_{l=1}^{\infty}\cup_{k=1}^{\infty}\cap_{n\geq k} f_n^{-1}\left((b+\frac{1}{l},\infty]\right) \\ &\subseteq \cup_{l=1}^{\infty} f^{-1}\left(\left[b+\frac{1}{l},\infty\right]\right) = f^{-1}((b,\infty]) \end{aligned}$$
The messy term on the middle is measurable because it consists of countable unions and intersections of measurable sets. It equals $f^{-1}((b,\infty])$ and so this last set is also measurable. By Lemma 14.5.2, f is measurable. ■

A convenient way to check measurability is in terms of limits of simple functions.

Definition 14.5.5 *Let $(\Omega,\mathscr{F})$ be a measurable space. A simple function is one which is measurable but has only finitely many values.*

Note that a simple function achieves each value on a measurable set.

Theorem 14.5.6 *Let $f \geq 0$ be measurable. Then there exists a sequence of nonnegative simple functions $\{s_n\}$ satisfying*

$$0 \leq s_n(\omega) \tag{14.3}$$

$$\cdots s_n(\omega) \leq s_{n+1}(\omega) \cdots$$

$$f(\omega) = \lim_{n\to\infty} s_n(\omega) \text{ for all } \omega \in \Omega. \tag{14.4}$$

If f is bounded, the convergence is actually uniform. Conversely, if f is nonnegative and is the pointwise limit of such simple functions, then f is measurable.

Proof: Letting $I \equiv \{\omega : f(\omega) = \infty\}$, define

$$t_n(\omega) = \sum_{k=0}^{2^n} \frac{k}{n} \mathscr{X}_{f^{-1}\left(\left[\frac{k}{n}, \frac{k+1}{n}\right)\right)}(\omega) + 2^n \mathscr{X}_I(\omega).$$

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 $t_n(\omega) = 2^n$ for $\omega \in I$ and if $f(\omega) \in [0, \frac{2^n+1}{n})$, then

$$0 \leq f(\omega) - t_n(\omega) \leq \frac{1}{n}. \tag{14.5}$$

Thus whenever $\omega \notin I$, the above inequality will hold for all n large enough. Let

$$s_1 = t_1,\ s_2 = \max(t_1, t_2),\ s_3 = \max(t_1, t_2, t_3), \cdots.$$

Then the sequence $\{s_n\}$ satisfies 14.3-14.4. Also each s_n has finitely many values and is measurable. To see this, note that

$$s_n^{-1}((a,\infty]) = \cup_{k=1}^n t_k^{-1}((a,\infty]) \in \mathscr{F}$$

To verify the last claim, note that in this case the term $2^n \mathscr{X}_I(\omega)$ is not present and for n large enough, $2^n/n$ is larger than all values of f. Therefore, for all n large enough, 14.5 holds for all ω. Thus the convergence is uniform.

Now consider the converse assertion. Why is f measurable if it is the pointwise limit of an increasing sequence simple functions?

$$f^{-1}((a,\infty]) = \cup_{n=1}^{\infty} s_n^{-1}((a,\infty])$$

because $\omega \in f^{-1}((a,\infty])$ if and only if $\omega \in s_n^{-1}((a,\infty])$ for all n sufficiently large. ■

Observation 14.5.7 *If $f : \Omega \to \mathbb{R}$ then the above definition of measurability holds with no change. In this case, f never achieves the value ∞. This is actually the case of most interest.*

Corollary 14.5.8 *If $f : \Omega \to (-\infty, \infty)$ is measurable, then there exists a sequence of simple functions $\{s_n(\omega)\}$ such that $|s_n(\omega)| \leq |f(\omega)|$ and $s_n(\omega) \to f(\omega)$.*

Proof: Let $f_+(\omega) \equiv \frac{|f(\omega)|+f(\omega)}{2}, f_-(\omega) \equiv \frac{|f(\omega)|-f(\omega)}{2}$. Thus $f = f_+ - f_-$ and $|f| = f_+ + f_-$. Also $f = f_+$ when $f \geq 0$ and $f = -f_-$ when $f \leq 0$. Both f_+, f_- are measurable functions. Indeed, if $a \geq 0$, $f_+^{-1}((a,\infty)) = f^{-1}((a,\infty)) \in \mathscr{F}$. If $a < 0$ then $f_+^{-1}((a,\infty)) = \Omega$. Similar considerations hold for f_-. Now let $s_n^+(\omega) \uparrow f_+(\omega), s_n^-(\omega) \uparrow f_-(\omega)$ meaning these are simple functions converging respectively to f_+ and f_- which are both increasing in n and nonnegative. Thus if $s_n(\omega) \equiv s_n^+(\omega) - s_n^-(\omega)$, this converges to $f_+(\omega) - f_-(\omega)$. Also

$$|s_n(\omega)| = s_n^+(\omega) + s_n^-(\omega) \leq f_+(\omega) + f_-(\omega) = |f(\omega)| \blacksquare$$

Definition 14.5.9 *$\mathbb{R}^p$ consists of the mappings from $(1,2,\cdots,p)$ to $\mathbb{R}$. We usually write it as follows. $\mathbf{x} \in \mathbb{R}^p$ means $\mathbf{x}$ is an ordered list of p real numbers. Thus*

$$\mathbf{x} = (x_1, \cdots, x_p)$$

$(1,2,3)$ is in $\mathbb{R}^3$. Note that $(1,2,3) \neq (3,2,1)$ because the two have different real numbers in some locations. In terms of functions, $x(1) \neq x(3)$.

Definition 14.5.10 *If $\mathbf{x}_n = \left(x_1^n, \cdots, x_p^n\right)$ is a sequence of points in $\mathbb{R}^p$, then we say $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$ if and only if $\lim_{n\to\infty} x_i^n = x_i$ for each i. In other words, convergence takes place if and only if the component entries of $\mathbf{x}_n$ converge to the corresponding component entries of $\mathbf{x}$. We say that $g : D \to \mathbb{R}$ is continuous for $D \subseteq \mathbb{R}^p$ if whenever $\mathbf{x}_n \to \mathbf{x}$ with each $\mathbf{x}_n \in D$ and $\mathbf{x} \in D$, then $g(\mathbf{x}_n) \to g(\mathbf{x})$.*

In other words, it is essentially the same as what was presented earlier for continuous functions of one variable.

Proposition 14.5.11 *Let $f_i : \Omega \to \mathbb{R}$ be measurable, $(\Omega, \mathscr{F})$ a measurable space, and let $g : \mathbb{R}^p \to \mathbb{R}$ be continuous. If $\mathbf{f}(\omega) = \begin{pmatrix} f_1(\omega) & \cdots & f_p(\omega) \end{pmatrix}^T$, then $g \circ \mathbf{f}$ is measurable.*

Proof: From Corollary 14.5.8 above, there are $s_i^n(\omega)$, simple functions

$$\lim_{n\to\infty} s_i^n(\omega) = f_i(\omega)$$

such that $|s_i^n(\omega)| \leq |f_i(\omega)|$. Let $\mathbf{s}_n(\omega) \equiv \begin{pmatrix} s_1^n(\omega) & \cdots & s_p^n(\omega) \end{pmatrix}$ thus, by continuity, $g(\mathbf{s}_n(\omega)) \to g(\mathbf{f}(\omega))$ for each ω. It remains to verify that $g \circ \mathbf{s}_n$ is measurable.

$$(g \circ \mathbf{s}_n)^{-1}(a,\infty) \equiv \{\omega : g(\mathbf{s}_n(\omega)) > a\}.$$

This is the finite union of measurable sets since each s_i^n is a simple function having finitely many values. Thus there are finitely many possible values for $g \circ \mathbf{s}_n$, each value corresponding to the intersection of p measurable sets. Therefore, $g \circ \mathbf{s}_n$ is measurable. By Corollary 14.5.4, it follows that $g \circ \mathbf{f}$, being the pointwise limit of measurable functions is also measurable. $\blacksquare$

Note how this shows as a very special case that linear combinations of measurable real valued functions are measurable because you could take $g(x,y) \equiv ax + by$ and then if you have two measurable functions f_1, f_2, it follows that $af_1 + bf_2, f_1 f_2$ are measurable. Also, if f is measurable, then so is $|f|$. Just let $g(x) = |x|$. In general, you can do pretty much any algebraic combination of measurable functions and get one which is measurable. This is very different than the case of generalized integrable functions.

14.6 Riemann Integrals for Decreasing Functions

This continues the abstract development but here it is tied in to the ordinary theory of Riemann integration for real valued functions. A decreasing function is always Riemann integrable with respect to the integrator function $F(t) = t$. This is because the function is of bounded variation and the integrator function is continuous. You can also show directly that there is a unique number between the upper and lower sums. I will define the Lebesgue integral for a nonnegative function in terms of an improper Riemann integral which involves a decreasing function.

Definition 14.6.1 *Let $f : [a,b] \to [0,\infty]$ be decreasing. Define*

$$\int_a^b f(\lambda)\, d\lambda \equiv \lim_{M\to\infty} \int_a^b M \wedge f(\lambda)\, d\lambda = \sup_M \int_a^b M \wedge f(\lambda)\, d\lambda$$

where $A \wedge B$ means the minimum of A and B. Note that for f bounded,

$$\sup_M \int_a^b M \wedge f(\lambda)\, d\lambda = \int_a^b f(\lambda)\, d\lambda$$

where the integral on the right is the usual Riemann integral because eventually $M > f$. For f a nonnegative decreasing function defined on $[0,\infty)$,

$$\int_0^\infty f d\lambda \equiv \lim_{R\to\infty} \int_0^R f d\lambda = \sup_{R>1} \int_0^R f d\lambda = \sup_R \sup_{M>0} \int_0^R f \wedge M d\lambda$$

Now here is an obvious property.

Lemma 14.6.2 *Let f be a decreasing nonnegative function defined on an interval $[a,b]$. Then if $[a,b] = \cup_{k=1}^m I_k$ where $I_k \equiv [a_k, b_k]$ and the intervals I_k are non overlapping, it follows*

$$\int_a^b f d\lambda = \sum_{k=1}^m \int_{a_k}^{b_k} f d\lambda.$$

Proof: This follows from the computation,

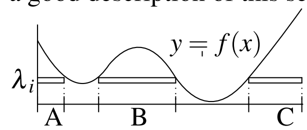
$$\int_a^b f d\lambda \equiv \lim_{M\to\infty} \int_a^b f \wedge M d\lambda = \lim_{M\to\infty} \sum_{k=1}^m \int_{a_k}^{b_k} f \wedge M d\lambda = \sum_{k=1}^m \int_{a_k}^{b_k} f d\lambda$$

Note both sides could equal $+\infty$. ■

14.7 Lebesgue Integrals of Nonnegative Functions

Here is the definition of the Lebesgue integral of a function which is measurable and has values in $[0,\infty]$. The idea is motivated by the following picture in which $f^{-1}(\lambda_i, \infty)$ is $A \cup B \cup C$ and we take the measure of this set, multiply by $\lambda_i - \lambda_{i-1}$ and do this for each λ_i in an increasing sequence of points, $\lambda_0 \equiv 0$. Then we add the "areas" of the little horizontal "rectangles" in order to approximate the "area" under the curve. The difference here is that the "rectangles" in the sum are horizontal whereas with the Riemann integral, they are vertical. Note how it is important to be able to measure $f^{-1}(\lambda, \infty) \equiv \{x : f(x) > \lambda\} \equiv$

$[f > \lambda]$ which is what it means for f to be measurable. Also note that, in spite of the picture, in general we don't know a good description of this set other than that it is measurable.



Definition 14.7.1 *Let $(\Omega, \mathscr{F}, \mu)$ be a measure space and suppose $f : \Omega \to [0, \infty]$ is measurable. Then define*

$$\int f d\mu \equiv \int_0^\infty \mu([f > \lambda]) d\lambda = \int_0^\infty \mu\left(f^{-1}(\lambda, \infty)\right) d\lambda$$

which makes sense because $\lambda \to \mu([f > \lambda])$ is nonnegative and decreasing. On the right you have an improper Riemann integral like what was discussed above.

Note that if $f \leq g$, then $\int f d\mu \leq \int g d\mu$ because $\mu([f > \lambda]) \leq \mu([g > \lambda])$. Next I point out that the integral is a limit of lower sums.

Lemma 14.7.2 *In the situation of the above definition,*

$$\int f d\mu = \sup_{h>0} \sum_{i=1}^{\infty} \mu([f > hi]) h$$

Proof: Let $m(h, R) \in \mathbb{N}$ satisfy $R - h < hm(h, R) \leq R$. Then $\lim_{R \to \infty} hm(h, R) = \infty$ and so

$$\int f d\mu \equiv \int_0^\infty \mu([f > \lambda]) d\lambda \equiv \sup_M \sup_R \int_0^R \mu([f > \lambda]) \wedge M d\lambda$$

$$\sup_M \sup_R \int_0^{hm(h,R)} \mu([f > \lambda]) \wedge M d\lambda = \sup_M \sup_{R>0} \sup_{h>0} \sum_{k=1}^{m(h,R)} (\mu([f > kh]) \wedge M) h$$

because the sum is just a lower sum for the integral $\int_0^{hm(h,R)} \mu([f > \lambda]) \wedge M d\lambda$. Hence, switching the order of the sups, this equals

$$\sup_{R>0} \sup_{h>0} \sup_M \sum_{k=1}^{m(h,R)} (\mu([f > kh]) \wedge M) h = \sup_{R>0} \sup_{h>0} \lim_{M \to \infty} \sum_{k=1}^{m(h,R)} (\mu([f > kh]) \wedge M) h$$

$$= \sup_{h>0} \sup_R \sum_{k=1}^{m(R,h)} \mu([f > kh]) h = \sup_{h>0} \sum_{k=1}^{\infty} \mu([f > kh]) h. \blacksquare$$

14.8 Nonnegative Simple Functions

To begin with, here is a useful lemma.

Lemma 14.8.1 *If $f(\lambda) = 0$ for all $\lambda > a$, where f is a decreasing nonnegative function, then $\int_0^\infty f(\lambda) d\lambda = \int_0^a f(\lambda) d\lambda$.*

Proof: From the definition,

$$\begin{aligned}\int_0^{\infty} f(\lambda)\,d\lambda &= \lim_{R\to\infty}\int_0^R f(\lambda)\,d\lambda = \sup_{R>1}\int_0^R f(\lambda)\,d\lambda = \sup_{R>1}\sup_M\int_0^R f(\lambda)\wedge M d\lambda \\ &= \sup_M\sup_{R>1}\int_0^R f(\lambda)\wedge M d\lambda = \sup_M\sup_{R>1}\int_0^a f(\lambda)\wedge M d\lambda \\ &= \sup_M\int_0^a f(\lambda)\wedge M d\lambda \equiv \int_0^a f(\lambda)\,d\lambda. \blacksquare\end{aligned}$$

Now the Lebesgue integral for a nonnegative function has been defined, what does it do to a nonnegative simple function? Recall a nonnegative simple function is one which has finitely many nonnegative real values which it assumes on measurable sets. Thus a simple function can be written in the form $s(\omega) = \sum_{i=1}^n c_i \mathscr{X}_{E_i}(\omega)$ where the c_i are each nonnegative, the distinct values of s.

Lemma 14.8.2 *Let $s(\omega) = \sum_{i=1}^p a_i \mathscr{X}_{E_i}(\omega)$ be a nonnegative simple function where the E_i are distinct but the a_i might not be. Then*

$$\int s\,d\mu = \sum_{i=1}^p a_i \mu(E_i). \tag{14.6}$$

Proof: Without loss of generality, assume $0 \equiv a_0 < a_1 \le a_2 \le \cdots \le a_p$ and that $\mu(E_i) < \infty, i > 0$. Here is why. If $\mu(E_i) = \infty$, then the left side would be

$$\begin{aligned}\int_0^{a_p} \mu([s>\lambda])\,d\lambda &\ge \int_0^{a_i} \mu([s>\lambda])\,d\lambda \\ &= \sup_M \int_0^{a_i} \mu([s>\lambda])\wedge M d\lambda \ge \sup_M M a_i = \infty\end{aligned}$$

and so both sides are equal to ∞. Thus it can be assumed that for each $i, \mu(E_i) < \infty$. Then it follows from Lemma 14.8.1 and Lemma 14.6.2,

$$\int_0^{\infty} \mu([s>\lambda])\,d\lambda = \int_0^{a_p} \mu([s>\lambda])\,d\lambda = \sum_{k=1}^p \int_{a_{k-1}}^{a_k} \mu([s>\lambda])\,d\lambda$$

$$= \sum_{k=1}^p (a_k - a_{k-1}) \sum_{i=k}^p \mu(E_i) = \sum_{i=1}^p \mu(E_i) \sum_{k=1}^i (a_k - a_{k-1}) = \sum_{i=1}^p a_i \mu(E_i) \ \blacksquare$$

Lemma 14.8.3 *If $a, b \ge 0$ and if s and t are nonnegative simple functions, then*

$$\int (as+bt)\,d\mu = a\int s\,d\mu + b\int t\,d\mu.$$

Proof: Let $s(\omega) = \sum_{i=1}^n \alpha_i \mathscr{X}_{A_i}(\omega)$, $t(\omega) = \sum_{i=1}^m \beta_j \mathscr{X}_{B_j}(\omega)$ 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 because it has finitely many values on measurable sets. In fact,

$$(as+bt)(\omega) = \sum_{j=1}^m \sum_{i=1}^n (a\alpha_i + b\beta_j) \mathscr{X}_{A_i \cap B_j}(\omega)$$

where the sets $A_i \cap B_j$ are disjoint and measurable. By Lemma 14.8.2,

$$\begin{aligned}\int as+bt d\mu &= \sum_{j=1}^{m}\sum_{i=1}^{n}(a\alpha_i+b\beta_j)\mu(A_i\cap B_j)\\ &= \sum_{i=1}^{n}a\sum_{j=1}^{m}\alpha_i\mu(A_i\cap B_j)+b\sum_{j=1}^{m}\sum_{i=1}^{n}\beta_j\mu(A_i\cap B_j)\\ &= a\sum_{i=1}^{n}\alpha_i\mu(A_i)+b\sum_{j=1}^{m}\beta_j\mu(B_j)=a\int s d\mu+b\int t d\mu.\ \blacksquare\end{aligned}$$

14.9 The Monotone Convergence Theorem

The following is called the monotone convergence theorem also Beppo Levi's theorem. This theorem and related convergence theorems are the reason for using the Lebesgue integral. If $\lim_{n\to\infty} f_n(\omega) = f(\omega)$ and $f_n(\omega)$ is increasing in n, then clearly f is also measurable because

$$f^{-1}((a,\infty]) = \cup_{k=1}^{\infty} f_k^{-1}((a,\infty]) \in \mathscr{F}$$

The version of this theorem given here will be much simpler than what was done with the generalized Riemann integral and it will be easier to state and remember and use.

Theorem 14.9.1 *(Monotone Convergence theorem) Let f have values in $[0,\infty]$ and suppose $\{f_n\}$ is a sequence of nonnegative measurable functions having values in $[0,\infty]$ and satisfying*

$$\lim_{n\to\infty} f_n(\omega) = f(\omega) \text{ for each } \omega.$$

$$\cdots f_n(\omega) \le f_{n+1}(\omega)\cdots$$

Then f is measurable and

$$\int f d\mu = \lim_{n\to\infty}\int f_n d\mu.$$

Proof: By Lemma 14.7.2 and interchange of supremums,

$$\lim_{n\to\infty}\int f_n d\mu = \sup_n \int f_n d\mu$$

$$\begin{aligned}&= \sup_n \sup_{h>0}\sum_{k=1}^{\infty}\mu([f_n>kh])h = \sup_{h>0}\sup_N\sup_n\sum_{k=1}^{N}\mu([f_n>kh])h\\ &= \sup_{h>0}\sup_N\sum_{k=1}^{N}\mu([f>kh])h = \sup_{h>0}\sum_{k=1}^{\infty}\mu([f>kh])h = \int f d\mu.\ \blacksquare\end{aligned}$$

The next theorem, known as Fatou's lemma is another important theorem which justifies the use of the Lebesgue integral.

Theorem 14.9.2 *(Fatou's lemma) Let f_n be a nonnegative measurable function. Let $g(\omega) = \liminf_{n\to\infty} f_n(\omega)$. Then g is measurable and*

$$\int g d\mu \le \lim\inf_{n\to\infty}\int f_n d\mu.$$

In other words,

$$\int\left(\lim\inf_{n\to\infty} f_n\right)d\mu \le \lim\inf_{n\to\infty}\int f_n d\mu$$

Proof: Let $g_n(\omega) = \inf\{f_k(\omega) : k \geq n\}$. Then $g_n^{-1}([a,\infty]) = \cap_{k=n}^{\infty} f_k^{-1}([a,\infty]) \in \mathscr{F}$ Thus g_n is measurable. Now the functions g_n form an increasing sequence of nonnegative measurable functions. Thus $g^{-1}((a,\infty)) = \cup_{n=1}^{\infty} g_n^{-1}((a,\infty)) \in \mathscr{F}$ so g is measurable also. By monotone convergence theorem, $\int g d\mu = \lim_{n\to\infty} \int g_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu$. The last inequality holding because $\int g_n d\mu \leq \int f_n d\mu$. (Note that it is not known whether $\lim_{n\to\infty} \int f_n d\mu$ exists.) ■

14.10 The Integral's Righteous Algebraic Desires

The monotone convergence theorem shows the integral wants to be linear. This is the essential content of the next theorem.

Theorem 14.10.1 *Let f,g be nonnegative measurable functions and let a,b be nonnegative numbers. Then $af+bg$ is measurable and*

$$\int (af+bg)\, d\mu = a\int f d\mu + b\int g d\mu. \tag{14.7}$$

Proof: By Theorem 14.5.6 on Page 334 there exist increasing sequences of nonnegative simple functions, $s_n \to f$ and $t_n \to g$. Then $af+bg$, being the pointwise limit of the simple functions $as_n + bt_n$, is measurable. Now by the monotone convergence theorem and Lemma 14.8.3, $\int (af+bg)\, d\mu =$

$$\lim_{n\to\infty} \int as_n + bt_n d\mu = \lim_{n\to\infty} \left(a\int s_n d\mu + b\int t_n d\mu \right) = a\int f d\mu + b\int g d\mu. \ ■$$

14.11 Integrals of Real Valued Functions

As long as you are allowing functions to take the value $+\infty$, you cannot consider something like $f + (-g)$ and so you can't very well expect a satisfactory statement about the integral being linear until you restrict yourself to functions which have values in a vector space. To be linear, a function must be defined on a vector space. The integral of real valued functions is next.

Definition 14.11.1 *Let $(\Omega, \mathscr{F}, \mu)$ be a measure space and let $f : \Omega \to \mathbb{R}$ be measurable. Then it is said to be in $L^1(\Omega,\mu)$ when $\int_\Omega |f(\omega)|\, d\mu < \infty$.*

Lemma 14.11.2 *If $g - h = \hat{g} - \hat{h}$ where $g, \hat{g}, h, \hat{h}$ are measurable and nonnegative, with all integrals finite, then*

$$\int_\Omega g d\mu - \int_\Omega h d\mu = \int_\Omega \hat{g} d\mu - \int_\Omega \hat{h} d\mu$$

Proof: From Theorem 14.10.1,

$$\int \hat{g} d\mu + \int h d\mu = \int (\hat{g} + h)\, d\mu = \int \left(g + \hat{h}\right) d\mu = \int g d\mu + \int \hat{h} d\mu$$

and so,

$$\int \hat{g} d\mu - \int \hat{h} d\mu = \int g d\mu - \int h d\mu \ ■$$

The functions you can integrate are those which have $|f|$ integrable, and then you can make sense of $\int f d\mu$ for f having values in $\mathbb{R}$ or $\mathbb{C}$ although here, I will emphasize $\mathbb{R}$.

Definition 14.11.3 *Let $f \in L^1(\Omega, \mu)$. Define $\int f d\mu \equiv \int f_+ d\mu - \int f_- d\mu$.*

Proposition 14.11.4 *The definition of $\int f d\mu$ is well defined and if a, b are real numbers*

$$\int (af + bg)\, d\mu = a \int f d\mu + b \int g d\mu$$

Proof: First of all, it is well defined because f_+, f_- are both no larger than $|f|$. Therefore, $\int f_+ d\mu, \int f_- d\mu$ are both real numbers. Next, why is the integral linear. First consider the sum.

$$\int (f+g)\, d\mu \equiv \int (f+g)_+\, d\mu - \int (f+g)_-\, d\mu$$

Now $(f+g)_+ - (f+g)_- = f + g = f_+ - f_- + g_+ - g_-$. By Lemma 14.11.2 and Theorem 14.10.1

$$\int (f+g)\, d\mu \equiv \int (f+g)_+\, d\mu - \int (f+g)_-\, d\mu = \int (f_+ + g_+)\, d\mu - \int (f_- + g_-)\, d\mu$$

$$= \int f_+ d\mu - \int f_- d\mu + \int g_+ d\mu - \int g_- d\mu \equiv \int f d\mu + \int g d\mu$$

Next note that if a is real and $a \geq 0, (af)_+ = af_+, (af)_- = af_-$ and if $a < 0, (af)_+ = -af_-, (af)_- = -af_+$. This follows from a simple computation involving the definition of f_+, f_-. Therefore, if $a < 0$,

$$\int af d\mu \equiv \int (af)_+\, d\mu - \int (af)_-\, d\mu = \int (-a)\, f_- d\mu - \int (-a)\, f_+ d\mu$$

By Theorem 14.10.1,

$$= -a \left(\int f_- d\mu - \int f_+ d\mu \right) = a \left(\int f_+ d\mu - \int f_- d\mu \right) \equiv a \int f d\mu$$

The case where $a \geq 0$ is easier. ■

Note how attractive this is. If you have a measurable function f and it is absolutely integrable, then it is integrable. This is just like the situation with series.

Now that we understand how to integrate real valued functions, it is time for another great convergence theorem, the dominated convergence theorem.

Theorem 14.11.5 *(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 $[0, \infty]$,[1] such that*

$$|f_n(\omega)| \leq g(\omega) \text{ and } \int g(\omega) d\mu < \infty.$$

Then $f \in L^1(\Omega)$ and $0 = \lim_{n\to\infty} \int |f_n - f|\, d\mu = \lim_{n\to\infty} \left| \int f d\mu - \int f_n d\mu \right|$

Proof: f is measurable by Corollary 14.5.4. Since $|f| \leq g$, it follows that

$$f \in L^1(\Omega) \text{ and } |f - f_n| \leq 2g.$$

[1] Note that, since g is allowed to have the value ∞, it is not known that $g \in L^1(\Omega)$.

By Fatou's lemma (Theorem 14.9.2),

$$\int 2g d\mu \leq \liminf_{n\to\infty} \int 2g - |f - f_n| d\mu = \int 2g d\mu - \limsup_{n\to\infty} \int |f - f_n| d\mu.$$

Subtracting $\int 2g d\mu$, $0 \leq -\limsup_{n\to\infty} \int |f - f_n| d\mu$. Hence

$$\begin{aligned} 0 &\geq \limsup_{n\to\infty} \left(\int |f - f_n| d\mu \right) \geq \liminf_{n\to\infty} \left(\int |f - f_n| d\mu \right) \\ &\geq \liminf_{n\to\infty} \left| \int f d\mu - \int f_n d\mu \right| \geq 0. \end{aligned}$$

This proves the theorem by Theorem 4.10.10 because the lim sup and lim inf are equal. ■

Example 14.11.6 *Let $\Omega \equiv \mathbb{N}$ and let $\mathscr{F}$ be the set of all subsets of Ω. Let $\mu(E) \equiv$ number of entries in E. Then $(\mathbb{N}, \mathscr{F}, \mu)$ is a measure space and the Lebesgue integral is summation. Thus all the convergence theorems mentioned above apply to sums.*

First, why is μ a measure? If $\{E_i\}$ are disjoint, then if infinitely many are nonempty, say $E_{n_1}, E_{n_2}, \cdots$. Then $\cup_i E_i$ is infinite and so

$$\mu(\cup_i E_i) = \infty = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{k=1}^{\infty} E_{n_k} \geq \sum_{k=1}^{\infty} 1 = \infty$$

The alternative is that only finitely many E_i are nonempty and in this case, the assertion that $\mu(\cup_i E_i) = \sum_{i=1}^{\infty} \mu(E_i)$ is obvious. Hence μ is indeed a measure. Now let $f : \mathbb{N} \to \mathbb{R}$. It is obviously measurable because the inverse image of anything is a subset of $\mathbb{N}$. So if $f(n) \geq 0$ for all n, what is $\int f d\mu$?

$$f(i) = \sum_{k=1}^{\infty} f(k) \mathscr{X}_{\{k\}}(i) = \lim_{n\to\infty} \sum_{k=1}^{n} f(k) \mathscr{X}_{\{k\}}(i) \equiv f_n(i)$$

Now f_n is a nonnegative simple function and there is exactly one thing in $\{k\}$. Therefore, $\int f_n d\mu = \sum_{k=1}^{n} f(k)$. Then, by the monotone convergence theorem,

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu = \lim_{n\to\infty} \sum_{k=1}^{n} f(k) \equiv \sum_{k=1}^{\infty} f(k)$$

When $\sum_k |f(k)| < \infty$, one has $\int f d\mu = \sum_{k=1}^{\infty} f(k)$.

This example illustrates how the Lebesgue integral pertains to absolute summability and absolute integrability. It is not a theory which can include conditional convergence. The generalized Riemann integral can do this. However, the Lebesgue integral is very easy to use because of this restriction. Of course one could make the same restriction and consider those functions for which $|f|$ is integrable in the context of the generalized Riemann integral.

How does this new integral compare to the generalized Riemann Stieltjes integral? Let μ_F be the measure of Theorem 14.3.6 in what follows.

Theorem 14.11.7 *Let F be an increasing integrator function and let f be continuous on $[a,b]$. Then*

$$\int_a^b f dF = \int f d\mu_F$$

Here the integral on the left is the Riemann Stieltjes integral.

Proof: Since f is continuous, there is a sequence of step functions $\{s_n\}$ which converges uniformly to f. In fact, you could partition $[a,b]$ as $a = x_0^n < \cdots < x_n^n = b$ and consider

$$s_n(x) = \sum_{k=1}^{m_n} f(x_{k-1}) \mathscr{X}_{[x_{k-1},x_k)}(x), \ |x_k - x_{k-1}| = \frac{b-a}{n}$$

Then if you consider the generalized Riemann integral of $s_n(x)$ it equals

$$\int_a^b s_n(x)\, dF = \sum_{k=1}^{n} f(x_{k-1})(F(x_k-) - F(x_{k-1}-)) = \int s_n d\mu_F$$

The integrals on the left converge as $n \to \infty$ to $\int_a^b f dF$ where this is the generalized Riemann integral of f because of the uniform convergence of the step functions to f. But this equals the Riemann Stieltjes integral since f is continuous. On the right, you get convergence to $\int f d\mu_F$ again because of uniform convergence. ■

This shows that the new integral coincides with the generalized Riemann Stieltjes integral on all continuous functions continuous on a closed interval which is the same as the Riemann Stieltjes integral. They give the same answer on continuous functions. The advantage of the Lebesgue integral is that it has all those wonderful convergence theorems which are particularly easy to understand and use. You don't have these theorems with the Riemann Stieltjes integral and the version of these convergence theorems which hold for the generalized Riemann integral are much more difficult to prove. Compare the easy monotone convergence theorem with the much more difficult one in the Chapter on the generalized Riemann integral.

Also, there is the following definition of the integral over a measurable subset.

Definition 14.11.8 *Let $(\Omega, \mathscr{F}, \mu)$ be a measure space and let E be a measurable subset of Ω. Then if $|f\mathscr{X}_E|$ is integrable, $\int_E f d\mu \equiv \int \mathscr{X}_E f d\mu$.*

When we are considering the Lebesgue Stieltjes measures, we might write the following for f Borel measurable:

$$\int_a^b f d\mu_F \equiv \int \mathscr{X}_{[a,b]} f d\mu_F$$

Of course the notation on the right is preferable because you might have $\mu_F(a) > 0$. The notation on the left might seem a little ambiguous about whether the end points are considered. Of course the main example is when $F(t) = t$ and in this case, there is no ambiguity. As pointed out, this gives the Riemann Stieltjes integral provided f is continuous on $[a,b]$ and vanishes near a and b. One must remember that single points can have positive measure in the Stieltjes case where you just have an increasing integrator function which could have jumps.

14.12 The Vitali Covering Theorems

These theorems are remarkable and fantastically useful. I will use m for μ_F where $F(t) = t$. This yields Lebesgue measure which gives the measure of an interval to be the length of the interval. As earlier, $B(x,r)$ will be the interval $(x-r, x+r)$. Thus $m(B(x,\alpha r)) = \alpha m(x,r)$ whenever $\alpha > 0$. Also note that the closure of $B(x,r)$ is the closed interval $[x-r, x+r]$. This is because the closure simply adds in the limit points which are not in the open interval and these are the end points. This all extends to many dimensions and so I am using the notation which is appropriate for this generalization. The main ideas are all present in one dimension. A ball will just be an interval. It might have no endpoints, one endpoint, or both endpoints.

Lemma 14.12.1 *Let $\mathscr{F}$ be a countable collection of balls satisfying*

$$\infty > M \equiv \sup\{r : B(p,r) \in \mathscr{F}\} > 0$$

and let $k \in (0,\infty)$. Then there exists $\mathscr{G} \subseteq \mathscr{F}$ such that

$$\text{If } B(p,r) \in \mathscr{G} \text{ then } r > k, \tag{14.8}$$

$$\text{If } B_1, B_2 \in \mathscr{G} \text{ then } \overline{B_1} \cap \overline{B_2} = \emptyset, \tag{14.9}$$

$$\mathscr{G} \text{ is maximal with respect to 14.8 and 14.9.} \tag{14.10}$$

By this is meant that if $\mathscr{H}$ is a collection of balls satisfying 14.8 and 14.9, then $\mathscr{H}$ cannot properly contain $\mathscr{G}$.

Proof: If no ball of $\mathscr{F}$ has radius larger than k, let $\mathscr{G} = \emptyset$. Assume therefore, that some balls have radius larger than k. Let $\mathscr{F} \equiv \{B_i\}_{i=1}^{\infty}, B_i \cap B_j = \emptyset$ if $i \neq j$. Now let B_{n_1} be the first ball in the list which has radius greater than k. If every ball having radius larger than k has closure which intersects $\overline{B_{n_1}}$, then stop. The maximal set is $\{B_{n_1}\}$. Otherwise, let B_{n_2} be the next ball having radius larger than k for which $\overline{B_{n_2}} \cap \overline{B_{n_1}} = \emptyset$. Continue this way obtaining $\{B_{n_i}\}_{i=1}^{N}$, a finite or infinite sequence of balls having radius larger than k whose closures are disjoint. $N = \infty$ if the process never stops. N is some number if the process does stop. Then let $\mathscr{G} \equiv \{B_{n_i}\}_{i=1}^{N}$. To see $\mathscr{G}$ is maximal with respect to 14.8 and 14.9, suppose $B \in \mathscr{F}$, B has radius larger than k, and $\mathscr{G} \cup \{B\}$ satisfies 14.8 and 14.9. Then at some point in the process, B would have been chosen because it would be the ball of radius larger than k which has the smallest index at some point in the construction. Therefore, $B \in \mathscr{G}$ and this shows $\mathscr{G}$ is maximal with respect to 14.8 and 14.9. ■

Lemma 14.12.2 *Let $\mathscr{F}$ be a collection of open balls, and let*

$$A \subseteq \cup\{B : B \in \mathscr{F}\}.$$

Suppose $\tilde{B}$ denotes the closed ball with the same center as B but four times the radius. Let $\hat{B}$ denote the open ball with same center as B but five times the radius.

$$\infty > M \equiv \sup\{r : B(p,r) \in \mathscr{F}\} > 0.$$

Then there exists $\mathscr{G} \subseteq \mathscr{F}$ such that $\mathscr{G}$ consists of balls whose closures are disjoint and

$$A \subseteq \cup\left\{\tilde{B} : B \in \mathscr{G}\right\} \subseteq \cup\{\widehat{B} : B \in \mathscr{G}\}$$

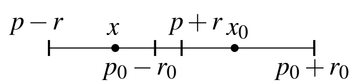
Proof: First of all, it follows from Theorem 4.8.20 on Page 73 that there is a countable subset $\tilde{\mathscr{F}}$ of $\mathscr{F}$ which also covers A. Thus $\cup\tilde{\mathscr{F}} \supseteq A \supseteq \cup\mathscr{F} \supseteq \cup\tilde{\mathscr{F}}$ and so $\cup\tilde{\mathscr{F}} = A$. Thus it can be assumed that $\mathscr{F}$ is countable.

By Lemma 14.12.1, there exists $\mathscr{G}_1 \subseteq \mathscr{F}$ which satisfies 14.8, 14.9, and 14.10 with $k = \frac{2M}{3}$. That is, $\mathscr{G}_1$ consists of balls having radii larger than $\frac{2M}{3}$ and their closures are disjoint and $\mathscr{G}_1$ is as large as possible.

Suppose $\mathscr{G}_1, \cdots, \mathscr{G}_{m-1}$ have been chosen for $m \geq 2$. Let $\overline{\mathscr{G}_i}$ denote the collection of closures of the balls of $\mathscr{G}_i$. Then let $\mathscr{F}_m$ be those balls of $\mathscr{F}$, such that if B is one of these balls, $\overline{B}$ has empty intersection with every closed ball of $\overline{\mathscr{G}_i}$ for each $i \leq m-1$. Then using Lemma 14.12.1, let $\mathscr{G}_m$ be a maximal collection of balls from $\mathscr{F}_m$ with the property that each ball has radius larger than $\left(\frac{2}{3}\right)^m M$ and their closures are disjoint. Let $\mathscr{G} \equiv \cup_{k=1}^{\infty} \mathscr{G}_k$. Thus the closures of balls in $\mathscr{G}$ are disjoint. Let $x \in B(p,r) \in \mathscr{F} \setminus \mathscr{G}$. Choose m such that

$$\left(\frac{2}{3}\right)^m M < r \leq \left(\frac{2}{3}\right)^{m-1} M$$

Then $\overline{B(p,r)}$ must have nonempty intersection with the closure of some ball from $\mathscr{G}_1 \cup \cdots \cup \mathscr{G}_m$ because if it didn't, then $\mathscr{G}_m$ would fail to be maximal. Denote by $B(p_0, r_0)$ a ball in $\mathscr{G}_1 \cup \cdots \cup \mathscr{G}_m$ whose closure has nonempty intersection with $\overline{B(p,r)}$. Thus $r_0, r > \left(\frac{2}{3}\right)^m M$. Consider the picture, in which $w \in \overline{B(p_0, r_0)} \cap \overline{B(p,r)}$.



Then for $x \in \overline{B(p,r)}$, $|x - p_0| \leq |x-p| + |p-w| + \overbrace{|w-p_0|}^{\leq r_0}$

$$\leq r + r + r_0 \leq 2\left(\frac{2}{3}\right)^{m-1} M + r_0 \leq 2\left(\frac{3}{2}\right)\overbrace{\left(\frac{2}{3}\right)^m M}^{<r_0} + r_0 \leq 4r_0$$

Thus $B(p,r)$ is contained in $\overline{B(p_0, 4r_0)}$. This shows that $A \subseteq \left\{\tilde{B} : B \in \mathscr{G}\right\} \subseteq \cup\{\widehat{B} : B \in \mathscr{G}\}$. ■

You can easily generalize this proposition to include the case where the balls are not necessarily open. These balls could be either open, closed, or neither open nor closed.

Proposition 14.12.3 *Let $\mathscr{F}$ be a collection of balls, open, closed, or neither open nor closed, and let*

$$A \equiv \cup\{B : B \in \mathscr{F}\}.$$

Let $\hat{B}$ denote the open ball with same center as B but five times the radius.

$$\infty > M \equiv \sup\{r : B(p,r) \in \mathscr{F}\} > 0.$$

Then there exists $\mathscr{G} \subseteq \mathscr{F}$ such that $\mathscr{G}$ consists of balls whose closures are disjoint and

$$A \subseteq \cup\{\widehat{B} : B \in \mathscr{G}\}$$

Proof: Let $\mathscr{F}'$ consist of the open balls obtained from $\mathscr{F}$ by keeping the center the same and multiplying the radius by $\frac{20}{19}$. Denote these open balls by $\{B' : B \in \mathscr{F}\}$. Thus B' is an open ball in which the radius is slightly expanded but the center is the same as $B \in \mathscr{F}$. Then from Lemma 14.12.2, there is $\mathscr{G}' \subseteq \mathscr{F}'$ such that the closures of the balls in $\mathscr{G}'$ are disjoint and $\left\{\tilde{B}' : B' \in \mathscr{G}'\right\}$ covers A. Here $\tilde{B}'$ is the closed ball obtained by multiplying the radius of B' by four. If $B = B(x,r)$, then $B' = B\left(x, \frac{20}{19}r\right)$ and so $\tilde{B}' = \overline{B\left(x, \frac{80}{19}r\right)} \subseteq B(x, 5r)$ since $5 > \frac{80}{19}$. Thus, letting $\mathscr{G}$ be the balls B such that $B' \in \mathscr{G}'$, it follows the closures of these balls in $\mathscr{G}$ are disjoint since these balls are smaller than the balls of $\mathscr{G}'$ and $A \subseteq \cup\{\widehat{B} : B \in \mathscr{G}\}$ ■

The notion of an outer measure allows the consideration of arbitrary sets, measurable or not.

Definition 14.12.4 *Let E be any set in $\mathbb{R}$. Then*

$$\bar{m}(E) \equiv \inf\{m(F) : F \text{ is a Borel set and } F \supseteq E\}.$$

Proposition 14.12.5 *The following hold.*

1. $\bar{m}(F) \leq \bar{m}(G)$ *if* $F \subseteq G$
2. *If E is Borel, then* $\bar{m}(E) = m(E)$
3. $\bar{m}\left(\cup_{k=1}^{\infty} E_k\right) \leq \sum_k \bar{m}(E_k)$
4. *If $\bar{m}(E) = 0$, then there exists F Borel such that $F \supseteq E$ and $m(F) = 0$.*
5. *For any E, there exists F Borel such that $F \supseteq E$ and $\bar{m}(E) = m(F)$.*

Proof: The first claim is obvious. Consider the second. By definition, $\bar{m}(E) \leq m(E)$ because $E \supseteq E$. If $F \supseteq E$ and F is Borel, then $m(E) \leq m(F)$ and so, taking the inf of all such F containing E, it follows that $m(E) \leq \bar{m}(E)$. Now consider the third assertion.

If any E_k has $\bar{m}(E_k) = \infty$, then there is nothing to show. Assume then that $\bar{m}(E_k) < \infty$ for all k. Then by definition, there is F_k Borel, containing E_k such that $\bar{m}(E_k) + \frac{\varepsilon}{2^k} \geq m(F_k)$. Then $\bar{m}(\cup_k E_k) \leq$

$$\bar{m}(\cup_k F_k) = m(\cup_k F_k) \leq \sum_{k=1}^{\infty} m(F_k) \leq \sum_k \bar{m}(E_k) + \frac{\varepsilon}{2^k} = \sum_k \bar{m}(E_k) + \varepsilon$$

Since ε is arbitrary, this shows 3.

Finally consider 4.,5. Of course 4. is a special case of 5. If $\bar{m}(E) = \infty$, let $F = \mathbb{R}$. Otherwise there exists $G_n \supseteq E$ such that $m(G_n) < \bar{m}(E) + 1/2^n$ and G_n is a Borel set. Then let $F_n \equiv \cap_{k=1}^{n} G_n$. It follows that $\cap_n F_n \equiv F$ is a Borel set containing E and $m(F) = \lim_{n\to\infty} m(F_n) = 0$. ■

Definition 14.12.6 *Let $|f| \mathscr{X}_B \in L^1(\mathbb{R}, m)$ for every bounded Borel B. The Hardy Littlewood maximal function $Mf(x)$ is defined as*

$$Mf(x) \equiv \sup_{1\geq r>0} \frac{1}{m(B(0,r))} \int_{B(x,r)} |f|\, dm$$

You can try and show this function is measurable, but I will not do so here. The fundamentally important inequality involving this maximal function is the following major result which comes from the above Vitali covering theorem and is one of the main applications of the above Vitali covering theorem.

Theorem 14.12.7 *Let $|f| \in L^1(\mathbb{R}, m)$. Then if $\lambda > 0$,*

$$\bar{m}([Mf > \lambda]) \leq \frac{5}{\lambda} \int |f|\, dm$$

Proof: If $x \in [Mf > \lambda]$, then there is $r_x < 1$ such that

$$\frac{1}{m(B(0,r_x))} \int_{B(x,r_x)} |f|\, dm > \lambda,\ m(B(0,r_x)) \leq \frac{1}{\lambda} \int_{B(x,r_x)} |f|\, dm$$

Then $\{B(x,r_x) : x \in [Mf > \lambda]\}$ is an open cover of $[Mf > \lambda]$. By the Vitali covering theorem, there are countably many of these balls $\{B_i\}_{i=1}^{\infty}$ which are disjoint and $\cup_i \hat{B}_i \supseteq [Mf > \lambda]$. Therefore,

$$\begin{aligned} \bar{m}([Mf > \lambda]) &\leq m\left(\cup_i \hat{B}_i\right) \leq \sum_{i=1}^{\infty} m\left(\hat{B}_i\right) \\ &= 5\sum_{i=1}^{\infty} m(B_i) \leq 5\sum_{i=1}^{\infty} \frac{1}{\lambda}\int_{B_i} |f|\,dm \leq \frac{5}{\lambda}\int |f|\,dm \end{aligned}$$

the last step resulting from the fact that the balls B_i are disjoint. ■

Next is a version of the Vitali covering theorem which involves covering a set with disjoint closed balls in such a way that what is left over has measure 0. Here is the concept of a Vitali covering.

Definition 14.12.8 *Let S be a set and let $\mathscr{C}$ be a covering of S meaning that every point of S is contained in a set of $\mathscr{C}$. This covering is said to be a Vitali covering if for each $\varepsilon > 0$ and $x \in S$, there exists a set $B \in \mathscr{C}$ containing x, the diameter of B is less than ε, and there exists an upper bound to the set of diameters of sets of $\mathscr{C}$.*

In the following, F will just be a bounded set measurable or not. Actually, you don't even need to consider it to be bounded but this will not be needed here.

Corollary 14.12.9 *Let F be a bounded set and let $\mathscr{C}$ be a Vitali covering of F consisting of balls. Let $r(B)$ denote the radius of one of these balls. Then assume also that $\sup\{r(B) : B \in \mathscr{C}\} = M < \infty$. Then there is a countable subset of $\mathscr{C}$ denoted by $\{B_i\}$ such that $\bar{m}\left(F \setminus \cup_{i=1}^{N} B_i\right) = 0, N \leq \infty$, and $B_i \cap B_j = \emptyset$ whenever $i \neq j$.*

Proof: If $\bar{m}(F) = 0$, there is nothing to show. Assume then that $\bar{m}(F) \neq 0$. Let U be a bounded open set containing F such that U approximates F so well that $m(U) \leq \frac{10}{9}\bar{m}(F)$. This is possible because of regularity of m shown earlier. Since this is a Vitali covering, for each $x \in F$, there is one of these balls B containing x such that $\hat{B} \subseteq U$. Recall this means you keep the same center but make the radius 5 times as large. Let $\widehat{\mathscr{C}}$ denote those balls of $\mathscr{C}$ such that $\hat{B} \subseteq U$ also. Thus, this is also a cover of F. By the Vitali covering theorem above, Proposition 14.12.3, there are disjoint balls from $\mathscr{C}$ $\{B_i\}$ such that $\{\hat{B}_i\}$ covers F. Here the $\hat{B}_i$ are open balls. Thus

$$\begin{aligned} \bar{m}\left(F \setminus \cup_{j=1}^{\infty} B_j\right) &\leq m(U) - m\left(\cup_{j=1}^{\infty} B_j\right) < \frac{10}{9}\bar{m}(F) - \sum_{j=1}^{\infty} m(B_j) \\ &= \frac{10}{9}\bar{m}(F) - \frac{1}{5}\sum_{j=1}^{\infty} m\left(\widehat{B_j}\right) \leq \frac{10}{9}\bar{m}(F) - \frac{1}{5}\bar{m}(F) = \bar{m}(F)\,\theta_1 \end{aligned}$$

Here $\theta_1 = \frac{41}{45} < 1$.

Now $F \setminus \cup_{k=1}^{n} B_k \subseteq \left(F \setminus \cup_{k=1}^{\infty} B_k\right) \cup \left(\cup_{k=1}^{n} B_k\right)$. Indeed, if x is in F but not in any of the B_k for $k \leq n$, then if it fails to be in $F \setminus \cup_{k=1}^{\infty} B_k$, so it must be in the complement of this set. Hence it is in $\left(\cap_{k=1}^{\infty}\left(F \cap B_k^C\right)\right)^C = \cup_{k=1}^{\infty}\left(F^C \cup B_k\right)$ but it is not in any of the B_k for $k \leq n$ while being in F so it must be in $\cup_{k=n+1}^{\infty} B_k$. It follows, since the B_k are disjoint, that

$$\bar{m}\left(F \setminus \cup_{j=1}^{n} B_j\right) \leq \bar{m}\left(F \setminus \cup_{j=1}^{\infty} B_j\right) + \sum_{k=n}^{\infty} m(B_k) \leq \theta_1 \bar{m}(F) + \sum_{k=n}^{\infty} m(B_k), \theta_1 < 1$$

Now these balls B_k are disjoint and contained in a bounded set so letting $1 > \theta > \theta_1$, if n is large enough, and since the sum on the right is the tail of a convergent series,

$$\sum_{k=n}^{\infty} m(B_k) \leq (\theta - \theta_1)\bar{m}(F)$$

Thus there exists n_1 such that $\bar{m}\left(F \setminus \cup_{j=1}^{n_1} B_j\right) < \theta \bar{m}(F)$. If $\bar{m}\left(F \setminus \cup_{j=1}^{n_1} B_j\right) = 0$, stop. Otherwise, do for $F \setminus \cup_{j=1}^{n_1} B_j$ exactly the same thing that was done for F. Since $\cup_{j=1}^{n_1} B_j$ is closed, you can arrange to have the approximating open set be contained in the open set $\left(\cup_{j=1}^{n_1} B_j\right)^C$. It follows there exist disjoint closed balls from $\mathscr{C}$ called $B_{n_1+1}, \cdots, B_{n_2}$ such that

$$\bar{m}\left(\left(F \setminus \cup_{j=1}^{n_1} B_j\right) \setminus \cup_{j=n_{1+1}}^{n_2} B_j\right) < \theta \bar{m}\left(F \setminus \cup_{j=1}^{n_1} B_j\right) < \theta^2 \bar{m}(F)$$

continuing this way and noting that $\lim_{n\to\infty} \theta^n = 0$ while $m(F) < \infty$, this shows the desired result. Either the process stops because $\bar{m}\left(F \setminus \cup_{j=1}^{n_k} B_j\right) = 0$ or else you obtain an infinite sequence $\{B_k\}$ and $\bar{m}\left(F \setminus \cup_{j=1}^{\infty} B_j\right) \leq \bar{m}\left(F \setminus \cup_{j=1}^{n_k} B_j\right) \leq \theta^k \bar{m}(F)$ for each k, showing that $\bar{m}\left(F \setminus \cup_{j=1}^{\infty} B_j\right) = 0$. ■

14.13 Differentiation of Increasing Functions

As a spectacular application of the covering theorem, is the famous theorem that an increasing function has a derivative a.e. Here the a.e. refers to Lebesgue measure, the Stieltjes measure from the increasing function $F(x) = x$.

I will write $y_n \uparrow x$ to mean $\lim_{n\to\infty} y_n = x$ and $y_n < x, y_n \leq y_{n+1}$. I will write $y_n \downarrow x$ to mean $\lim_{n\to\infty} y_n = x$ and $y_n > x, y_n \geq y_{n+1}$.

Definition 14.13.1 *The Dini derivates are as follows. In these formulas, f is a real valued function defined on $\mathbb{R}$. $y_n \downarrow x$ refers to a decreasing sequence as just described.*

$$\begin{aligned}
D^+ f(x) &\equiv \sup\left\{\limsup_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} : y_n \downarrow x\right\}, \\
D_+ f(x) &\equiv \inf\left\{\liminf_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} : y_n \downarrow x\right\}, \\
D^- f(x) &\equiv \sup\left\{\limsup_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} : y_n \uparrow x\right\}, \\
D_- f(x) &\equiv \inf\left\{\liminf_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} : y_n \uparrow x\right\}
\end{aligned}$$

The notation means that the sup *and* inf *refer to all sequences of the sort described in* $\{\}$.

Lemma 14.13.2 *The function $f : \mathbb{R} \to \mathbb{R}$ has a derivative if and only if all the Dini derivates are equal for any sequence just described.*

Proof: If $D^+ f(x) = D_+ f(x)$, then for any $y_n \downarrow x$, it must be the case that

$$\begin{aligned}
\limsup_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} &= \liminf_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} \\
&= \lim_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} = D^+ f(x) = D_+ f(x)
\end{aligned}$$

whenever $y_n \downarrow x$. Therefore, it would follow that the limit of these difference quotients would exist and f would have a right derivative at x. Therefore, $D^+ f(x) > D_+ f(x)$ if and only if there is no right derivative. Similarly $D^- f(x) > D_- f(x)$ if and only if there is no derivative from the left at x. Also, there is a derivative if and only if there is a derivative from the left, right and the two are equal. Thus this happens if and only if all Dini derivates are equal. ■

The Lebesgue measure of single points is 0 and so we do not need to worry about whether the intervals are closed in using Corollary 14.12.9.

Let $\Delta f(I) = f(b) - f(a)$ if I is an interval having end points $a < b$. Now suppose $\{J_j\}$ are disjoint intervals contained in I. Then, since f is increasing, $\Delta f(I) \geq \sum_j \Delta f(J_j)$. In this notation, the above lemma implies that if $D^- f(x) > b$ or $D^+ f(x) > b$, then for each $\varepsilon > 0$ there is an interval J of length less than ε which is centered at x and $\frac{\Delta f(J)}{m(J)} > b$ where $m(J)$ is the Lebesgue measure of J which is the length of J. If either $D_- f(x)$ or $D_+ f(x) < a$, the above lemma implies that for each $\varepsilon > 0$ there exists I centered at x with $m(I) < \varepsilon$ and $\frac{\Delta f(I)}{m(I)} < a$. For example, if $D^- f(x) < a$, there exists a sequence $y_n \uparrow x$ with

$$\frac{f(y_n) - f(x)}{y_n - x} = \frac{f(x) - f(y_n)}{x - y_n} < a$$

so let I_n be the interval (y_n, x). For large n it is smaller than ε.

Theorem 14.13.3 *Let $f : \mathbb{R} \to \mathbb{R}$ be increasing. Then $f'(x)$ exists for all x off a set of measure zero.*

Proof: Let N_{ab} for $0 < a < b$ denote either

$$\left\{x : D^+ f(x) > b > a > D_+ f(x)\right\}, \left\{x : D^- f(x) > b > a > D_- f(x)\right\},$$

or

$$\left\{x : D^- f(x) > b > a > D_+ f(x)\right\}, \left\{x : D^+ f(x) > b > a > D_- f(x)\right\}.$$

The function f is increasing and so it is a Borel measurable function. Indeed, $f^{-1}(a, \infty)$ is either open or closed. Therefore, all these derivates are also Borel measurable, hence Lebesgue measurable. Assume that N_{ab} is bounded and let V be open with

$$V \supseteq N_{ab}, \quad m(N_{ab}) + \varepsilon > m(V)$$

By Corollary 14.12.9 and the above discussion, there are open, disjoint intervals $\{I_n\}$, each centered at a point of N_{ab} such that

$$\frac{\Delta f(I_n)}{m(I_n)} < a, \; m(N_{ab}) = m(N_{ab} \cap \cup_i I_i) = \sum_i m(N_{ab} \cap I_i)$$

Now do for $N_{ab} \cap I_i$ what was just done for N_{ab} and get disjoint intervals J_i^j contained in I_i with

$$\frac{\Delta f\left(J_i^j\right)}{m\left(J_i^j\right)} > b, \; m(N_{ab} \cap I_i) = \sum_j m\left(N_{ab} \cap I_i \cap J_i^j\right)$$

Then

$$\begin{aligned} a\left(m\left(N_{ab}\right)+\varepsilon\right) &> am\left(V\right)\geq a\sum_i m\left(I_i\right)>\sum_i \Delta f\left(I_i\right)\geq \sum_i\sum_j \Delta f\left(J_i^j\right) \\ &\geq b\sum_i\sum_j m\left(J_i^j\right)\geq b\sum_i\sum_j m\left(J_i^j\cap N_{ab}\right) \\ &= b\sum_i m\left(N_{ab}\cap I_i\right)=bm\left(N_{ab}\right) \end{aligned}$$

Since ε is arbitrary and $a<b$, this shows $m\left(N_{ab}\right)=0$. If N_{ab} is not bounded, apply the above to $N_{ab}\cap(-r,r)$ and conclude this has measure 0. Hence so does N_{ab}.

The countable union of N_{ab} for a,b positive rational numbers is an exceptional set off which

$$D^{+}f\left(x\right)=D_{+}f\left(x\right)\geq D^{-}f\left(x\right)\geq D_{-}f\left(x\right)\geq D^{+}f\left(x\right)$$

and so these are all equal. This shows that off a set of measure zero, the function has a derivative a.e. ■

14.14 Exercises

1. Show carefully that if $\mathfrak{S}$ is a set whose elements are σ algebras which are subsets of $\mathscr{P}\left(\Omega\right)$, the set of all subsets of Ω, then $\cap\mathfrak{S}$ is also a σ algebra. Now let $\mathscr{G}\subseteq\mathscr{P}\left(\Omega\right)$ satisfy property P if $\mathscr{G}$ is closed with respect to complements and countable disjoint unions as in Dynkin's lemma, and contains $\emptyset$ and Ω. If $\mathfrak{H}\subseteq\mathscr{G}$ is any set whose elements are subsets of $\mathscr{P}\left(\Omega\right)$ which satisfies property P, then $\cap\mathfrak{H}$ also satisfies property P. Thus there is a smallest subset of $\mathscr{G}$ satisfying P. Show these things.

2. Show $f:(\Omega,\mathscr{F})\to\mathbb{R}$ is measurable if and only if $f^{-1}\left(\text{open}\right)\in\mathscr{F}$. Show that if E is any set in $\mathscr{B}\left(\mathbb{R}\right)$, then $f^{-1}\left(E\right)\in\mathscr{F}$. Thus, inverse images of Borel sets are measurable. Next consider $f:(\Omega,\mathscr{F})\to\mathbb{R}$ being measurable and $g:\mathbb{R}\to\mathbb{R}$ is Borel measurable, meaning that $g^{-1}\left(open\right)\in\mathscr{B}\left(\mathbb{R}\right)$. Explain why $g\circ f$ is measurable. **Hint:** You know that $\left(g\circ f\right)^{-1}\left(U\right)=f^{-1}\left(g^{-1}\left(U\right)\right)$. For your information, it does not work the other way around. That is, measurable composed with Borel measurable is not necessarily measurable. In fact examples exist which show that if g is measurable and f is continuous, then $g\circ f$ may fail to be measurable. This is in the chapter, but show it anyway.

3. You have two finite measures defined on $\mathscr{B}\left(\mathbb{R}\right)$ μ,ν. Suppose these are equal on every open set. Show that these must be equal on every Borel set. **Hint:** You should use Dynkin's lemma to show this very easily.

4. You have two measures defined on $\mathscr{B}\left(\mathbb{R}\right)$ which are finite and equal on every open set. Show, using Dynkin's lemma that these are the same on all Borel sets.

5. Let $\mu\left(E\right)=1$ if $0\in E$ and $\mu\left(E\right)=0$ if $0\notin E$. Show this is a measure on $\mathscr{P}\left(\mathbb{R}\right)$.

6. Give an example of a measure μ and a measure space and a decreasing sequence of measurable sets $\{E_i\}$ such that $\lim_{n\to\infty}\mu\left(E_n\right)\neq\mu\left(\cap_{i=1}^{\infty}E_i\right)$.

7. You have a measure space $(\Omega,\mathscr{F},P)$ where P is a probability measure on $\mathscr{F}$. Thus $P\left(\Omega\right)=1$. Then you also have a measurable function $X:\Omega\to\mathbb{R}$, meaning that

$X^{-1}(U) \in \mathscr{F}$ whenever U is open. Now define a measure on $\mathscr{B}(\mathbb{R})$ denoted by λ_X and defined by $\lambda_X(E) = P(\{\omega : X(\omega) \in E\})$. Explain why this yields a well defined probability measure on $\mathscr{B}(\mathbb{R})$. This is called the distribution measure.

8. Let $(\Omega, \mathscr{F})$ be a measurable space and let $f : \Omega \to \mathbb{R}$ be a measurable function meaning that $f^{-1}(U) \in \mathscr{F}$ whenever U is open. Then $\sigma(f)$ denotes the smallest σ algebra such that f is measurable with respect to this σ algebra. Show that $\sigma(f) = \left\{f^{-1}(E) : E \in \mathscr{B}(\mathbb{R})\right\}$.

9. There is a monumentally important theorem called the Borel Cantelli lemma. It says the following. If you have a measure space $(\Omega, \mathscr{F}, \mu)$ and if $\{E_i\} \subseteq \mathscr{F}$ is such that $\sum_{i=1}^{\infty} \mu(E_i) < \infty$, then there exists a set N of measure 0 ($\mu(N) = 0$) such that if $\omega \notin N$, then ω is in only finitely many of the E_i. **Hint:** You might look at the set of all ω which are in infinitely many of the E_i. First explain why this set is of the form $\cap_{n=1}^{\infty} \cup_{k \geq n} E_k$.

10. Let $(\Omega, \mathscr{F}, \mu)$ be a measure space. A sequence of functions $\{f_n\}$ is said to converge in measure to a measurable function f if and only if for each $\varepsilon > 0$,
$$\lim_{n\to\infty} \mu(\{\omega : |f_n(\omega) - f(\omega)| > \varepsilon\}) = 0$$
Show that if this happens, then there exists a subsequence $\left\{f_{n_k}\right\}$ and a set of measure N such that if $\omega \notin N$, then $\lim_{n_k\to\infty} f_{n_k}(\omega) = f(\omega)$. Also show that if μ is finite and $\lim_{n\to\infty} f_n(\omega) = f(\omega)$, then f_n converges in measure to f.

11. Prove Chebyshev's inequality $\mu(\{\omega : |f(\omega)| > \lambda\}) \leq \frac{1}{\lambda} \int |f| \, d\mu$.

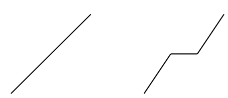
12. $\uparrow$Use the above inequality to show that if $\lim_{n\to\infty} \int |f_n| \, d\mu = 0$, then there is a set of measure zero N and a subsequence, still called $\{f_n\}$ such that for $\omega \notin N, f_n(\omega) \to 0$. **Hint:** Get a subsequence using the above Chebyshev inequality, still denoted as $\{f_n\}$, such that $\mu\left(\left[|f_n(\omega)| > \frac{1}{n}\right]\right) < C2^{-n}$ Then use the Borel Cantelli lemma of Problem 9 to conclude that off a set of measure zero, $|f_n(\omega)| < \frac{1}{n}$ for all n large enough.

13. Let $\{r_n\}_{n=1}^{\infty}$ be an enumeration of the rational numbers in $[0,1]$ meaning that every rational number is included in $\{r_n\}_{n=1}^{\infty}$ for some n and let $f_n(x) = 0$ except for when $x \in \{r_1, \cdots, r_n\}$ when it is 1. Explain why f_n is Riemann integrable and has Riemann integral 0. However, $\lim_{n\to\infty} f_n(x) \equiv f(x)$ is 1 on rationals and 0 elsewhere so this isn't even Riemann integrable with respect to the integrator $F(t) = t$. It can be shown that the two integrals give the same answer whenever the function is Riemann integrable. Thus the Lebesgue integral of f_n will be 0. So what is the Lebesgue integral of the function which is 1 on the rationals and 0 on the irrationals? Explain why this is so.

14. In fact, μ_F being only defined on $\mathscr{B}(\mathbb{R})$ might not be a complete measure. This means that you can have $A \subseteq B \subseteq C$ and $\mu(A) = \mu(C)$ with both A, C measurable but B is not. Give a way to enlarge $\mathscr{B}(\mathbb{R})$ to a larger σ algebra, extending μ so that the result is a σ algebra with measure which is a complete measure space, meaning that if $A \subseteq B$ and $\mu(B) = 0$ with B measurable, then A is also measurable. **Hint:** Let a null set be one which is contained in a measurable set of measure zero. Denoting such sets with N, let $\mathscr{F} \equiv \{A \cup N$ where N is a null set, $A \in \mathscr{B}(\mathbb{R})\}$ and let $\hat{\mu}(A \cup N) \equiv \mu(A)$.

15. Lebesgue measure was discussed. Recall that $m((a,b)) = b-a$ and it is defined on a σ algebra which contains the Borel sets. Show, using Dynkin's lemma, that m is translation invariant meaning that $m(E) = m(E+a)$. for all Borel sets E and then explain why this will be true for all Lebesgue measurable sets described in Problem 14. Let $x \sim y$ if and only if $x - y \in \mathbb{Q}$. Show this is an equivalence relation. Now let W be a set of positive measure which is contained in $(0,1)$. For $x \in W$, let $[x]$ denote those $y \in W$ such that $x \sim y$. Thus the equivalence classes partition W. Use axiom of choice to obtain a set $S \subseteq W$ such that S consists of exactly one element from each equivalence class. Let $\mathbb{T}$ denote the rational numbers in $[-1,1]$. Consider $\mathbb{T} + S \subseteq [-1,2]$. Explain why $\mathbb{T} + S \supseteq W$. For $\mathbb{T} \equiv \{r_j\}$, explain why the sets $\{r_j + S\}_j$ are disjoint. Explain why S cannot be measurable. Explain why this shows that for any Lebesgue measurable set of positive measure, there is a subset of this set which is not measurable.

16. Consider the sequence of functions defined in the following way. 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. The process of modifying a nonconstant section of the graph of this function is illustrated in the following picture.



Show $\{f_n\}$ converges uniformly on $[0,1]$. If $f(x) = \lim_{n\to\infty} f_n(x)$, show that $f(0) = 0$, $f(1) = 1$, f is continuous, and $f'(x) = 0$ for all $x \notin P$ where P is the Cantor set of Problem 14. This function is called the Cantor function.It is a very important example to remember. Note it has derivative equal to zero a.e. and yet it succeeds in climbing from 0 to 1. Explain why this function cannot be recovered as an integral of its derivative even though the derivative exists everywhere but on a set of measure zero. **Hint:** This isn't too hard if you focus on getting a careful estimate on the difference between two successive functions in the list considering only a typical small interval in which the change takes place. The above picture should be helpful.

17. $\uparrow$ This problem gives a very interesting example found in the book by McShane [23]. Let $g(x) = x + f(x)$ where f is the strange function of Problem 16. Let P be the Cantor set of Problem 14. Let $[0,1] \setminus P = \cup_{j=1}^{\infty} I_j$ where I_j is open and $I_j \cap I_k = \emptyset$ if $j \neq k$. These intervals are the connected components of the complement of the Cantor set. Show $m(g(I_j)) = m(I_j)$ so

$$m(g(\cup_{j=1}^{\infty} I_j)) = \sum_{j=1}^{\infty} m(g(I_j)) = \sum_{j=1}^{\infty} m(I_j) = 1.$$

Thus $m(g(P)) = 1$ because $g([0,1]) = [0,2]$. By Problem 15 there exists a set, $A \subseteq g(P)$ which is non measurable. Define $\phi(x) = \mathscr{X}_A(g(x))$. Thus $\phi(x) = 0$ unless $x \in P$. Tell why ϕ is measurable. (Recall $m(P) = 0$ and Lebesgue measure is complete.) Now show that $\mathscr{X}_A(y) = \phi(g^{-1}(y))$ for $y \in [0,2]$. Tell why g^{-1} is continuous but $\phi \circ g^{-1}$ is not measurable. (This is an example of measurable $\circ$ continuous $\neq$

measurable.) Show there exist Lebesgue measurable sets which are not Borel measurable. **Hint**: The function, ϕ is Lebesgue measurable. Now recall that Borel $\circ$ measurable = measurable.

18. Suppose you have a function f which is of bounded variation on every closed interval. Explain why this function has a derivative off some Borel set of measure zero.

19. Suppose $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz. This means $|f(x) - f(y)| \leq K|x-y|$. Then let $g(x) \equiv (K+1)x - f(x)$. Explain why g is monotone. Note $f(x) = (K+1)x - g(x)$. Explain why $f'(x)$ exists off a Borel set of measure zero. This is a case of Rademacher's theorem which says that a Lipschitz map is differentiable almost everywhere. In the general case, the function is vector valued and defined on $\mathbb{R}^p$ rather than $\mathbb{R}$.

20. Suppose $f \geq 0$, is Lebesgue measurable, see Problem 14, or if you didn't do that one, let it be Borel measurable. Also suppose $f\mathscr{X}_{[a,b]}$ has a finite integral for any finite interval $[a,b]$. Let $F(x) \equiv \int_0^x f(t)\,dm$. Show that this function of x is continuous everywhere on $[a,b]$ and then explain why this function has a derivative off some Borel set of measure zero. **Hint:** To show continuous, note that for $x < b$, $\mathscr{X}_{[x,x+h]}f(x) \to 0$ as $h \to 0$. Consider using the dominated convergence theorem.

21. Let $S \neq \emptyset$ and define $\operatorname{dist}(x,S) \equiv \inf\{|x-y| : y \in S\}$. Show $x \to \operatorname{dist}(x,S)$ is continuous. In fact, show that
$$|\operatorname{dist}(x,S) - \operatorname{dist}(y,S)| \leq |x-y|$$

22. $\uparrow$From regularity theorems for measures on $\mathscr{B}(\mathbb{R})$, you know that if $\mu(E) < \infty$, then there is a compact K and open V such that $K \subseteq E \subseteq V$ and $\mu(V \setminus K) < \delta$. Here μ is a Borel measure which is finite on bounded sets. Show that for such E, there is a continuous function f which is 1 on K, 0 off V, and has values in $[0,1]$. Use this to show that there exists a continuous function h such that $\int |\mathscr{X}_E - h|\,d\mu < \varepsilon$. **Hint:** You might get such compact sets K_n and open sets V_n such that $\mu(V_n \setminus K_n) < 1/2^n$ and let h_n be as just shown. Then do some sort of argument involving the dominated convergence theorem.

23. $\uparrow$Let μ be a measure on $\mathscr{B}(\mathbb{R})$ which is finite on bounded intervals and $\int |f|\,d\mu < \infty$. Show there exists a simple function s such that $|s(x)| \leq |f(x)|$ for all x and
$$\int |f(x) - s(x)|\,d\mu < \varepsilon/2$$
Next show there is a continuous function h such that
$$\int |s(x) - h(x)|\,d\mu < \varepsilon/2$$
Conclude that if f is real valued and measurable with $\int |f|\,d\mu < \infty$, then there is h continuous such that $\int |f-h|\,d\mu < \varepsilon$.

24. $\uparrow$Let μ be a measure on $\mathscr{B}(\mathbb{R})$ which is finite on bounded intervals and $\int |f|\,d\mu < \infty$. Show that if $f_n \equiv f\mathscr{X}_{[-n,n]}$, then if n is large enough,
$$\int |f - f_n|\,d\mu < \varepsilon$$

Modify the above problem slightly to show that it can be assumed that h vanishes outside of some finite interval.

25. A Lebesgue measurable function f is said to be in L^1_{loc} if $f\mathscr{X}_{[a,b]}$ has a finite integral. Show that there is a Borel set of measure zero off which $x \to \int_0^x f(t)\,dm$ has a derivative a.e.

26. This is on the Lebesgue fundamental theorem of calculus. Let $f \in L^1(\mathbb{R}, m)$. By Problem 24 for each ε, there is g continuous which vanishes off a finite interval such that $\int |f-g|\,dm < \varepsilon$. Such a g is uniformly continuous by Theorem 6.7.2. Fill in details to the following argument. For g as just described,

$$\lim_{r\to 0} \frac{1}{m(B(0,r))} \int_{B(x,r)} |g(x) - g(y)|\,dm(y) = 0$$

Here x is fixed and the function integrated is $y \to |g(x) - g(y)|$. Now let $r_n \to 0+$ be an arbitrary sequence each less than 1. Then

$$\limsup_{n\to\infty} \left(\frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |f(y) - f(x)|\,dm(y) \right) \leq$$

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |f(y) - g(y)|\,dm(y) \\ &+\limsup_{n\to\infty} \frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |g(y) - g(x)|\,dm(y) \\ &+\limsup_{n\to\infty} \frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |g(x) - f(x)|\,dm(y) \\ &\leq M(f-g)(x) + |g(x) - f(x)| \end{aligned}$$

Then

$$\bar{m}\left[x : \limsup_{n\to\infty} \left(\frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |f(y) - f(x)|\,dm(y) \right) > \lambda \right] \leq$$

$$\bar{m}\left[M(f-g) > \lambda/2\right] + \bar{m}\left[|g-f| > \lambda/2\right]$$

By Theorem 14.12.7 and Problem 11

$$\leq \frac{10}{\lambda} \int |f-g|\,dm + \frac{2}{\lambda} \int |g-f|\,dm$$

We are free to choose how close g is to f and so

$$\bar{m}\left(\left[x : \limsup_{n\to\infty} \left(\frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |f(y) - f(x)|\,dm(y) \right) > \lambda \right] \right) = 0$$

It follows that

$$\bar{m}\left(\left[x : \limsup_{n\to\infty} \left(\frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |f(y) - f(x)|\,dm(y) \right) > 0 \right] \right) = 0$$

and so, there is a Borel set of measure zero N such that if $x \notin N$,

$$\lim_{n\to\infty} \frac{1}{m(B(0,r_n))} \int_{B(x,r_n)} |f(y) - f(x)|\, dm(y) = 0$$

and so

$$\lim_{r\to\infty} \frac{1}{m(B(0,r))} \int_{B(x,r)} |f(y) - f(x)|\, dm(y) = 0$$

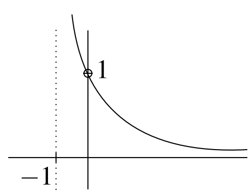
which implies for $x \notin N$,

$$\lim_{r\to 0} \frac{1}{m(B(0,r))} \int_{B(x,r)} f(y)\, dm(y) = f(x)$$

27. If you only know that $f\mathscr{X}_{[a,b]} \in L^1(\mathbb{R}, m)$ for all finite intervals $[a,b]$, show exactly the same conclusion holds.

28. Use the above result to show that if $f\mathscr{X}_{[a,b]} \in L^1(\mathbb{R}, m)$ for all finite intervals $[a,b]$, then for a.e. x,
$$\lim_{h\to 0} \frac{1}{h} \int_0^{x+h} f(t)\, dm = f(x)$$
which can also be termed the Lebesgue fundamental theorem of calculus.

29. This problem outlines an approach to Stirling's formula following [24] and [9]. An easier approach comes from the traditional Stirling's formula and Problem 16 on Page 250 From the above problems, $\Gamma(n+1) = n!$ for $n \geq 0$. Consider more generally $\Gamma(x+1)$ for $x > 0$. Actually, we will always assume $x > 1$ since it is the limit as $x \to \infty$ which is of interest. $\Gamma(x+1) = \int_0^\infty e^{-t} t^x dt$. Change variables letting $t = x(1+u)$ to obtain $\Gamma(x+1) = x^{x+1} e^{-x} \int_{-1}^\infty \left((1+u)e^{-u}\right)^x du$ Next let $h(u)$ be such that $h(0) = 1$ and $(1+u)e^{-u} = \exp\left(-\frac{u^2}{2} h(u)\right)$ Show the thing which works is $h(u) = \frac{2}{u^2}(u - \ln(1+u))$. Use L'Hospital's rule to verify that the limit of $h(u)$ as $u \to 0$ is 1. The graph of h is illustrated in the following picture. Verify that its graph is like this, with an asymptote at $u = -1$ decreasing and equal to 1 at 0 and converging to 0 as $u \to \infty$.



Next change the variables again letting $u = s\sqrt{\frac{2}{x}}$. This yields, from the original description of h, $\Gamma(x+1) = x^x e^{-x}\sqrt{2x} \int_{-\sqrt{x/2}}^\infty \exp\left(-s^2 h\left(s\sqrt{\frac{2}{x}}\right)\right) ds$. For $s < 1$,

$$h\left(s\sqrt{\frac{2}{x}}\right) > 2 - 2\ln 2 = 0.61371$$

so the above integrand is dominated by $e^{-(2-2\ln 2)s^2}$. Consider the integrand in the above for $s \geq 1$. The exponent part is $-\left(\sqrt{2}\sqrt{x}s - x\ln\left(1 + s\sqrt{\frac{2}{x}}\right)\right)$. Now for

$x > 1$, $\left(\sqrt{2}\sqrt{x}s - x\ln\left(1+s\sqrt{\frac{2}{x}}\right)\right) - \left(\sqrt{2}s - \ln\left(1+s\sqrt{2}\right)\right)$ equals the expression $\sqrt{2}s\left(\sqrt{x}-1\right) + \ln\left(\frac{1+s\sqrt{2}}{\left(1+s\sqrt{\frac{2}{x}}\right)^x}\right) > 0$.

Therefore, $\left(\sqrt{2}\sqrt{x}s - x\ln\left(1+s\sqrt{\frac{2}{x}}\right)\right) > \left(\sqrt{2}s - \ln\left(1+s\sqrt{2}\right)\right)$ and so, for $s \geq 1$,

$$\exp\left(-s^2 h\left(s\sqrt{\frac{2}{x}}\right)\right) \leq \exp\left(-\left(\sqrt{2}s - \ln\left(1+s\sqrt{2}\right)\right)\right) = \left(1+s\sqrt{2}\right)e^{-\sqrt{2}s}$$

Thus, there exists a dominating function for $\mathscr{X}_{\left[-\sqrt{\frac{x}{2}},\infty\right]}(s)\exp\left(-s^2 h\left(s\sqrt{\frac{2}{x}}\right)\right)$ and $\lim_{x\to\infty}\mathscr{X}_{\left[-\sqrt{\frac{x}{2}},\infty\right]}(s)\exp\left(-s^2 h\left(s\sqrt{\frac{2}{x}}\right)\right) = \exp\left(-s^2\right)$ so by the dominated convergence theorem,

$$\lim_{x\to\infty}\int_{-\sqrt{x/2}}^{\infty}\exp\left(-s^2 h\left(s\sqrt{\frac{2}{x}}\right)\right)ds = \int_{-\infty}^{\infty}e^{-s^2}ds = \sqrt{\pi}$$

See Problem 49 on Page 227 or Theorem 9.9.4. This yields a general Stirling's formula, $\lim_{x\to\infty}\frac{\Gamma(x+1)}{x^x e^{-x}\sqrt{2x}} = \sqrt{\pi}$.

Chapter 15

Construction of Real Numbers

The purpose of this chapter is to give a construction of the real numbers from the rationals. This was first done by Dedekind in 1858 although not published till 1872. He used a totally different approach than what is used here. Dedekind's construction was based on an early idea of Eudoxus who lived around 350 B.C. in which points on the real line divided the rationals above and below them. Dedekind made this cut into what was meant by a real number. See Rudin [24] for a description of this approach. I am using equivalence classes of Cauchy sequences rather than Dedekind cuts because this approach applies to metric spaces and Dedekind cuts do not. This idea of constructing the real numbers from such equivalence classes is due to Cantor, also published in 1872. Hewitt and Stromberg [17] also use this approach. Hobson [18] has a description of both of these methods.

Why do this? Why not continue believing that a real number is a point on a number line and regard its existence as geometrically determined? This is essentially what was done till Dedekind and, as mentioned, seems to have been part of the belief system of Greeks thousands of years earlier. I think the reason that such a construction is needed is algebra. Is there a way to carry the algebraic notions and order axioms of the rational numbers to the real numbers, defined geometrically as points on the number line? Till Dedekind, this was simply assumed, as it has been in this book till now. Also, it is desirable to remove the last vestiges of geometry from analysis; hence this construction. I think this is why Dedekind and Cantor's work was very important.

Definition 15.0.1 *Let* $\mathbf{R}$ *denote the set of Cauchy sequences of rational numbers. If* $\{x_n\}_{n=1}^{\infty}$ *is such a Cauchy sequence, this will be denoted by* $\mathbf{x}$ *for the sake of simplicity of notation. A Cauchy sequence* $\mathbf{x}$ *will be called a null sequence if* $\lim_{n\to\infty} x_n = 0$. *Denote the collection of such null Cauchy sequences as* $\mathbf{N}$.*Then define* $\mathbf{x} \sim \mathbf{y}$ *if and only if* $\mathbf{x}-\mathbf{y} \in \mathbf{N}$. *Here* $\mathbf{x}-\mathbf{y}$ *signifies the Cauchy sequence* $\{x_n - y_n\}_{n=1}^{\infty}$. *Also, for simplicity of notation, let* $\mathbf{Q}$ *denote the collection of Cauchy sequences equivalent to some constant Cauchy sequence* $\{a_n\}_{n=1}^{\infty}$ *for* $a_n = a \in \mathbb{Q}$. *Thus* $\mathbf{Q} \subseteq \mathbf{R}$.

Notice that whether $\mathbf{x}-\mathbf{y} \in \mathbf{N}$ is determined completely by the tail of $\mathbf{x}-\mathbf{y}$, the terms of the sequence larger than some number. Then the following proposition is very easy to establish and is left to the reader.

Proposition 15.0.2 $\sim$ *is an equivalence relation on* $\mathbf{R}$.

Definition 15.0.3 *Define* $\mathbb{R}$ *as the set of equivalence classes of* $\mathbf{R}$. *For* $[\mathbf{x}], [\mathbf{y}], [\mathbf{z}] \in \mathbb{R}$, *define* $[\mathbf{x}][\mathbf{y}] \equiv [\mathbf{xy}]$ *where* $\mathbf{xy}$ *is the sequence* $\{x_n y_n\}_{n=1}^{\infty}$. *Also define* $[\mathbf{x}]+[\mathbf{y}] \equiv [\mathbf{x}+\mathbf{y}]$.

This leads to the following theorem. Note that this is enlarging the field $\mathbb{Q}$ obtaining a larger field $\mathbb{R}$. Enlarging fields is done frequently in algebra using the machinery of field extensions. It also uses equivalence classes. However, this is very different, resulting in an enlargement of $\mathbb{Q}$ which essentially goes all the way at once. It emphasizes completeness and order rather than inclusion of roots of various polynomials.

Lemma 15.0.4 *If* $\mathbf{x} \notin \mathbf{N}$, *then there is* $\delta > 0$ *and* N *such that* $|x_k| > \delta$ *for all* $k \geq N$.

Proof: If the conclusion does not hold, then for each $\delta > 0$, there exist infinitely many k such that $|x_k| \leq \delta$. Thus there is a subsequence which converges to 0. By Theorem 4.5.4, $\mathbf{x} \in \mathbf{N}$ after all. ■

Theorem 15.0.5 *With the two operations defined above, $\mathbb{R}$ is a field. The operations are well defined.*

Proof: Why are these operations well defined? Consider multiplication because it is fairly obvious that addition is well defined. If $\mathbf{x} \sim \mathbf{x}'$ and $\mathbf{y} \sim \mathbf{y}'$, is it true that $[\mathbf{x}'\mathbf{y}'] = [\mathbf{xy}]$? Is $\{x_n' y_n' - x_n y_n\}_{n=1}^{\infty} \in \mathbf{N}$?

$$\left|x_n' y_n' - x_n y_n\right| \leq \left|x_n' y_n' - x_n' y_n\right| + \left|x_n' y_n - x_n y_n\right| \leq C\left(\left|y_n' - y_n\right| + \left|x_n' - x_n\right|\right)$$

where C is a constant which bounds all terms of all four given Cauchy sequences, the constant existing because these are all Cauchy sequences. See Theorem 4.5.2. By assumption, the last expression converges to 0 as $n \to \infty$ and so $\{x_n' y_n' - x_n y_n\}_{n=1}^{\infty} \in \mathbf{N}$ which verifies that $[\mathbf{x}'\mathbf{y}'] = [\mathbf{xy}]$ as hoped. The case for addition is similar but easier.

Now it is necessary to verify that the two operations are binary operations on $\mathbb{R}$. This is obvious for addition. The question for multiplication reduces to whether $\mathbf{xy}$ is a Cauchy sequence.

$$|x_n y_n - x_m y_m| \leq |x_n y_n - x_m y_n| + |x_m y_n - x_m y_m| \leq C\left(|x_n - x_m| + |y_n - y_m|\right)$$

for some constant which is independent of n, m. This follows because $\mathbf{x}, \mathbf{y}$ Cauchy implies that these sequences are both bounded. See Theorem 4.5.2.

Commutative and associative laws for addition and multiplication are all obvious because these hold pointwise for individual terms of the sequence. So is the distributive law. The existence of an additive identity is clear. You just use $[\mathbf{0}]$. Similarly $[\mathbf{1}]$ is a multiplicative identity. For $[\mathbf{x}] \neq [\mathbf{0}]$, let $y_n = 1$ if $x_n = 0$ and $y_n = x_n^{-1}$ if $x_n \neq 0$. Is $\mathbf{y} \in \mathbf{R}$? Since $[\mathbf{x}] \neq [\mathbf{0}]$, Lemma 15.0.4 implies that there exists $\delta > 0$ and N such that $|x_k| > \delta$ for all $k \geq N$. Now for $m, n > N$,

$$|y_n - y_m| = \left|\frac{1}{x_n} - \frac{1}{x_m}\right| = \frac{|x_n - x_m|}{|x_n|\,|x_m|} \leq \frac{1}{\delta^2}\,|x_n - x_m|$$

which shows that $\{y_n\}_{n=1}^{\infty} \in \mathbf{R}$. Then clearly $[\mathbf{y}] = [\mathbf{x}]^{-1}$ because $[\mathbf{y}]\,[\mathbf{x}] = [\mathbf{yx}]$ and $\mathbf{yx}$ is a Cauchy sequence which equals 1 for all n large enough. Therefore, $[\mathbf{xy}] = [\mathbf{1}]$ as required, because $\mathbf{xy} - \mathbf{1} \in \mathbf{N}$. It is obvious that an additive inverse $[-\mathbf{x}] \equiv -[\mathbf{x}]$ exists for each $[\mathbf{x}] \in \mathbb{R}$. Thus $\mathbb{R}$ is a field as claimed. ■

It might be of interest to note that with the operations described, $\mathbf{R}$ is a commutative ring and $\mathbf{N}$ is a maximal ideal. Thus from algebra $\mathbf{R}/\mathbf{N}$ is a field. Showing $\mathbf{N}$ is maximal is essentially done above where if $[\mathbf{x}] \neq [\mathbf{0}]$, then the multiplicative inverse exists which gets 1 in any ideal containing $\mathbf{N}$ making $\mathbf{N}$ maximal. You do the same thing with algebraic field extensions but the argument is harder there.

Of course there are two other properties which need to be considered. Is $\mathbb{R}$ ordered? Is $\mathbb{R}$ complete? First recall what it means for $\mathbb{R}$ to be ordered. There is a subset of $\mathbb{R}$ called the positive numbers, such that

The sum of positive numbers is positive.

The product of positive numbers is positive.

$[\mathbf{x}]$ is either positive $[\mathbf{0}]$, or $-[\mathbf{x}]$ is positive.

Definition 15.0.6 *Define* $[\mathbf{x}] > [\mathbf{0}]$ *means that there exists* $\delta > 0, \delta \in \mathbb{Q}$ *and a subsequence of* $\mathbf{x}$ $\{x_{n_k}\}_{k=1}^{\infty}$ *with the property that* $x_{n_k} > \delta$ *for all* k. *Since* $\mathbf{x}$ *is a Cauchy sequence, this requires that for all* n *large enough,* $x_n > \delta/2$. *Thus we could also say* $[\mathbf{x}] > [\mathbf{0}]$ *means* $x_k > \eta$ *for some positive* η *whenever* k *is large enough. It has already been shown that* $[\mathbf{x}] = [\mathbf{0}]$ *means* $\lim_{k\to\infty} x_k = 0$.

Theorem 15.0.7 *The above definition makes* $\mathbb{R}$ *into an ordered field.*

Proof: The first two order axioms are obvious. As to the third, if $[\mathbf{x}] \neq [\mathbf{0}]$, it follows that for all k large enough, $x_k > \delta$ for some $\delta > 0$ or $-x_k < \delta$ for all k large enough. In the first case, $[\mathbf{x}] > [\mathbf{0}]$ and in the second case, $-[\mathbf{x}] > [\mathbf{0}]$. The only other case is where $[\mathbf{x}] = [\mathbf{0}]$. Thus $\mathbb{R}$ is an ordered field. ■

Now that we know this is an ordered field, the usual notions apply to $<$.

Observation 15.0.8 *To say* $[\mathbf{x}] < \varepsilon \in \mathbb{Q}$ *is to say that* $[\varepsilon] - [\mathbf{x}] = [\varepsilon - \mathbf{x}] > [\mathbf{0}]$ *where* ε *is the constant sequence every term equal to* $\varepsilon \in \mathbb{Q}$. *Thus* $\varepsilon > [\mathbf{x}]$ *says that for all* k *large enough,* $x_k < \varepsilon$.

Lemma 15.0.9 $\mathbb{Q}$ *is dense in* $\mathbb{R}$ *in the sense that given* $\varepsilon > 0$ *where* $\varepsilon \in \mathbb{Q}$, *and* $\mathbf{x} \in \mathbf{R}$, *there is* $r \in \mathbb{Q}$ *such that* $[|\mathbf{x} - \mathbf{r}|] < \varepsilon$ *where* $\mathbf{r}$ *is the constant sequence always equal to* $r \in \mathbb{Q}$.

Proof: Note that $|\mathbf{x} - \mathbf{r}| \in \mathbf{R}$ from the triangle inequality. It remains to verify $\mathbb{Q}$ is dense in $\mathbb{R}$. I need to verify that if $\mathbf{x} \in \mathbf{R}$, then there is $r \in \mathbb{Q}$ such that for large enough k, $r - x_k < \varepsilon \in \mathbb{Q}$ and $x_k - r < \varepsilon$. It suffices to get an r such that $|x_k - r| < \varepsilon$. Since $\mathbf{x} \in \mathbf{R}$, $|x_k - x_m| < \varepsilon/4$ for all $k, m \geq n$ for some n. Let $r \equiv x_n$ so $k \geq n$ implies $|x_k - r| < \varepsilon/4$. Then $\varepsilon - |x_k - r| > 3\varepsilon/4$ so $[|\mathbf{x} - \mathbf{r}|] < \varepsilon$. ■

Definition 15.0.10 *Define* $|[\mathbf{x}]| \equiv [|\mathbf{x}|]$. *This makes sense because of the triangle inequality* $||x_k| - |y_k|| \leq |x_k - y_k|$. *Thus* $|\mathbf{x}| \equiv \{|x_k|\}_k \in \mathbf{R}$ *and if* $[\mathbf{x}] = [\mathbf{y}]$ *then* $[|\mathbf{x}|] = [|\mathbf{y}|]$.

Theorem 15.0.11 $|\cdot|$ *as just defined satisfies the usual properties of the absolute value. Also, if* $[\mathbf{q}] > 0$ *and if* $[\mathbf{x}] \in \mathbb{R}$ *then there is* $r \in \mathbb{Q}$ *such that* $|[\mathbf{x}] - [\mathbf{r}]| < [\mathbf{q}]$. *Thus* $\mathbb{Q}$ *is dense in* $\mathbb{R}$. *Also* $\mathbb{R}$ *is complete.*

Proof: If $[\mathbf{x}] \neq [\mathbf{0}]$, then eventually, for large enough k either $x_k \geq \delta > 0$ or $x_k < -\delta$ for some rational $\delta > 0$. Therefore, $|[\mathbf{x}]| = [|\mathbf{x}|] > \delta$.

$$|[\mathbf{x}][\mathbf{y}]| \equiv |[\mathbf{x}\mathbf{y}]| \equiv [|\mathbf{x}\mathbf{y}|] \text{ and } |[\mathbf{x}]||[\mathbf{y}]| \equiv [|\mathbf{x}|][|\mathbf{y}|] \equiv [|\mathbf{x}\mathbf{y}|]$$

Thus $|[\mathbf{x}][\mathbf{y}]| = |[\mathbf{x}]||[\mathbf{y}]|$. Also $[\mathbf{x}]^2 = |[\mathbf{x}]|^2$ and $|[\mathbf{x}][\mathbf{y}]| \geq [\mathbf{x}][\mathbf{y}]$.

$$\begin{aligned} |[\mathbf{x}] + [\mathbf{y}]|^2 &= ([\mathbf{x}] + [\mathbf{y}])^2 = [\mathbf{x}]^2 + [\mathbf{y}]^2 + 2[\mathbf{x}][\mathbf{y}] \\ &\leq |[\mathbf{x}]|^2 + |[\mathbf{y}]|^2 + 2|[\mathbf{x}]||[\mathbf{y}]| = (|[\mathbf{x}]| + |[\mathbf{y}]|)^2 \end{aligned}$$

and so $|[\mathbf{x}] + [\mathbf{y}]| \leq |[\mathbf{x}]| + |[\mathbf{y}]|$. The usual properties of absolute value are obtained.

Now since $[\mathbf{q}] > 0$, we have $q_k > \delta > 0$ for some $\delta \in \mathbb{Q}$ whenever k is large enough. Let n be still larger such that for $k, m \geq n, |q_k - q_m| < \delta/3$. Thus $k \geq n$ implies $q_k - q_n > -\frac{\delta}{3}, q_k > q_n - \frac{\delta}{3} > 2\delta/3$. From Theorem 15.0.7, if $[\mathbf{x}] \in \mathbb{R}$ then there is $r \in \mathbb{Q}$ such that $[|\mathbf{x} - \mathbf{r}|] < \delta/2 < 2\delta/3 < [\mathbf{q}]$. It follows $\mathbb{Q}$ is dense in $\mathbb{R}$.

To show $\mathbb{R}$ is complete, let $\{[\mathbf{x}]^n\}_n$ be a sequence of elements of $\mathbb{R}$. By definition, $[\mathbf{x}]^n = [\mathbf{x}^n]$ where $\mathbf{x}^n$ is a representative of $[\mathbf{x}]^n$. Thus $\mathbf{x}^n$ is in $\mathbf{R}$. Then we are given that for every $\varepsilon > 0, \varepsilon \in \mathbb{Q}$ there is n_ε such that if $m,n > n_\varepsilon$, then

$$|[\mathbf{x}]^n - [\mathbf{x}]^m| \equiv |[\mathbf{x}^n] - [\mathbf{x}^m]| \equiv |[\mathbf{x}^n - \mathbf{x}^m]| = [|\mathbf{x}^n - \mathbf{x}^m|] < \varepsilon$$

We can therefore, obtain a subsequence, still denoted with n such that

$$\left|[\mathbf{x}]^n - [\mathbf{x}]^{n+1}\right| = \left[\left|\mathbf{x}^n - \mathbf{x}^{n+1}\right|\right] < 4^{-n}$$

It will be this subsequence in what follows. Thus there is an increasing sequence $\{k_n\}$ such that $\left|x_k^n - x_l^{n+1}\right| < 2^{-n}$ if $k,l \geq k_n$ where k_n is increasing in n. Let $\mathbf{y} = \left\{x_{k_n}^n\right\}_{n=1}^\infty$. Then $\mathbf{y} \in \mathbf{R}$ and if $m \geq k_n$,

$$\begin{aligned} |x_m^n - y_m| &= \left|x_m^n - x_{k_m}^m\right| \leq \left|x_m^n - x_{k_n}^n\right| + \left|x_{k_n}^n - x_{k_m}^m\right| \\ &< 2^{-n} + \sum_{j=n}^{m-1} \left|x_{k_n}^j - x_{k_m}^{j+1}\right| < 2^{-n} + \sum_{j=n}^{\infty} 2^{-j} = 2^{-n} + 2^{-(n-1)} \end{aligned}$$

Then from the above,

$$|[\mathbf{x}]^n - [\mathbf{y}]| = [|\mathbf{x}^n - \mathbf{y}|] < 2^{-(n-2)}$$

This says that $\lim_{n\to\infty} [\mathbf{x}]^n = [\mathbf{y}]$ by definition. The original Cauchy sequence converges to the same thing thanks to Theorem 4.5.4. ■

It follows that you can consider each real number in $\mathbb{R}$ as an equivalence class of Cauchy sequences. One can show that any two ordered, complete, separable, fields are isomorphic so there is essentially only one of them.

There are other ways to construct the real numbers from the rational numbers. The technique of Dedekind cuts might be a little shorter and easier to understand. However, the above technique of the consideration of equivalence classes of Cauchy sequences can also be used to complete any metric space and this is a common problem. The technique of Dedekind cuts cannot do this because it depends on the order of $\mathbb{Q}$ and there is no order in a general metric space.

A metric space is a nonempty set X on which is defined a distance function (metric) which satisfies the following axioms for $x,y,z \in X$.

$$d(x,y) = d(y,x), \ \infty > d(x,y) \geq 0$$

$$d(x,y) + d(y,z) \geq d(x,z)$$

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

Its completion is a larger metric space with the property that Cauchy sequences converge. It will also consist of equivalence classes of Cauchy sequences. The idea of a Cauchy sequence makes sense in a metric space.

Appendix A

Classification of Real Numbers

Dedekind and Cantor constructed the real numbers in 1872. Then in 1882 and 1884 Lindermann and Weierstrass were able to classify certain important real numbers like logarithms, sines and cosines and π. This was an amazing achievement.

Recall that algebraic numbers are those which are roots of a polynomial with rational or integer coefficients. (Note that if the coefficients are rational, you could simply multiply by the product of the denominators and reduce to one which has all integer coefficients.) This of course includes many complex numbers. For example, x^2+1 has roots $\pm i$. The algebraic numbers include all rational numbers. For example the root of $mx-n=0$ is $\frac{n}{m}$. Numbers which are not algebraic are called transcendental.

Most numbers are transcendental. This follows from Problem 14 on Page 53 and Problem 16 on Page 53. However, it is very difficult to show that a particular number is transcendental. Lindermann and Weierstrass made some progress on this in 1882 and 1884. In particular, Lindermann showed that π is transcendental. This solved the ancient problem about whether one could square the circle. If you start with the unit circle, its area is π and the question was whether you could construct with compass and unmarked straight edge only, a square of area π.

You can't do it because all constructible numbers are algebraic. In fact they all involve roots of quadratic polynomials and linear polynomials which is essentially why you cannot trisect an arbitrary angle either, such as a $60°$ angle. If you could square the circle, then you would end up needing the sides of the square to be $\sqrt{\pi}$ which, if algebraic, would require π to also be algebraic. This is explained below. It turns out that doing algebra to algebraic numbers results in algebraic numbers.

This theorem of Lindermann is a very significant result and it seems to be neglected these days. This is why I am including a treatment of it which I hope will be somewhat understandable. It is very technical however. I have not seen the original proof of this theorem. I suspect it is not what is about to be presented which depends on work of Steinberg and Redheffer dating from 1950. However, the use of the symmetric polynomial theorem used here seems an interesting way to proceed. This symmetric polynomial theorem is very important for its own sake.

I will use the concept of a vector space and a basis for it in what follows. A beginning linear algebra course which is not restricted to row operations should contain sufficient background. However, if you have not seen the notion of an abstract vector space and basis, it would be better to learn this first. It is in any of my books on linear algebra.

A.1 Algebraic Numbers

a is an algebraic number when there is a polynomial $p(x) \equiv x^n + a_{n-1}x^{n-1} + \cdots + a_0$ with each a_k rational having a as a root. Out of all such polynomials, the one which has n as small as possible is called the minimum polynomial for the algebraic number a and n is called the degree of this algebraic number. This minimum polynomial is unique. Indeed, if $p(x)$ and $\hat{p}(x)$ are two such, then by the division algorithm,

$$\hat{p}(x) = p(x)\,q(x) + r(x)$$

where $r(x)$ has smaller degree than n or is 0. The first case cannot happen because $r(a)=0$ and so $r(x)=0$. Now matching coefficients shows that $\hat{p}(x)$ is a multiple of $p(x)$ and so

$q(x)$ can only be a scalar. $q(x)$ must be 1 because both $p(x)$, and $\hat{p}(x)$ have leading coefficient equal to 1.

If $m \geq n$, then by the division algorithm for polynomials, $x^m = p(x)q(x) + r(x)$ where $r(x)$ is either 0 or has degree less than n the degree of $p(x)$. Then $a^m = r(a)$ and so this shows that if $q(x)$ is a polynomial then $q(a)$ can always be written in the form $r(a)$ where either $r(a) = 0$ or $r(x)$ has smaller degree than $p(x)$ the minimum polynomial. Incidentally, this works for linear transformations in place of a for exactly the same reasons. We denote by $\mathbf{k} \equiv \begin{pmatrix} k_1 & \cdots & k_r \end{pmatrix}$ an ordered list of nonnegative integers.

Let $\mathbb{Q}[a_1,...,a_r]$ be all finite sums of the form $\sum_{\mathbf{k}} a_{\mathbf{k}} a_1^{k_1} \cdots a_r^{k_r}$ where the k_i are nonnegative integers, the $a_{\mathbf{k}}$ are rational numbers, and the a_i are nonzero algebraic numbers. Then from what was just observed, this is always of the form $\sum_{\{\mathbf{k} \text{ such that } k_i \leq n_i\}} a_{\mathbf{k}} a_1^{k_1} \cdots a_r^{k_r}$ where the n_i are degrees of the algebraic numbers a_i. Then $\mathbb{Q}[a_1,...,a_r]$ is a vector space over the field of scalars $\mathbb{Q}$ or more generally, you would have $\mathbb{F}[a_1,...,a_r]$ a vector space over the field of scalars $\mathbb{F}$. It follows that $\mathbb{Q}[a_1,...,a_r]$ has a spanning set

$$\left\{a_1^{k_1} \cdots a_r^{k_r}, 0 \leq k_i \leq n_i - 1\right\}$$

Therefore, the dimension of $\mathbb{Q}[a_1,...,a_r]$, as such a vector space, is no more than $\prod_{i=1}^r n_i \equiv m$, the product of the degrees of the algebraic numbers. Letting $g(a_1,...,a_r)$ be a polynomial in which the a_i are algebraic, it follows that with m as just defined,

$$1, g(a_1,...,a_r), g(a_1,...,a_r)^2, ..., g(a_1,...,a_r)^m$$

cannot be linearly independent. There are too many of them. It follows that there exist scalars b_k in $\mathbb{Q}$ or the field of scalars, such that $b_0 + \sum_{k=1}^m b_k g(a_1,...,a_r)^k = 0$ and so, $g(a_1,...,a_r)$ is an algebraic number. This shows that if you have algebraic numbers, you can multiply them, add them, raise them to positive integer powers and so forth, and the result will still be an algebraic number. Sloppily expressed, doing algebra to algebraic numbers yields algebraic numbers.

In fact, you can exploit the existence of a polynomial of minimum degree for which a number in $\mathbb{Q}[a_1,...,a_r]$ is a root, to show that $\mathbb{Q}[a_1,...,a_r]$ is actually a field. However, this is not needed here. It is enough to note that $\mathbb{Q}[a_1,...,a_r]$ is a commutative ring, discussed below. Note how this compares with the previous section about extending $\mathbb{Q}$ to $\mathbb{R}$. This extends $\mathbb{Q}$ to a larger field $\mathbb{Q}[a_1,...,a_r] \subseteq \mathbb{C}$ but not anywhere near all the way to the field $\mathbb{C} = \mathbb{R} + i\mathbb{R}$. In fact, as shown in an early problem the set of algebraic numbers, is countable whereas $\mathbb{R}$ and $\mathbb{C}$ are not. Of course, if you had a which is not a root of any polynomial having rational coefficients, then $\mathbb{Q}[a]$ would not be finite dimensional and the above process will fail. When this occurs, we say that a is transcendental.

A.2 The Symmetric Polynomial Theorem

First here is a definition of polynomials in many variables which have coefficients in a commutative ring. A commutative ring would be a field except it lacks the axiom which gives multiplicative inverses for nonzero elements of the ring. A good example of a commutative ring is the integers. In particular, every field is a commutative ring. Thus, a commutative ring satisfies the following axioms. They are just the field axioms with one omission just mentioned. You might not have x^{-1} if $x \neq 0$. We will assume that the ring has 1, the multiplicative identity.

Axiom A.2.1 *Here are the axioms for a commutative ring.*

1. $x+y=y+x$, *(commutative law for addition)*
2. *There exists* 0 *such that* $x+0=x$ *for all* x, *(additive identity).*
3. *For each* $x\in\mathbb{F}$, *there exists* $-x\in\mathbb{F}$ *such that* $x+(-x)=0$, *(existence of additive inverse).*
4. $(x+y)+z=x+(y+z)$ *,(associative law for addition).*
5. $xy=yx$, *(commutative law for multiplication). You could write this as* $x\times y=y\times x$*.*
6. $(xy)z=x(yz)$ *,(associative law for multiplication).*
7. *There exists* 1 *such that* $1x=x$ *for all* x*,(multiplicative identity).*
8. $x(y+z)=xy+xz$*.(distributive law).*

The example of most interest here is where the commutative ring is the integers $\mathbb{Z}$ or $\mathbb{Q}[a_1,...,a_r]$. Next is a definition of what is meant by a polynomial.

Definition A.2.2 *Let* $\mathbf{k}\equiv(k_1,k_2,\cdots,k_n)$ *where each* k_i *is a nonnegative integer. Let* $|\mathbf{k}|\equiv\sum_i k_i$*. Polynomials of degree* p *in the variables* $x_1,x_2,\cdots,x_n$ *are expressions of the form*

$$g(x_1,x_2,\cdots,x_n)=\sum_{|\mathbf{k}|\leq p}a_{\mathbf{k}}x_1^{k_1}\cdots x_n^{k_n}$$

where each $a_{\mathbf{k}}$ *is in a commutative ring. If all* $a_{\mathbf{k}}=0$*, the polynomial has no degree. Such a polynomial is said to be symmetric if whenever* σ *is a permutation of* $\{1,2,\cdots,n\}$*,*

$$g\left(x_{\sigma(1)},x_{\sigma(2)},\cdots,x_{\sigma(n)}\right)=g(x_1,x_2,\cdots,x_n)$$

An example of a symmetric polynomial is $s_1(x_1,x_2,\cdots,x_n)\equiv\sum_{i=1}^n x_i$. Another one is $s_n(x_1,x_2,\cdots,x_n)\equiv x_1x_2\cdots x_n$.

Definition A.2.3 *The elementary symmetric polynomial*

$$s_k(x_1,x_2,\cdots,x_n),k=1,\cdots,n$$

is the coefficient of $(-1)^k x^{n-k}$ *in the following polynomial.*

$$(x-x_1)(x-x_2)\cdots(x-x_n)=x^n-s_1x^{n-1}+s_2x^{n-2}-\cdots\pm s_n$$

Thus

$$s_1=x_1+x_2+\cdots+x_n$$

$$s_2=\sum_{i<j}x_ix_j,\ s_3=\sum_{i<j<k}x_ix_jx_k,\ldots,s_m=\sum_{i_1<i_2\cdots<i_m}x_{i_1}x_{i_2}\cdots x_{i_m},\ s_n=x_1x_2\cdots x_n$$

These special elementary polynomials are symmetric because switching two of the variables x_i and x_j is equivalent to switching the corresponding factors in the product $(x-x_1)(x-x_2)\cdots(x-x_n)$ and using the same process to collect terms which multiply x^{n-k}. The polynomial in x does not change.

Example A.2.4

$$\begin{aligned}&(x-x_1)(x-x_2)(x-x_3)\\ =\;& x^3-x^2(x_1+x_2+x_3)+x(x_1x_2+x_1x_3+x_2x_3)-x_1x_2x_3.\end{aligned}$$

Thus the symmetric polynomials are $x_1+x_2+x_3, x_1x_2+x_1x_3+x_2x_3$, *and* $x_1x_2x_3$.

Note that it follows from the above definition that

$$\alpha^k s_k(x_1,x_2,\cdots,x_n)=s_k(\alpha x_1,\cdots,\alpha x_n)$$

Then the following result is the fundamental theorem in the subject. It is the symmetric polynomial theorem. This is a very remarkable theorem.

Theorem A.2.5 *Let* $g(x_1,x_2,\cdots,x_n)$ *be a symmetric polynomial. Then this symmetric polynomial* $g(x_1,x_2,\cdots,x_n)$ *equals a polynomial in the elementary symmetric polynomials.*

$$g(x_1,x_2,\cdots,x_n)=\sum_{\mathbf{k}} a_{\mathbf{k}} s_1^{k_1}\cdots s_n^{k_n}$$

and the $a_{\mathbf{k}}$ *in the commutative ring are unique.*

Proof: The proof is by induction on the number of variables. If $n=1$, it is obviously true because $s_1=x_1$ and $g(x_1)$ can only be a polynomial in x_1. Suppose the theorem is true for $n-1$ variables and $g(x_1,x_2,\cdots,x_n)$ has degree d. Thus in the sum for the polynomial, $|\mathbf{k}|\leq d$. By induction, there is a polynomial

$$Q(\tilde{s}_1,\cdots,\tilde{s}_{n-1})=\sum_{|\mathbf{k}|\leq p} a_{\mathbf{k}}\tilde{s}_1^{k_1}\cdots\tilde{s}_{n-1}^{k_{n-1}}=g(x_1,x_2,\cdots,x_{n-1},0) \tag{1.1}$$

where $\tilde{s}_k$ is a symmetric polynomial for the variables $\{x_1,x_2,\cdots,x_{n-1}\}$. Now let

$$p(x_1,x_2,\cdots,x_n)\equiv g(x_1,x_2,\cdots,x_n)-Q(s_1,\cdots,s_{n-1}) \tag{1.2}$$

Thus $p(x_1,x_2,\cdots,x_n)$ is a symmetric polynomial because each s_j is symmetric and g is given to be symmetric. Notice how $\tilde{s}_k$ was replaced with s_k.

If x_n is set equal to 0, the right side reduces to 0 because $s_k(x_1,x_2,\cdots,x_{n-1},0)=\tilde{s}_k(x_1,x_2,\cdots,x_{n-1})$. This follows from the definition of these symmetric polynomials. Indeed, the coefficient of x^{n-k} in $(x-x_1)(x-x_2)\cdots(x-x_{n-1})(x-0)$ is the same as the coefficient of $x^{(n-1)-k}$ in $(x-x_1)(x-x_2)\cdots(x-x_{n-1})$. Thus, the right side of 1.2 reduces to $g(x_1,x_2,\cdots,x_{n-1},0)-Q(\tilde{s}_1,\cdots,\tilde{s}_{n-1})=0$ from 1.1 when $x_n=0$.

Thus x_n divides $p(x_1,x_2,\cdots,x_n)$ so every term in $p(x_1,x_2,\cdots,x_n)$ has a factor of x_n. The same must be true with x_j since otherwise, the symmetric polynomial $p(x_1,x_2,\cdots,x_n)$ would change if you switched x_j and x_n. Hence there exists a symmetric polynomial $g_1(x_1,x_2,\cdots,x_n)$ such that

$$s_n g_1(x_1,x_2,\cdots,x_n)=g(x_1,x_2,\cdots,x_n)-Q(s_1,\cdots,s_{n-1})$$

Recall $s_n=x_1x_2\cdots x_n$. Thus

$$g(x_1,x_2,\cdots,x_n)=s_n g_1(x_1,x_2,\cdots,x_n)+Q(s_1,\cdots,s_{n-1})$$

Now if g_1 is not constant, do for g_1 what was just done for g. Obtain

$$g(x_1, x_2, \cdots, x_n) = s_n \begin{pmatrix} s_n g_2(x_1, x_2, \cdots, x_n) \\ +Q_2(s_1, \cdots, s_{n-1}) \end{pmatrix} + Q(s_1, \cdots, s_{n-1})$$

Continue this way, obtaining a sequence of g_k till the process stops with some g_m being a constant. This must happen because the degree of g_k becomes strictly smaller with each iteration. This yields a polynomial in the elementary symmetric polynomials for $\{x_1, x_2, \cdots, x_n\}$. ■

Example A.2.6 *Let $g(x,y) = x^3 + y^3$. It is clear that $g(x,y) = g(y,x)$ so g is a symmetric polynomial. Write as a polynomial in the elementary functions.*

The above proof tells how to do this. First note that $x^3 = \bar{s}_1^3$ where s_1 is the symmetric polynomial associated with the single variable x. Thus $p(x,y) = x^3 + y^3 - s_1^3$ where this s_1 is $x + y$. Then $p(x,y) = x^3 + y^3 - (x+y)^3 = -3x^2y - 3xy^2$ and this equals $(-xy)(3x+3y) = -3s_2s_1$. Thus $-3s_1s_2 = x^3 + y^3 - s_1^3$ and so $g(x,y) = s_1^3 - 3s_1s_2$.

You can see that if you have a symmetric polynomial in more variables, you could use a process of reducing one variable at a time in $g(x_1, ..., x_{n-1}, 0)$ to eventually obtain this function as a polynomial in the symmetric polynomials in variables $\{x_1, ..., x_{n-1}\}$.

Note that if you have $\prod_{j=1}^m (x - x_j)$ then by definition, it is the sum of terms like $g(x_1, \cdots, x_m) x^{m-k}$. If you replace x with x_i and sum over all i, you would obtain an expression of the form $\sum_{i=1}^m g(x_1, \cdots, x_m) x_i^{m-k}$ which would also be a symmetric polynomial. It is of the form

$$g(x_1, \cdots, x_m) x_1^{m-k} + g(x_1, \cdots, x_m) x_2^{m-k} + \cdots + g(x_1, \cdots, x_m) x_m^{m-k}$$

so when you switch some variables in this, you get the same thing.

Here is a very interesting result which I saw claimed in a paper by Steinberg and Redheffer on Lindermannn's theorem which follows from the above theorem. It is a very useful property of symmetric polynomials and is the main tool for proving the Lindermann Weierstrass theorem.

Theorem A.2.7 *Let $\alpha_1, \cdots, \alpha_n$ be roots of the polynomial equation*

$$p(x) \equiv a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \qquad (*)$$

*where each a_i is an integer. Then **any** symmetric polynomial in the quantities*

$$a_n \alpha_1, \cdots, a_n \alpha_n$$

having integer coefficients is also an integer. Also any symmetric polynomial with rational coefficients in the quantities $\alpha_1, \cdots, \alpha_n$ is a rational number.

Proof: Let $f(x_1, \cdots, x_n)$ be the symmetric polynomial having integer coefficients. From Theorem A.2.5 it follows there are integers $a_{k_1 \cdots k_n}$ such that

$$f(x_1, \cdots, x_n) = \sum_{k_1 + \cdots + k_n \leq m} a_{k_1 \cdots k_n} p_1^{k_1} \cdots p_n^{k_n} \qquad (1.3)$$

where the p_i are elementary symmetric polynomials defined as the coefficients of $\hat{p}(x) = \prod_{j=1}^{n}(x - x_j)$ with $p_k(x_1, ..., x_n)$ of degree k since it is the coefficient of x^{n-k}. Earlier we had them $\pm$ these coefficients. Thus

$$f(a_n\alpha_1, \cdots, a_n\alpha_n) = \sum_{k_1+\cdots+k_n=d} a_{k_1\cdots k_n} p_1^{k_1}(a_n\alpha_1, \cdots, a_n\alpha_n) \cdots p_n^{k_n}(a_n\alpha_1, \cdots, a_n\alpha_n)$$

Now the given polynomial in $*$, $p(x)$ is of the form

$$a_n \prod_{j=1}^{n}(x - \alpha_j) \equiv a_n \left(\sum_{k=0}^{n} p_k(\alpha_1, \cdots, \alpha_n) x^{n-k} \right)$$
$$= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

Thus, equating coefficients, $a_n p_k(\alpha_1, \cdots, \alpha_n) = a_{n-k}$. Multiply both sides by a_n^{k-1}. Thus

$$p_k(a_n\alpha_1, \cdots, a_n\alpha_n) = a_n^{k-1} a_{n-k}$$

an integer. Therefore,

$$f(a_n\alpha_1, \cdots, a_n\alpha_n) = \sum_{k_1+\cdots+k_n=d} a_{k_1\cdots k_n} p_1^{k_1}(a_n\alpha_1, \cdots, a_n\alpha_n) \cdots p_n^{k_n}(a_n\alpha_1, \cdots, a_n\alpha_n)$$

and each $p_k(a_n\alpha_1, \cdots, a_n\alpha_n)$ is an integer. Thus $f(a_n\alpha_1, \cdots, a_n\alpha_n)$ is indeed an integer. From this, it is obvious that $f(\alpha_1, \cdots, \alpha_n)$ is rational. Indeed, from 1.3,

$$f(\alpha_1, \cdots, \alpha_n) = \sum_{k_1+\cdots+k_n=d} a_{k_1\cdots k_n} p_1^{k_1}(\alpha_1, \cdots, \alpha_n) \cdots p_n^{k_n}(\alpha_1, \cdots, \alpha_n)$$

Now multiply both sides by a_n^M, an integer where M is chosen large enough that

$$\begin{aligned} & a_n^M f(\alpha_1, \cdots, \alpha_n) \\ = & \sum_{k_1+\cdots+k_n=d} a_n^{h(k_1,\ldots,k_n)} a_{k_1\cdots k_n} p_1^{k_1}(a_n\alpha_1, \cdots, a_n\alpha_n) \cdots p_n^{k_n}(a_n\alpha_1, \cdots, a_n\alpha_n) \end{aligned}$$

where $h(k_1, ..., k_n)$ is some nonnegative integer. Then the right side is an integer. Thus $f(\alpha_1, \cdots, \alpha_n)$ is rational. If the f had rational coefficients, then mf would have integer coefficients for a suitable m and so $mf(\alpha_1, \cdots, \alpha_n)$ would be rational which yields $f(\alpha_1, \cdots, \alpha_n)$ is rational. ■

A.3 Transcendental Numbers

Most numbers are like this, transcendental. Here the algebraic numbers are those which are roots of a polynomial equation having rational numbers as coefficients, equivalently integer coefficients. By the fundamental theorem of algebra, all these numbers are in $\mathbb{C}$ and they constitute a countable collection of numbers in $\mathbb{C}$. Therefore, most numbers in $\mathbb{C}$ are transcendental. Nevertheless, it is very hard to prove that a particular number is transcendental. Probably the most famous theorem about this is the Lindermannn Weierstrass theorem, 1884.

Theorem A.3.1 *Let the α_i be distinct nonzero algebraic numbers and let the a_i be nonzero algebraic numbers. Then $\sum_{i=1}^{n} a_i e^{\alpha_i} \neq 0$.*

I am following the interesting Wikepedia article on this subject. You can also look at the book by Baker [5], Transcendental Number Theory, Cambridge University Press. There are also many other treatments which you can find on the web including an interesting article by Steinberg and Redheffer which appeared in about 1950.

The proof makes use of the following identity. For $f(x)$ a polynomial,

$$I(s) \equiv \int_0^s e^{s-x} f(x)\, dx = e^s \sum_{j=0}^{\deg(f)} f^{(j)}(0) - \sum_{j=0}^{\deg(f)} f^{(j)}(s). \tag{1.4}$$

where $f^{(j)}$ denotes the j^{th} derivative. It is like the convolution integral discussed earlier with Laplace transforms. In this formula, $s \in \mathbb{C}$ and the integral is defined in the natural way as

$$\int_0^1 s f(ts)\, e^{s-ts} dt \tag{1.5}$$

The identity follows from integration by parts.

$$\int_0^1 s f(ts)\, e^{s-ts} dt = s e^s \int_0^1 f(ts)\, e^{-ts} dt = s e^s \left[-\frac{e^{-ts}}{s} f(ts)\, |_0^1 + \int_0^1 \frac{e^{-ts}}{s} s f'(st)\, dt \right]$$

$$= s e^s \left[-\frac{e^{-s}}{s} f(s) + \frac{1}{s} f(0) + \int_0^1 e^{-ts} f'(st)\, dt \right] = e^s f(0) - f(s) + \int_0^1 s e^{s-ts} f'(st)\, dt$$

$$\equiv e^s f(0) - f(s) + \int_0^s e^{s-x} f'(x)\, dx$$

Continuing this way establishes the identity since the right end looks just like what we started with except with a derivative on the f.

Lemma A.3.2 *Let $(x_1, \ldots, x_n) \to g(x, x_1, \ldots, x_n)$ be symmetric and let*

$$x \to g(x, x_1, \ldots, x_n)$$

be a polynomial. Then $\frac{d^m}{dx^m} g(x, x_1, \ldots, x_n)$ is symmetric in the variables $\{x_1, \ldots, x_n\}$. If $(x_1, \ldots, x_n) \to h(x, x_1, \ldots, x_n)$ is symmetric, then for r some nonnegative integer, it follows that $\sum_{k=1}^n h(x_k, x_1, \ldots, x_n) x_k^r$ is symmetric. In particular, $\sum_{k=1}^n \frac{d^l}{dx^l} g(\cdot, x_1, \ldots, x_n)(x_k) x_k^r$ is symmetric in $\{x_1, \ldots, x_n\}$.

Proof: The coefficients of the polynomial $x \to g(x, x_1, \ldots, x_n)$ are symmetric functions of $\{x_1, \ldots, x_n\}$. Differentiating with respect to x multiple times just gives another polynomial in x having coefficients which are symmetric functions. Thus the first part is proved. For the second part, the sum is of the form

$$h(x_1, x_1, \ldots, x_n) x_1^r + h(x_2, x_1, \ldots, x_n) x_2^r + \cdots + h(x_n, x_1, \ldots, x_n) x_n^r$$

You see that this is unchanged from switching two variables. For example, switch x_1 and x_2. By assumption, nothing changes in the terms after the first two. The first term then becomes

$$h(x_2, x_2, x_1 \ldots, x_n) x_2^r = h(x_2, x_1, x_2, \ldots, x_n) x_2^r$$

and the second term becomes

$$h(x_1, x_2, x_1, \ldots, x_n) x_1^r = h(x_1, x_1, x_2, \ldots, x_n) x_1^r$$

which are the same two terms, just added in a different order. The situation works the same way with any other pair of variables. ■

Recall that every algebraic number is a root of a polynomial having integer coefficients.

Lemma A.3.3 *Let $Q(x) = vx^m + \cdots + u$ have integer coefficients with roots $\beta_1, ..., \beta_m$ listed according to multiplicity. Let*

$$f(x) \equiv \frac{v^{(m+1)p} Q^p(x) x^{p-1}}{(p-1)!} \tag{1.6}$$

a polynomial of degree $n = pm + p - 1$. Then

$$\sum_{j=0}^{n} f^{(j)}(0) = v^{p(m+1)} u^p + m_1(p) p \tag{1.7}$$

$$\sum_{i=1}^{m} \sum_{j=0}^{n} f^{(j)}(\beta_i) = m_2(p) p \tag{1.8}$$

where $m_1(p), m_2(p)$ are integers and p will be a large prime.

Proof: First consider 1.7. $f(x) = \frac{v^{(m+1)p}(vx^m+\cdots+u)^p x^{p-1}}{(p-1)!}$. Then $f^j(0) = 0$ unless $j \geq p-1$ because otherwise, that x^{p-1} term will result in some $x^r, r > 0$ and everything is zero when you plug in $x = 0$. Now say $j = p-1$. Then it is clear that you get a $(p-1)!$ which cancels the denominator and letting $x = 0$, you get the integer $f^{(p-1)}(0) = u^p v^{(m+1)p}$. So what if $j > p-1$?

$$\begin{aligned}
&\frac{d^j}{dx^j}\left((vx^m + \cdots + u)^p x^{p-1}\right) \\
= &\sum_{r=0}^{j} \binom{j}{i} \frac{d^i}{dx^i}\left((vx^m + \cdots + u)^p\right) \frac{d^{j-i}}{dx^{j-i}} x^{p-1}
\end{aligned}$$

and, since eventually $x = 0$, only $j - i = p - 1$ is of interest, so $i = j - p + 1$ where $j \geq p$ as just mentioned. Since $i \geq 1$, there will be a factor of p and a factor of $(p-1)!$ from $\frac{d^{j-i}}{dx^{j-i}} x^{p-1}$. Thus when $x = 0$, this reduces to $m_1(p) p (p-1)!$ and so this yields 1.7.

Next consider 1.8 which says that $\sum_{i=1}^{m} \sum_{j=0}^{n} f^{(j)}(\beta_i) = m_2(p) p$. The factorization of $Q(x)$ is $v(x-\beta_1)\cdots(x-\beta_m)$. Replace $Q(x)$ with its factorization in 1.6 to get

$$f(x)(p-1)! = v^p v^{(m+1)p} \left((x-\beta_1)(x-\beta_2)\cdots(x-\beta_m)\right)^p x^{p-1} \tag{1.9}$$

First notice that $(p-1)! f^{(j)}(\beta_i) = 0$ unless $j \geq p$. Thus all terms in computing

$$f^{(j)}(\beta_i)(p-1)!$$

for $j \geq p$ have a factor of $p!$. If you have

$$g(x, \beta_1, \cdots, \beta_m) \equiv v^p v^{(m+1)p} \left((x-\beta_1)(x-\beta_2)\cdots(x-\beta_m)\right)^p x^{p-1},$$

it is symmetric in the β_i so all derivatives with respect to x are also symmetric in these β_i by Lemma A.3.2. By the same lemma, for $j \geq p$

$$\sum_{i=1}^{m} \frac{d^j}{dx^j}\left(g(\cdot, \beta_1, \cdots, \beta_m)(\beta_i) \frac{1}{(p-1)!}\right) = \sum_{i=1}^{m} f^{(j)}(\beta_i)$$

is symmetric in the $\beta_1, \cdots, \beta_m$. Thanks to the factor $v^p v^{(m+1)p}$ and the factor $p!$ coming from $j \geq p$, it is a symmetric polynomial in the $v\beta_i$ with integer coefficients, each multiplied by p with the β_i roots of $Q(x) = vx^m + \cdots + u$. By Theorem A.2.7 this is an integer. As noted earlier, it equals 0 unless $j \geq p$ when it contains a factor of p. Thus the sum of these integers is also an integer times p. It follows that

$$\sum_{i=1}^{m} \sum_{j=0}^{n} f^{(j)}(\beta_i) = m_2(p)\, p,\ m_2(p) \text{ an integer. } \blacksquare$$

Note that no use was made of p being a large prime number. This will come next.

Lemma A.3.4 *If K and c are nonzero integers, and $\beta_1, \cdots, \beta_m$ are the roots of a single polynomial with integer coefficients,*

$$Q(x) = vx^m + \cdots + u$$

where $v, u \neq 0$, then,

$$K + c\left(e^{\beta_1} + \cdots + e^{\beta_m}\right) \neq 0.$$

Letting

$$f(x) \equiv \frac{v^{(m+1)p} Q^p(x)\, x^{p-1}}{(p-1)!}$$

and $I(s)$ be defined in terms of $f(x)$ as above,

$$I(s) \equiv \int_0^s e^{s-x} f(x)\, dx = e^s \sum_{j=0}^{\deg(f)} f^{(j)}(0) - \sum_{j=0}^{\deg(f)} f^{(j)}(s),$$

it follows,

$$\lim_{p \to \infty} \sum_{i=1}^{m} I(\beta_i) = 0 \tag{1.10}$$

and for n the degree of $f(x)$, $n = pm + p - 1$, where $m_i(p)$ is some integer for p a large prime number.

Proof: The first step is to verify 1.10 for $f(x)$ as given in 1.6 for p large prime numbers. Let p be a large prime number. Then 1.10 follows right away from the definition of $I\left(\beta_j\right)$ and the definition of $f(x)$.

$$\left|I\left(\beta_j\right)\right| \leq \int_0^1 \left|\beta_j f\left(t\beta_j\right) e^{\beta_j - t\beta_j}\right| dt \leq \int_0^1 \left|\frac{|v|^{(m-1)p} \left|Q\left(t\beta_j\right)\right|^p t^{p-1} \left|\beta_j\right|^{p-1}}{(p-1)!} dt\right|$$

which clearly converges to 0 using considerations involving convergent series which show the integrand converges uniformly to 0. The degree of $f(x)$ is $n \equiv pm + p - 1$ where p will be a sufficiently large prime number from now on.

From 1.4,

$$c \sum_{i=1}^{m} I(\beta_i) = c \sum_{i=1}^{m} \left(e^{\beta_i} \sum_{j=0}^{n} f^{(j)}(0) - \sum_{j=0}^{n} f^{(j)}(\beta_i)\right)$$

$$= \left(K + c \sum_{i=1}^{m} e^{\beta_i} \right) \sum_{j=0}^{n} f^{(j)}(0) - \left(K \sum_{j=0}^{n} f^{(j)}(0) + c \sum_{i=1}^{m} \sum_{j=0}^{n} f^{(j)}(\beta_i) \right) \tag{1.11}$$

Here $K \sum_{j=0}^{n} f^{(j)}(0)$ is added and subtracted. From Lemma A.3.3,

$$v^{p(m+1)} u^p + m_1(p) p + m_2(p) p = K \sum_{j=0}^{n} f^{(j)}(0) + c \sum_{i=1}^{m} \sum_{j=0}^{n} f^{(j)}(\beta_i)$$

Thus, if p is very large,

$$c \sum_{i=1}^{m} I(\beta_i) = \text{small} = K v^{p(m+1)} u^p + M(p) p + \left(K + c \sum_{i=1}^{m} e^{\beta_i} \right) \sum_{j=0}^{n} f^{(j)}(0)$$

Let p be prime and larger than $\max(K, v, u)$. If $K + c \sum_{i=1}^{m} e^{\beta_i} = 0$, the above is impossible because it would require

$$\text{small} = K v^{p(m+1)} u^p + M(p) p$$

Now the right side is a nonzero integer because p cannot divide $K v^{p(m+1)} u^p$ so the right side cannot equal something small. ■

Note that this shows π is irrational. If $\pi = k/m$ where k, m are integers, then both $i\pi$ and $-i\pi$ are roots of the polynomial with integer coefficients, $m^2 x^2 + k^2$ which would require, from what was just shown that $0 \neq 2 + e^{i\pi} + e^{-i\pi}$ which is not the case since the sum on the right equals 0.

The following corollary follows from this. It is like the above lemma except it involves several polynomials. First is a lemma.

Lemma A.3.5 *Let v_k, u_k, m_k be integers for $k = 1, 2..., m, u_k, v_k$ nonzero. Then for each k there exists α_k an integer such that $\alpha_k^{m_k+2} v_k^{m_k+1} u_k$ is U for some non zero integer.*

Proof: Let $U \equiv \left(\prod_{j=1}^{m} v_j^{m_j+1} u_j \right)^{\prod_{j=1}^{m} (m_j+2)^2} \equiv \alpha_k^{m_k+2} v_k^{m_k+1} u_k$ where α_k is an integer chosen to make this so.■

Corollary A.3.6 *Let K and c_i for $i = 1, \cdots, n$ be nonzero integers. For each k between 1 and n let $\{\beta(k)_i\}_{i=1}^{m_k}$ be the roots of a polynomial with integer coefficients,*

$$Q_k(x) \equiv v_k x^{m_k} + \cdots + u_k$$

where $v_k, u_k \neq 0$. Then

$$K + c_1 \left(\sum_{j=1}^{m_1} e^{\beta(1)_j} \right) + c_2 \left(\sum_{j=1}^{m_2} e^{\beta(2)_j} \right) + \cdots + c_n \left(\sum_{j=1}^{m_n} e^{\beta(n)_j} \right) \neq 0. \tag{$*$}$$

Proof: Let K_k be nonzero integers which add to K. It is certainly possible to obtain this since the K_k are allowed to change sign. They only need to be nonzero. Also let α_k be as in the above lemma such that $\alpha_k^{m_k+2} v_k^{m_k+1} u_k = U$ some integer. Thus, replacing each $Q_k(x)$ with $\alpha_k v_k x^{m_k} + \cdots + \alpha_k u_k$, it follows that for each large prime $p, (\alpha_k v)^{p(m_k+1)} (\alpha_k u)^p = \left(\alpha_k^{m_k+2} v^{m_k+1} \right)^p = U^p$. From now on, use the new $Q_k(x)$.

Defining $f_k(x)$ and $I_k(s)$ as in Lemma A.3.4,

$$f_k(x) \equiv \frac{v^{(m+1)p} Q_k^p(x) x^{p-1}}{(p-1)!}$$

and as before, let p be a very large prime number. It follows from Lemma A.3.4 that for each $k = 1, \cdots, n$,

$$\begin{aligned} c_k \sum_{i=1}^{m_k} I_k(\beta(k)_i) &= \left(K_k + c_k \sum_{i=1}^{m_k} e^{\beta(k)_i}\right) \sum_{j=0}^{\deg(f_k)} f_k^{(j)}(0) \\ &\quad - \left(K_k \sum_{j=0}^{\deg(f_k)} f_k^{(j)}(0) + c_k \sum_{i=1}^{m_k} \sum_{j=0}^{\deg(f_k)} f_k^{(j)}(\beta(k)_i)\right) \end{aligned}$$

This is exactly the same computation as in the beginning of that lemma except one adds and subtracts $K_k \sum_{j=0}^{\deg(f_k)} f_k^{(j)}(0)$ rather than $K \sum_{j=0}^{\deg(f_k)} f_k^{(j)}(0)$ where the K_k are chosen such that their sum equals K and the term on the left converges to 0 as $p \to \infty$. By Lemma A.3.4,

$$c_k \sum_{i=1}^{m_k} I_k(\beta(k)_i) = \left(K_k + c_k \sum_{i=1}^{m_k} e^{\beta(k)_i}\right)(U^p + N_k p)$$

$$-K_k(U^p + N_k p) - c_k N_k' p$$

$$= \left(K_k + c_k \sum_{i=1}^{m_k} e^{\beta(k)_i}\right) U^p - KU^p + M_k p$$

where M_k is some integer. Now add.

$$\sum_{k=1}^{m} c_k \sum_{i=1}^{m_k} I_k(\beta(k)_i) = U^p \left(K + \sum_{k=1}^{m} c_k \sum_{i=1}^{m_k} e^{\beta(k)_i}\right) - KmU^p + Mp$$

If $K + \sum_{k=1}^{m} c_k \sum_{i=1}^{m_k} e^{\beta(k)_i} = 0$, then if $p > \max(K, m, U)$ you would have $-KmU^p + Mp$ an integer so it cannot equal the left side which will be small if p is large. Therefore, $*$ follows. ■

Next is an even more interesting Lemma which follows from the above corollary.

Lemma A.3.7 *If $b_0, b_1, \cdots, b_n$ are non zero integers, and $\gamma_1, \cdots, \gamma_n$ are distinct algebraic numbers, then*

$$b_0 e^{\gamma_0} + b_1 e^{\gamma_1} + \cdots + b_n e^{\gamma_n} \neq 0$$

Proof: Assume

$$b_0 e^{\gamma_0} + b_1 e^{\gamma_1} + \cdots + b_n e^{\gamma_n} = 0 \tag{1.12}$$

Divide by e^{γ_0} and letting $K = b_0$,

$$K + b_1 e^{\alpha(1)} + \cdots + b_n e^{\alpha(n)} = 0 \tag{1.13}$$

where $\alpha(k) = \gamma_k - \gamma_0$. These are still distinct algebraic numbers. Therefore, $\alpha(k)$ is a root of a polynomial

$$Q_k(x) = v_k x^{m_k} + \cdots + u_k \tag{1.14}$$

having integer coefficients, $v_k, u_k \neq 0$. Recall algebraic numbers were defined as roots of polynomial equations having rational coefficients. Just multiply by the denominators to get one with integer coefficients. Let the roots of this polynomial equation be

$$\left\{\alpha(k)_1, \cdots, \alpha(k)_{m_k}\right\}$$

and suppose they are listed in such a way that $\alpha(k)_1 = \alpha(k)$. Thus, by Theorem A.2.7 every symmetric polynomial in these roots is rational.

Letting i_k be an integer in $\{1, \cdots, m_k\}$ it follows from the assumption 1.12 that

$$\prod_{\substack{(i_1, \cdots, i_n) \\ i_k \in \{1, \cdots, m_k\}}} \left(K + b_1 e^{\alpha(1)_{i_1}} + b_2 e^{\alpha(2)_{i_2}} + \cdots + b_n e^{\alpha(n)_{i_n}}\right) = 0 \tag{1.15}$$

This is because one of the factors is the one occurring in 1.13 when $i_k = 1$ for every k. The product is taken over all distinct ordered lists $(i_1, \cdots, i_n)$ where i_k is as indicated. Expand this possibly huge product. This will yield something like the following.

$$\begin{aligned} & K' + c_1\left(e^{\beta(1)_1} + \cdots + e^{\beta(1)_{\mu(1)}}\right) \\ & + c_2\left(e^{\beta(2)_1} + \cdots + e^{\beta(2)_{\mu(2)}}\right) + \cdots + \\ & c_N\left(e^{\beta(N)_1} + \cdots + e^{\beta(N)_{\mu(N)}}\right) = 0 \end{aligned} \tag{1.16}$$

These integers c_j come from products of the b_i and K. You group these exponentials according to which c_i they multiply. The $\beta(i)_j$ are the distinct exponents which result, each being a sum of some of the $\alpha(r)_{i_r}$. Since the product included all roots for each $Q_k(x)$, interchanging their order does not change the distinct exponents $\beta(i)_j$ which result. They might occur in a different order however, but you would still have the same distinct exponents associated with each c_s as shown in the sum. Thus any symmetric polynomial in the $\beta(s)_1, \beta(s)_2, \cdots, \beta(s)_{\mu(s)}$ is also a symmetric polynomial in the roots of $Q_k(x)$, $\alpha(k)_1, \alpha(k)_2, \cdots, \alpha(k)_{m_k}$ for each k.

Doesn't this contradict Corollary A.3.6? This is not yet clear because we don't know that the $\beta(i)_1, ..., \beta(i)_{\mu(i)}$ are roots of a polynomial having rational coefficients. For a given $r, \beta(r)_1, \cdots, \beta(r)_{\mu(r)}$ are roots of the polynomial

$$(x - \beta(r)_1)(x - \beta(r)_2) \cdots \left(x - \beta(r)_{\mu(r)}\right) \tag{1.17}$$

the coefficients of which are elementary symmetric polynomials in the $\beta(r)_i, i \leq \mu(r)$. Thus the coefficients are symmetric polynomials in the

$$\alpha(k)_1, \alpha(k)_2, \cdots, \alpha(k)_{m_k}$$

for each k. Say the polynomial is of the form

$$\sum_{l=0}^{\mu(r)} x^{n-l} B_l(A(1), \cdots, A(n))$$

where $A(k)$ signifies the roots of $Q_k(x)$, $\left\{\alpha(k)_1, \cdots, \alpha(k)_{m_k}\right\}$. Thus, by the symmetric polynomial theorem applied to the commutative ring $\mathbb{Q}[A(1), \cdots, A(n-1)]$, the above polynomial is of the form

$$\sum_{l=0}^{\mu(r)} x^{\mu(r)-l} \sum_{\mathbf{k}_l} B_{\mathbf{k}_l}(A(1), \cdots, A(n-1)) s_1^{k_1^l} \cdots s_{\mu(r)}^{k_n^l}$$

where the s_k is one of the elementary symmetric polynomials in

$$\left\{\alpha(n)_1, \cdots, \alpha(n)_{m_n}\right\}$$

and $B_{\mathbf{k}_l}$ is symmetric in $\alpha(k)_1, \alpha(k)_2, \cdots, \alpha(k)_{m_k}$ for each $k \leq n-1$ and

$$B_{\mathbf{k}_l} \in \mathbb{Q}[A(1), \cdots, A(n-1)].$$

Now do to $B_{\mathbf{k}_l}$ what was just done to B_l featuring $A(n-1)$ this time, and continue till eventually you obtain for the coefficient of $x^{\mu(r)-l}$ a large sum of rational numbers times a product of symmetric polynomials in $A(1), A(2)$, etc. By Theorem A.2.7 applied repeatedly, beginning with $A(1)$ and then to $A(2)$ and so forth, one finds that the coefficient of $x^{\mu(r)-l}$ is a rational number and so the $\beta(r)_j$ for $j \leq \mu(r)$ are algebraic numbers and roots of a polynomial which has rational coefficients, namely the one in 1.17, hence roots of a polynomial with integer coefficients. Now 1.16 contradicts Corollary A.3.6. ■

Note this lemma is sufficient to prove Lindermann's theorem that π is transcendental. Here is why. If π is algebraic, then so is $i\pi$ and so from this lemma, $e^0 + e^{i\pi} \neq 0$ but this is not the case because $e^{i\pi} = -1$.

The next theorem is the main result, the Lindermann Weierstrass theorem. It replaces the integers b_i in the above lemma with algebraic numbers.

Theorem A.3.8 *Suppose $a(1), \cdots, a(n)$ are nonzero algebraic numbers and suppose*

$$\alpha(1), \cdots, \alpha(n)$$

are distinct algebraic numbers. Then

$$a(1) e^{\alpha(1)} + a(2) e^{\alpha(2)} + \cdots + a(n) e^{\alpha(n)} \neq 0$$

Proof: Suppose $a(j) \equiv a(j)_1$ is a root of the polynomial

$$v_j x^{m_j} + \cdots + u_j$$

where $v_j, u_j \neq 0$. Let the roots of this polynomial be $a(j)_1, \cdots, a(j)_{m_j}$. Suppose to the contrary that

$$a(1)_1 e^{\alpha(1)} + a(2)_1 e^{\alpha(2)} + \cdots + a(n)_1 e^{\alpha(n)} = 0$$

Then consider the big product

$$\prod_{\substack{(i_1, \cdots, i_n) \\ i_k \in \{1, \cdots, m_k\}}} \left(a(1)_{i_1} e^{\alpha(1)} + a(2)_{i_2} e^{\alpha(2)} + \cdots + a(n)_{i_n} e^{\alpha(n)}\right) \tag{1.18}$$

the product taken over all ordered lists $(i_1, \cdots, i_n)$. Since one of the factors in this product equals 0, this product equals

$$0 = b_1 e^{\beta(1)} + b_2 e^{\beta(2)} + \cdots + b_N e^{\beta(N)} \tag{1.19}$$

where the $\beta(j)$ are the distinct exponents which result and the b_k result from combining terms corresponding to a single $\beta(k)$. The $\beta(i)$ are clearly algebraic because they are the sum of the $\alpha(i)$. I want to show that the b_k are actually rational numbers. Since the product in 1.18 is taken for all ordered lists as described above, it follows that for a given k,if $a(k)_i$ is switched with $a(k)_j$, that is, two of the roots of $v_k x^{m_k} + \cdots + u_k$ are switched, then the product is unchanged and so 1.19 is also unchanged. Thus each b_l is a symmetric polynomial in the $a(k)_j, j = 1, \cdots, m_k$ for each k. Consider then a particular b_k. It follows

$$b_k = \sum_{(j_1, \cdots, j_{m_n})} A_{j_1, \cdots, j_{m_n}} a(n)_1^{j_1} \cdots a(n)_{m_n}^{j_{m_n}}$$

and this is symmetric in the $\left\{a(n)_1, \cdots, a(n)_{m_n}\right\}$ (note n is distinguished) the coefficients $A_{j_1, \cdots, j_{m_n}}$ being in the commutative ring $\mathbb{Q}[A(1), \cdots, A(n-1)]$ where $A(p)$ denotes

$$a(k)_1, \cdots, a(k)_{m_p}$$

and so from Theorem A.2.5,

$$b_k = \sum_{(j_1, \cdots, j_{m_n})} B_{j_1, \cdots, j_{m_n}} p_1^{j_1}\left(a(n)_1 \cdots a(n)_{m_n}\right) \cdots p_{m_n}^{j_{m_n}}\left(a(n)_1 \cdots a(n)_{m_n}\right)$$

where the $B_{j_1, \cdots, j_{m_n}}$ are symmetric in $\left\{a(k)_j\right\}_{j=1}^{m_k}$ for each $k \le n-1$ and the p_k^l are elementary symmetri c polynomials. Now doing to $B_{j_1, \cdots, j_{m_n}}$ what was just done to b_k and continuing this way, it follows b_k is a finite sum of rational numbers times powers of elementary polynomials in the various $\left\{a(k)_j\right\}_{j=1}^{m_k}$ for $k \le n$. By Theorem A.2.7 this is a rational number. Thus b_k is a rational number as desired. Multiplying by the product of all the denominators, it follows there exist integers c_i such that

$$0 = c_1 e^{\beta(1)} + c_2 e^{\beta(2)} + \cdots + c_N e^{\beta(N)}$$

which contradicts Lemma A.3.7. ■

This theorem is sufficient to show e is transcendental. If it were algebraic, then

$$e e^{-1} + (-1) e^0 \neq 0$$

but this is not the case. If $a \neq 1$ is algebraic, then $\ln(a)$ is transcendental. To see this, note that

$$1 e^{\ln(a)} + (-1) a e^0 = 0$$

which cannot happen if $\ln(a)$ is algebraic according to the above theorem. If a is algebraic and $\sin(a) \neq 0$, then $\sin(a)$ is transcendental because

$$\frac{1}{2i} e^{ia} - \frac{1}{2i} e^{-ia} + (-1)\sin(a) e^0 = 0$$

which cannot occur if $\sin(a)$ is algebraic. There are doubtless other examples of numbers which are transcendental by this amazing theorem. For example, π is also transcendental. This is because $1 + e^{i\pi} = 0$. This couldn't happen if π were algebraic because then so would be $i\pi$.

Of course this marvelous theorem is insufficient to classify an arbitrary real number, even many which are well specified like $\pi + e$.

Appendix B

Integration on Rough Paths*

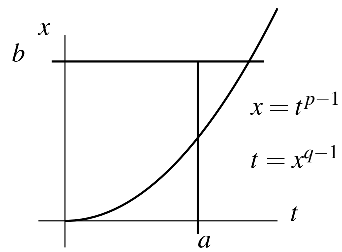
The material on Stieltjes integrals has a very important generalization called integration on rough functions. This chapter gives an introduction to this topic. In order to show this, we need a simple inequality called the triangle inequality. First here is a useful lemma.

As in the case of Stieltjes integrals all of this has generalizations to integrator functions which have values in various normed linear spaces but this is a book on single variable advanced calculus and so this level of generality is avoided.

Lemma B.0.1 *If $a,b \geq 0, p > 1$ and p' is defined by $\frac{1}{p}+\frac{1}{p'}=1$, then*

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

Proof: Let $p' = q$ to save on notation and 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 $a^p = b^q$. ■

The following is a case of Holder's inequality.

Lemma B.0.2 *Let $a_i, b_i \geq 0$. Then for $p \geq 1$,*

$$\sum_i a_i b_i \leq \left(\sum_i a_i^p\right)^{1/p} \left(\sum_i b_i^{p'}\right)^{1/p'}$$

Proof: From the above inequality,

$$\sum_{i=1}^n \frac{a_i}{\left(\sum_i a_i^p\right)^{1/p}} \frac{b_i}{\left(\sum_i b_i^{p'}\right)^{1/p'}} \leq \sum_{i=1}^n \frac{1}{p}\left(\frac{a_i^p}{\sum_i a_i^p}\right) + \frac{1}{p'}\left(\frac{b_i^{p'}}{\sum_i b_i^{p'}}\right)$$

$$= \frac{1}{p}\left(\frac{\sum_i a_i^p}{\sum_i a_i^p}\right) + \frac{1}{p'}\left(\frac{\sum_i b_i^{p'}}{\sum_i b_i^{p'}}\right) = \frac{1}{p} + \frac{1}{p'} = 1$$

Hence the inequality follows from multiplying both sides by $\left(\sum_i a_i^p\right)^{1/p} \left(\sum_i b_i^{p'}\right)^{1/p'}$. ■

Then with this lemma, here is the triangle inequality.

Theorem B.0.3 *Let $a_i, b_i \in \mathbb{R}$. Then for $p \geq 1$,*

$$\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |b_i|^p\right)^{1/p}$$

Proof: First note that from the definition, $p - 1 = p/p'$.

$$\sum_{i=1}^{n} |a_i + b_i|^p \leq \sum_{i=1}^{n} |a_i + b_i|^{p-1} (|a_i| + |b_i|)$$

$$\leq \sum_{i=1}^{n} |a_i + b_i|^{p/p'} |a_i| + \sum_{i=1}^{n} |a_i + b_i|^{p/p'} |b_i|$$

Now from Lemma B.0.2,

$$\leq \left(\sum_{i=1}^{n} \left(|a_i + b_i|^{p/p'}\right)^{p'}\right)^{1/p'} \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} \left(|a_i + b_i|^{p/p'}\right)^{p'}\right)^{1/p'} \left(\sum_{i=1}^{n} |b_i|^p\right)^{1/p}$$

$$= \left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/p'} \left(\left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |b_i|^p\right)^{1/p}\right)$$

In case $\sum_{i=1}^{n} |a_i + b_i|^p = 0$ there is nothing to show in the inequality. It is obviously true. If this is nonzero, then divide both sides of the above inequality by $\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/p'}$ to get

$$\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1-\frac{1}{p'}} = \left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |b_i|^p\right)^{1/p} \blacksquare$$

B.1 Finite p Variation

Instead of integrating with respect to a finite variation integrator function F, the function will be of finite p variation. This is more general than finite variation. Here it is assumed $p > 0$ rather than $p > 1$.

Definition B.1.1 *Define for a function $F : [0, T] \to \mathbb{R}$*

1. *α Holder continuous if $\sup_{0 \leq s < t \leq T} \frac{|F(t) - F(s)|}{|t-s|^{\alpha}} < C < \infty$. Then from this inequality, it follows that $|F(t) - F(s)| \leq C |t - s|^{\alpha}$.*
2. *The function F has finite p variation if for some $p > 0$,*

$$\|F\|_{p,[0,T]} \equiv \sup_{\mathscr{P}} \left(\sum_{i=1}^{m} |F(t_{i+1}) - F(t_i)|^p\right)^{1/p} < \infty$$

where $\mathscr{P}$ denotes a partition of $[0, T]$, $\mathscr{P} = \{t_0, t_1, \cdots, t_n\}$ for

$$0 = t_0 < t_1 < \cdots < t_n = T$$

also called a dissection in this subject. It was also called a division when discussing the generalized Riemann integral. $|\mathscr{P}|$ denotes the largest length in any of the sub intervals. It will be always assumed that actually $p \geq 1$.

Note that when $p=1$ having finite p variation is just the same as saying that it has finite total variation. Thus this is including more general considerations. Also, to simplify the notation, for $\mathscr{P}$ such a dissection, write

$$\sum_{\mathscr{P}} |F(t_{i+1}) - F(t_i)|^p \text{ instead of } \sum_{i=1}^{n} |F(t_{i+1}) - F(t_i)|^p$$

Definition B.1.2 *Let $C^{\alpha}([0,T];\mathbb{R})$ denote the α Holder functions and denote by $V^p([0,T],\mathbb{R})$ the continuous functions F which have finite p variation, V^p for short.*

It is routine to verify that if $\alpha > 1$, then any Holder continuous function is a constant. It is also easy to see that any $1/p$ Holder is p finite variation. To see this, note that you have $|F(t) - F(s)| \leq C|t-s|^{1/p}$ and so

$$\left(\sum_{i=1}^{m} |F(t_{i+1}) - F(t_i)|^p\right)^{1/p} \leq \left(\sum_{i=1}^{m} C^p |t_{i+1} - t_i|\right)^{1/p} = CT^{1/p}$$

From now on $p \geq 1$ and define $\|F\|_{V^p([0,T],\mathbb{R})} \equiv \|F\|_{p,[0,T]} + \sup_{t\in[0,T]} |F(t)|$ where as described above,

$$\|F\|_{p,[0,T]} \equiv \sup_{\mathscr{P}} \left(\sum_{i} |F(t_{i+1}) - F(t_i)|^p\right)^{1/p} = \left(\sup_{\mathscr{P}} \sum_{i} |F(t_{i+1}) - F(t_i)|^p\right)^{1/p}$$

To save notation, it is customary to write $\|F\|_{\infty} = \sup_{t\in[0,T]} |F(t)|$

Definition B.1.3 *Suppose you have a set of functions V defined on some interval I which satisfies $cX \in V$ whenever c is a number and $X \in V$ and that $X+Y \in V$ whenever $X,Y \in V$. Then $\|\cdot\| : V \to [0,\infty)$ is a norm if it satisfies.*

$$\|X\| \geq 0, \|X\| = 0 \text{ if and only if } X = 0$$

$$\|X+Y\| \leq \|X\| + \|Y\|$$

$$\text{For } c \text{ a number, } \|cX\| = |c| \, \|X\|$$

Proposition B.1.4 *For each $p \geq 1, V^p$ is a set of continuous functions. Also $\|\cdot\|_{V^p}$ is a norm. In addition, if $1 \leq p < q$,*

$$V^p([0,T];\mathbb{R}) \subseteq V^q([0,T];\mathbb{R}) \subseteq C^0([0,T];\mathbb{R})$$

The embeddings are continuous. In fact, $\|F\|_{\infty} \leq \|F\|_{V^q} \leq \|F\|_{V^p}$.

Note that, although $\|\cdot\|_{V^p}$ is a norm, $\|\cdot\|_{p,[0,T]}$ is not.

Proof: It is clear that $\|F\|_{V^p}$ equals 0 if and only if $F=0$. This follows from the inclusion in the definition for the norm, $\sup_{t\in[0,T]} |F(t)|$. It only remains to verify the other axioms of a norm. It suffices to consider $\|\cdot\|_{p,[0,T]}$. Does it satisfy the triangle inequality?

$$\|Z+Y\|_{p,[0,T]} \equiv \sup_{\mathscr{P}} \left(\sum_{i} |(Z+Y)(t_{i+1}) - (Z+Y)(t_i)|^p\right)^{1/p}$$

$$
\begin{aligned}
&\leq \sup_{\mathscr{P}}\left(\sum_i\left(\left|Z\left(t_{i+1}\right)-Z\left(t_i\right)\right|+\left|Y\left(t_{i+1}\right)-Y\left(t_i\right)\right|\right)^p\right)^{1/p} \\
&\leq \sup_{\mathscr{P}}\left[\left(\sum_i\left|Z\left(t_{i+1}\right)-Z\left(t_i\right)\right|^p\right)^{1/p}+\left(\sum_i\left|Y\left(t_{i+1}\right)-Y\left(t_i\right)\right|^p\right)^{1/p}\right]
\end{aligned}
$$

This is by

$$
\begin{aligned}
&\leq \sup_{\mathscr{P}}\left(\sum_i\left|Z\left(t_{i+1}\right)-Z\left(t_i\right)\right|^p\right)^{1/p}+\sup_{\mathscr{P}}\left(\sum_i\left|Y\left(t_{i+1}\right)-Y\left(t_i\right)\right|^p\right)^{1/p} \\
&= \|Z\|_{p,[0,T]}+\|Y\|_{p,[0,T]}
\end{aligned}
$$

Thus $\|\cdot\|_{V^p}$ clearly is a norm. What about those inclusions? Let $F \in V^p$. Is it also in V^q? Suppose $F \in V^p$ and $\|F\|_{p,[0,T]} = 1$. What about $\|F\|_{q,[0,T]}$? For any $\mathscr{P}$

$$
\left(\sum_i\left|F\left(t_{i+1}\right)-F\left(t_i\right)\right|^q\right)^{1/p} \leq \left(\sum_i\left|F\left(t_{i+1}\right)-F\left(t_i\right)\right|^p\right)^{1/p} \leq 1
$$

and so $\|F\|_{q,[0,T]}^{q/p} \leq 1$. Therefore, if $\|F\|_{p,[0,T]} < \infty$, $\left\|\frac{F}{\|F\|_{p,[0,T]}}\right\|_{q,[0,T]}^{q/p} \leq 1$ and so

$$
\left\|\frac{F}{\|F\|_{p,[0,T]}}\right\|_{q,[0,T]} \leq 1, \quad \|F\|_{q,[0,T]} \leq \|F\|_{p,[0,T]}
$$

as claimed. Thus $V^p \subseteq V^q$ and the inclusion map is continuous. There is nothing to verify for the uniform norm part of $\|F\|_{V^p}$. How about the embedding into C^0? $\|F\|_{C^0} \leq \|F\|_{V^p}$by definition. ■

In the above, there is nothing sacred about the interval $[0,T]$. You could use any other interval. Then we write $\|F\|_{V^p(I)}$, $\|F\|_{p,I}$ to denote the above with respect to the interval I.

Lemma B.1.5 *Suppose $a = t_0 < t_1 < \cdots < t_n = b$. Then $\sum_{i=0}^{n-1} \|F\|_{p,[t_i,t_{i+1}]}^p \leq \|F\|_{p,[a,b]}^p$*

Proof: It is sufficient to verify this with two intervals. Say $a < b < c$. Then let $\mathscr{P}_1$ be a dissection for $[a,b]$ and $\mathscr{P}_2$ a dissection for $[b,c]$. Then $\mathscr{P} = \mathscr{P}_1 \cup \mathscr{P}_2$ is clearly a dissection for $[a,c]$. Then from definition,

$$
\sum_{\mathscr{P}_1}\left|F\left(t_{i+1}\right)-F\left(t_i\right)\right|^p+\sum_{\mathscr{P}_2}\left|F\left(t_{i+1}\right)-F\left(t_i\right)\right|^p=\sum_{\mathscr{P}}\left|F\left(t_{i+1}\right)-F\left(t_i\right)\right|^p \leq \|F\|_{p,[a,c]}^p
$$

Then taking the sup over all such $\mathscr{P}_1$ one gets

$$
\|F\|_{p,[a,b]}^p+\sum_{\mathscr{P}_2}\left|F\left(t_{i+1}\right)-F\left(t_i\right)\right|^p \leq \|F\|_{p,[a,c]}^p
$$

Now take sup over all such $\mathscr{P}_2$. Note that in finding $\|F\|_{p,[a,c]}^p$ there is no guarantee that b will be in any of the dissections. That is, although $\mathscr{P}_1 \cup \mathscr{P}_2$ is a disection of $[a,c]$ it might not be any of the dissections needed to obtain $\|F\|_{p,[a,c]}^p$. Thus we can't expect to have this inequality an equation. ■

B.2 Piecewise Linear Approximation

Definition B.2.1 *Let $\mathscr{P}$ be a dissection of $[0,T]$ and let $F \in V^p([0,T])$. Let $F^{\mathscr{P}}$ denote the piecewise linear approximation of F. That is, it agrees with F at every point of $\mathscr{P}$ and in $[t_i, t_{i+1}]$ it is of the form $\frac{1}{t_{i+1}-t_i}[(F(t_i))(t_{i+1}-t)+F(t_{i+1})(t-t_i)] = F(t_i) + (t-t_i)\left(\frac{F(t_{i+1})-F(t_i)}{t_{i+1}-t_i}\right)$.*

To get the piecewise linear approximation, you could write

$$F^{\mathscr{P}}(t) \equiv \int_0^t \left(\frac{F(0)}{t_1-t_0} + \sum_{i=0}^{n-1} \frac{F(t_{i+1})-F(t_i)}{t_{i+1}-t_i}\right) \mathscr{X}_{[t_i,t_{i+1})}(s)\,ds$$

Note that the formula gives

$$\begin{aligned} F^{\mathscr{P}}(t_1) &= F(0)+F(t_1)-F(t_0) = F(t_1), \\ F^{\mathscr{P}}(t_2) &= F(t_1) + \frac{F(t_2)-F(t_1)}{t_2-t_1}(t_2-t_1) = F(t_2) \end{aligned}$$

etc.

Next is a fundamental approximation lemma which says that when you replace a function in V^p with its piecewise linear approximation the p variation gets smaller. First is a simple observation. Suppose $p \geq 1$ and $\sum_{i=1}^n r_i = r$ where each r_i is positive and less than 1. Then $r^p \geq \sum_{i=1}^n r_i^p$. To see this is so, note that $\sum_{i=1}^n \frac{r_i}{r} = 1$ and so $\sum_i \left(\frac{r_i}{r}\right)^p \leq \sum_i \frac{r_i}{r} = 1$ so the claim follows.

Lemma B.2.2 *Let $F \in V^p([0,T])$ and let $\mathscr{P}$ be a dissection. Then $\left\|F^{\mathscr{P}}\right\|_{p,[0,T]} \leq \|F\|_{p,[0,T]}$. Also, if $\mathscr{P}_\varepsilon$ is a dissection for which*

$$\sum_{\mathscr{P}_\varepsilon} \left|F^{\mathscr{P}}(u_{i+1}) - F^{\mathscr{P}}(u_i)\right|^p > \left\|F^{\mathscr{P}}\right\|_{p,[0,T]}^p - \varepsilon,$$

then this inequality continues to hold for $\mathscr{P}_\varepsilon$ in which $\mathscr{P}_\varepsilon \subseteq \mathscr{P}$.

Proof: Let $\mathscr{P}_\varepsilon$ be a dissection of $[0,T]$ such that

$$\sum_{u_i \in \mathscr{P}_\varepsilon} \left|F^{\mathscr{P}}(u_{i+1}) - F^{\mathscr{P}}(u_i)\right|^p > \left\|F^{\mathscr{P}}\right\|_{p,[0,T]}^p - \varepsilon \tag{2.1}$$

The idea is to show that, deleting some points of $\mathscr{P}_\varepsilon$ you can assume every grid point of $\mathscr{P}_\varepsilon$ is also a grid point of $\mathscr{P}$ while preserving the above inequality. This might not be too surprising because the variation of $F^{\mathscr{P}}$ is determined by the $F(t_i)$ where $t_i \in \mathscr{P}$. Let the mesh points of $\mathscr{P}_\varepsilon$ be denoted by $\{u_i\}$ and those of $\mathscr{P}$ are denoted by $\{t_j\}$.

Let (t_j, t_{j+1}) be an interval which contains points of $\mathscr{P}_\varepsilon$ say $u_i, u_{i+1}, \ldots, u_k$ in order. Modify $\mathscr{P}$ if necessary by including t_j, t_{j+1}. Then corresponding to this interval, the above sum gives

$$\sum_{l=i}^{k-1} |a(u_{l+1}-u_l)|^p + \left|a(t_{j+1}-u_k)\right|^p + \left|a(u_i - t_j)\right|^p$$

where $a \equiv \frac{F(t_{j+1})-F(t_j)}{t_{j+1}-t_j}$. Then this sum equals

$$\left|F\left(t_{j+1}\right)-F\left(t_j\right)\right|^p\left(\sum_{l=i}^{k-1}\left|\frac{u_{l+1}-u_l}{t_{j+1}-t_j}\right|^p+\left|\frac{t_{j+1}-u_k}{t_{j+1}-t_j}\right|^p+\left|\frac{u_i-t_j}{t_{j+1}-t_j}\right|^p\right).$$

If all the u_r were deleted from $\mathscr{P}_\varepsilon$ for $i<r<k$, this results in $\left|F\left(t_{j+1}\right)-F\left(t_j\right)\right|^p$ which must be at least as large from the above observation because $\sum_{l=i}^{k-1}\left|\frac{u_{l+1}-u_l}{t_{j+1}-t_j}\right|+\left|\frac{u_i-t_j}{t_{j+1}-t_j}\right|+\left|\frac{t_{j+1}-u_k}{t_{j+1}-t_j}\right|=1$and $p\geq 1$ and all these quotients are positive and smaller than 1. Thus we get a larger sum on the left in 2.1 by modifying $\mathscr{P}_\varepsilon$ as just described. This shows that it can be assumed that every point of $\mathscr{P}_\varepsilon$ is in $\mathscr{P}$. Therefore, since each $u_i\in\mathscr{P}_\varepsilon$

$$\|F\|_{p,[0,T]}^p\geq\sum_{u_i\in\mathscr{P}_\varepsilon}\left|F\left(u_{i+1}\right)-F\left(u_i\right)\right|^p=\sum_{u_i\in\mathscr{P}_\varepsilon}\left|F^{\mathscr{P}}\left(u_{i+1}\right)-F^{\mathscr{P}}\left(u_i\right)\right|^p>\left\|F^{\mathscr{P}}\right\|_{p,[0,T]}^p-\varepsilon$$

Since ε is arbitrary, this shows the conclusion of the lemma. ■

Next is to show that these piecewise linear functions approximate F in terms of p variation. In general, when you have a continuous function defined on a closed interval, the above piecewise continuous approximations always converge to the function uniformly. This is left to the reader to show. It is an exercise in uniform continuity of F.

Theorem B.2.3 *Let $F\in V^p([0,T]), p\geq 1$. Then $\lim_{|\mathscr{P}|\to 0}\left\|F^{\mathscr{P}}-F\right\|_{q,[0,T]}=0$ for every $q>p$.*

Proof: Let $\varepsilon>0$ be given and let $\mathscr{P}$ be a dissection of $[0,T]$.

Using what was just observed about uniform continuity of piecewise linear approximations,

$$\sum_{t_i\in\mathscr{P}}\left|F^{\mathscr{P}}\left(t_{i+1}\right)-F\left(t_{i+1}\right)-\left(F^{\mathscr{P}}\left(t_i\right)-F\left(t_i\right)\right)\right|^q$$

$$\begin{aligned}
&\leq \max_{t_i\in\mathscr{P}}\left(\left|F^{\mathscr{P}}\left(t_{i+1}\right)-F\left(t_{i+1}\right)-\left(F^{\mathscr{P}}\left(t_i\right)-F\left(t_i\right)\right)\right|^{q-p}\right)\cdot\\
&\quad\sum_{t_i\in\mathscr{P}}\left|F^{\mathscr{P}}\left(t_{i+1}\right)-F\left(t_{i+1}\right)-\left(F^{\mathscr{P}}\left(t_i\right)-F\left(t_i\right)\right)\right|^p\\
&\leq \max_{t_i\in\mathscr{P}}\left(\left|F^{\mathscr{P}}\left(t_{i+1}\right)-F\left(t_{i+1}\right)-\left(F^{\mathscr{P}}\left(t_i\right)-F\left(t_i\right)\right)\right|^{q-p}\right)\left\|F^{\mathscr{P}}-F\right\|_{p,[0,T]}^p\\
&\leq \max_{t_i\in\mathscr{P}}\left(\left(\left|F^{\mathscr{P}}\left(t_{i+1}\right)-F\left(t_{i+1}\right)\right|+\left|F^{\mathscr{P}}\left(t_i\right)-F\left(t_i\right)\right|\right)^{q-p}\right)\cdot\\
&\quad\left(\left\|F^{\mathscr{P}}\right\|_{p,[0,T]}+\|F\|_{p,[0,T]}\right)^p\\
&\leq 2^{q-p}\max_{t\in[0,T]}\left|F^{\mathscr{P}}(t)-F(t)\right|^{q-p}\left(2\|F\|_{p,[0,T]}\right)^p
\end{aligned}$$

from Lemma B.2.2 since $\left\|F^{\mathscr{P}}\right\|_{p,[0,T]}\leq\|F\|_{p,[0,T]}$. Now it follows that if $|\mathscr{P}|$ is small enough, the above is no more than ε. ■

The method of proof of the above yields the following useful lemma. To save space, $\|F\|_\infty\equiv\sup_t|F(t)|$.

Lemma B.2.4 *Let $F,Y \in V^p([0,T])$ and let $q > p \geq 1$. Then*

$$\|F-Y\|_{V^q} \leq 2^{(q-p)/q} \|F-Y\|_{\infty}^{\frac{q-p}{q}} \|F-Y\|_{p,[0,T]}^{p/q} + \|F-Y\|_{\infty}$$

Note that $2^{\frac{q-p}{q}} \leq 2$.

Proof: Let $\mathscr{P}$ be a dissection. Then

$$\sum_i |F(t_{i+1}) - Y(t_{i+1}) - (F(t_i) - Y(t_i))|^q$$

$$\leq \sup_i |F(t_{i+1}) - Y(t_{i+1}) - (F(t_i) - Y(t_i))|^{q-p} \sum_i |F(t_{i+1}) - Y(t_{i+1}) - (F(t_i) - Y(t_i))|^p$$

$$\leq 2^{q-p} \|F-Y\|_{\infty}^{q-p} \sum_i |F(t_{i+1}) - Y(t_{i+1}) - (F(t_i) - Y(t_i))|^p$$

Then taking the sup over all dissections,

$$\|F-Y\|_{q,[0,T]}^q \leq 2^{q-p} \|F-Y\|^{q-p} \|F-Y\|_{p,[0,T]}^p$$

$$\|F-Y\|_{q,[0,T]} \leq 2^{(q-p)/q} \|F-Y\|^{(q-p)/q} \|F-Y\|_{p,[0,T]}^{p/q}$$

Hence $\|F-Y\|_{V^q} \leq 2^{(q-p)/q} \|F-Y\|^{(q-p)/q} \|F-Y\|_{p,[0,T]}^{p/q} + \|F-Y\|_{\infty}$ ■

In all the above, one can replace $[0,T]$ with $[a,b]$ through simple modifications.

B.3 The Young Integral

This dates from about 1936. Basically, you can do $\int_0^T YdF$ if Y is continuous and F is of bounded variation or the other way around. This is the old Stieltjes integral. However, this integral has to do with $Y \in V^q$ and $F \in V^p$ and of course, these functions are not necessarily of bounded variation although they are continuous. First, here is a simple lemma which is used a little later.

Lemma B.3.1 *Let F be piecewise linear with respect to a dissection $\mathscr{P}$ and let Y be continuous. Then $t \to \int_0^t YdF$ is continuous.*

Proof: Say $\mathscr{P}$ is given by $\{t_0, t_1, \cdots, t_n\}$ where $0 = t_0 < \cdots < t_n = T$. Let $G(t) \equiv \int_0^t YdF$. Then on $[0,t_1]$,

$$G(t) = \int_0^t Y(s) \frac{F(t_1) - F(0)}{t_1 - 0} ds = \frac{F(t_1) - F(0)}{t_1 - 0} \int_0^t Y(s) ds$$

which is clearly continuous. Then on $[t_1, t_2]$ you have

$$G(t) = G(t_1) + \int_{t_1}^t Y(s) \frac{F(t_2) - F(t_1)}{t_2 - t_1} ds$$

which is again continuous. Continuing this way shows the desired conclusion. ■

Definition B.3.2 *Let $\mathscr{P}$ be a dissection of $[0,T]$ and let Y, F be continuous. Then*

$$\int_{\mathscr{P}} YdF \equiv \sum_{\mathscr{P}} Y(t_i)(F(t_{i+1}) - F(t_i))$$

where $\mathscr{P} = \{t_0, \cdots, t_n\}$. This is like a Riemann Stieltjes sum except that you don't have a bounded variation integrator function. $\int_0^T YdF$ is said to exist if there is $I \in \mathbb{R}$ such that

$$\lim_{|\mathscr{P}| \to 0} \left| \int_{\mathscr{P}} YdF - I \right| = 0$$

meaning that for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $|\mathscr{P}| < \delta$, it follows that $\left|\int_{\mathscr{P}} YdF - I\right| < \varepsilon$. It suffices to show that for every ε there exists δ such that if $|\mathscr{P}|, |\mathscr{P}'| < \delta$, then

$$\left| \int_{\mathscr{P}} YdF - \int_{\mathscr{P}'} YdF \right| < \varepsilon$$

This last condition says that the set of all these $\int_{\mathscr{P}}$ for $|\mathscr{P}|$ sufficiently small has small diameter.

Note how this looks just like the Stieltjes integral except here one is considering one sided sums just like Cauchy did in the 1820's. The following theorem is from Young.

Theorem B.3.3 *Let $1 \le p, q, \frac{1}{p} + \frac{1}{q} > 1$. Also suppose that $F \in V^p([0,T])$ and $Y \in V^q([0,T])$. Then for all sub interval $[a,b]$ of $[0,T]$ there exists $I_{[a,b]}$ such that*

$$\lim_{|\mathscr{P}| \to 0} \left| \int_{\mathscr{P}} YdF - I \right| = 0$$

exists where here $\mathscr{P} \subseteq [a,b]$ is a dissection of $[a,b]$. Also there exist estimates of the form

$$\left| \int_s^t YdF \right| \le C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[s,t]} \tag{2.2}$$

$$\left| \int_0^{(\cdot)} YdF \right|_{p,[0,T]} \le C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]} \tag{2.3}$$

$$\left| \int_s^t YdF \right| = \left| \int_0^t YdF - \int_0^s YdF \right| \le C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[s,t]}$$

$$\left| \int_0^{(\cdot)} YdF \right|_{V^p([0,T])} \le 2C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]} \tag{2.4}$$

where C_{pq} depends only on p, q. Also $t \to \int_0^t YdF$ is continuous.

Proof: Let $Z \in V^q([0,T]), 1 \le p, q, \frac{1}{p} + \frac{1}{q} > 1$. Define for $s \le t$

$$\omega(s,t) \equiv \|F\|_{p,[s,t]}^p + \|Z\|_{q,[s,t]}^q$$

Here Z will end up being $Y - Y(0)$. From Lemma B.1.5,

$$\omega(t_i, t_{i+1}) + \omega(t_{i+1}, t_{i+2}) + \cdots + \omega(t_{i+r}, t_{i+r+1}) \le \omega(t_i, t_{i+r+1})$$

Also let $\mathscr{P}$ be a dissection $0 = t_0 < t_1 < \cdots < t_r = T$.

Claim: Then there exists t_i such that $\omega(t_{i-1}, t_{i+1}) \leq \frac{5}{r}\omega(0,T)$.

Proof of claim: If $r = 2$ so there are only two intervals in the dissection,

$$\|F\|_{p,[0,T]}^p \leq \|F\|_{p,[0,T]}^p + \|Z\|_{p,[0,T]}^p = \omega(0,T) < \frac{5}{2}\omega(0,T).$$

So assume $r > 2$. Then how many disjoint intervals $[t_{i-1}, t_{i+1}]$ are there for either i odd or i even? Certainly fewer than r and at least as many as $\frac{r}{2}$. Then we have

$$\sum_{i \text{ even}} \omega(t_{i-1}, t_{i+1}) + \sum_{i \text{ odd}} \omega(t_{i-1}, t_{i+1}) \leq 2\omega(0,T)$$

Now if it is not true that there exists such an i, then for each i, $\omega(t_{i-1}, t_{i+1}) > \frac{5}{r}\omega(0,T)$.

$$\frac{5}{2}\omega(0,T) = \frac{5}{r}\omega(0,T)\frac{r}{2} \leq \sum_{i \text{ even}} \frac{5}{r}\omega(0,T) < \sum_{i \text{ even}} \omega(t_{i-1}, t_{i+1}) \leq \omega(0,T)$$

$$\frac{5}{2}\omega(0,T) = \frac{5}{r}\omega(0,T)\frac{r}{2} \leq \sum_{i \text{ odd}} \frac{5}{r}\omega(0,T) < \sum_{i \text{ odd}} \omega(t_{i-1}, t_{i+1}) \leq \omega(0,T)$$

Hence $5\omega(0,T) < 2\omega(0,T)$ which is absurd. Therefore, there exists such a t_i. This shows the claim.

Next points of $\mathscr{P}$ are deleted beginning with t_i where $\omega(t_{i-1}, t_{i+1}) \leq \frac{5}{r}\omega(0,T)$.

$$\begin{aligned}\int_{\mathscr{P}} ZdF - \int_{\mathscr{P}\setminus\{t_i\}} ZdF &= Z(t_{i-1})(F(t_i) - F(t_{i-1})) \\ &+ Z(t_i)(F(t_{i+1}) - F(t_i)) - Z(t_{i-1})(F(t_{i+1}) - F(t_{i-1}))\end{aligned}$$

This equals $Z(t_{i-1})F(t_i) + Z(t_i)F(t_{i+1}) - Z(t_i)F(t_i) - Z(t_{i-1})F(t_{i+1})$

$$\begin{aligned} &= Z(t_i)(F(t_{i+1}) - F(t_i)) + Z(t_{i-1})(F(t_i) - F(t_{i+1})) \\ &= (Z(t_i) - Z(t_{i-1}))(F(t_{i+1}) - F(t_i))\end{aligned}$$

It follows that

$$\left|\int_{\mathscr{P}} ZdF - \int_{\mathscr{P}\setminus\{t_i\}} ZdF\right| \leq |(Z(t_i) - Z(t_{i-1}))|\,|F(t_{i+1}) - F(t_i)|$$

$$\leq \omega(t_{i-1}, t_i)^{1/q}\omega(t_i, t_{i+1})^{1/p} \leq \omega(t_{i-1}, t_{i+1})^{1/p+1/q} \leq \left(\frac{5}{r}\omega(0,T)\right)^{1/p+1/q}$$

In particular, this yields

$$\begin{aligned}&\left|\int_{\mathscr{P}} (Y - Y(0))\,dF - \int_{\mathscr{P}\setminus\{t_i\}} (Y - Y(0))\,dF\right| \leq \left(\frac{5}{r}\omega(0,T)\right)^{1/p+1/q} \\ =\ &\left(\frac{5}{r}\left(\|F\|_{p,[0,T]}^p + \|Y\|_{q,[0,T]}^q\right)\right)^{1/p+1/q}\end{aligned}$$

It follows that removing the special point from the claim out of those t_i which remain, denoted as $t^1, t^2, t^3, \cdots$,

$$\begin{aligned}
&\left|\int_{\mathscr{P}} (Y - Y(0))\, dF - \int_{\mathscr{P}\setminus\{t^1\}} (Y - Y(0))\, dF\right| \\
&+ \left|\int_{\mathscr{P}\setminus\{t^1\}} (Y - Y(0))\, dF - \int_{\mathscr{P}\setminus\{t^1,t^2\}} (Y - Y(0))\, dF\right| + \\
&+ \cdots + \left|\int_{\mathscr{P}\setminus\{t^1,t^2,\cdots t^{r-2}\}} (Y - Y(0))\, dF - \int_{\mathscr{P}\setminus\{t^1,t^2,\cdots t^{r-1}\}} (Y - Y(0))\, dF\right| \\
&\leq 5^{1/p+1/q} (\omega(0,T))^{1/p+1/q} \left(\left(\frac{1}{r}\right)^{1/p+1/q} + \left(\frac{1}{r-1}\right)^{1/p+1/q} + \cdots + 1\right) \\
&\leq 5^{1/p+1/q} (\omega(0,T))^{1/p+1/q} \sum_{r=1}^{\infty} \left(\frac{1}{r}\right)^{1/p+1/q}
\end{aligned}$$

The triangle inequality applied to that sum of terms inside $||$ implies

$$\begin{aligned}
&\left|\int_{\mathscr{P}} (Y - Y(0))\, dF - \int_{\mathscr{P}\setminus\{t^1,t^2,\cdots t^{r-1}\}} (Y - Y(0))\, dF\right| \\
= &\left|\int_{\mathscr{P}} (Y - Y(0))\, dF\right| \leq 5^{1/p+1/q} (\omega(0,T))^{1/p+1/q} \sum_{r=1}^{\infty} \left(\frac{1}{r}\right)^{1/p+1/q}
\end{aligned}$$

since $\int_{\mathscr{P}\setminus\{t^1,t^2,\cdots t^{r-1}\}} (Y - Y(0))\, dF = 0$. This is true because by definition it equals

$$(Y(0) - Y(0))(Y(T) - Y(0))$$

Of course all of this works for F replaced with $\tilde{F}_t \equiv \frac{F_t}{\|F\|_{p,[0,T]}}$, and Y replaced with $\tilde{Y}$ where $\tilde{Y}_t \equiv \frac{Y_t}{\|Y\|_{q,[0,T]}}$. Then in this case it is very convenient because $\omega(0,T) \leq 2$. Then the above reduces to

$$\left|\int_{\mathscr{P}} \left(\tilde{Y} - \tilde{Y}(0)\right) d\tilde{F}\right| \leq 5^{1/p+1/q} (2)^{1/p+1/q} \sum_{r=1}^{\infty} \left(\frac{1}{r}\right)^{1/p+1/q} \equiv \hat{C}_{pq}$$

Then also,

$$\left|\int_{\mathscr{P}} (Y - Y(0))\, dF\right| \leq \hat{C}_{pq} \|Y\|_{q,[0,T]} \|F\|_{p,[0,T]}$$

Then this implies

$$\begin{aligned}
\left|\int_{\mathscr{P}} Y dF\right| &\leq \hat{C}_{pq} \|Y\|_{q,[0,T]} \|F\|_{p,[0,T]} + \|Y(0)\|_{\infty} \|F_T - F_0\| \\
&= \hat{C}_{pq} \|Y\|_{q,[0,T]} \|F\|_{p,[0,T]} + \|Y(0)\|_{\infty} \left(\|F_T - F_0\|^p\right)^{1/p}
\end{aligned}$$

$$\begin{aligned}
&\leq \hat{C}_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]} + \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]} \\
&= \left(\hat{C}_{pq} + 1\right) \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]} \equiv C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]}
\end{aligned}$$

Of course there was absolutely nothing special about $[0,T]$. We could have used the interval $[s,t]$ just as well and concluded that for $\mathscr{P}$ a dissection of $[s,t]$,

$$\left|\int_{\mathscr{P}} YdF\right| \leq C_{pq}\left\|Y\right\|_{V^q([0,T])}\left\|F\right\|_{p,[s,t]} \tag{2.5}$$

Now it is possible to show the existence of the integral. By Theorem B.2.3 there exists a sequence of piecewise linear functions $\{F_n\}$ which converge to F in $V^{p'}([0,T])$ where p' is chosen larger than p but still $\frac{1}{p'}+\frac{1}{q}>1$. Then letting $\varepsilon>0$ be given, there exists N such that if $n\geq N$,

$$\left|\int_{\mathscr{P}} Yd(F-F_n)\right| = \left|\int_{\mathscr{P}} YdF - \int_{\mathscr{P}} YdF_n\right| \leq C_{p'q}\left\|Y\right\|_{V^q}\left\|F-F_n\right\|_{p',[0,T]} < \frac{\varepsilon}{3}$$

because F_n converges to F in $V^{p'}$. Also, there exists δ such that if $|\mathscr{P}|,|\mathscr{P}'|<\delta$,

$$\left|\int_{\mathscr{P}} YdF_N - \int_{\mathscr{P}'} YdF_N\right| < \frac{\varepsilon}{3}$$

Then you have, for such $\mathscr{P}$, $\mathscr{P}'$,

$$\left|\int_{\mathscr{P}} YdF - \int_{\mathscr{P}'} YdF\right| \leq \left|\int_{\mathscr{P}} YdF - \int_{\mathscr{P}} YdF_N\right|$$

$$+\left|\int_{\mathscr{P}} YdF_N - \int_{\mathscr{P}'} YdF_N\right| + \left|\int_{\mathscr{P}'} YdF_N - \int_{\mathscr{P}'} YdF\right| < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon$$

Showing that the limit of the $\int_{\mathscr{P}} YdF$ exists as $|\mathscr{P}|\to 0$.

All of the above applies for any sub interval of $[0,T]$. Also from 2.5, you can take a limit as $|\mathscr{P}|\to 0$ and conclude that $\int_s^t YdF$ exists and that

$$\left|\int_s^t YdF\right| \leq C_{pq}\left\|Y\right\|_{V^q([0,T])}\left\|F\right\|_{p,[s,t]} \tag{2.6}$$

This proves the existence of the integral. Does it have the usual properties of an integral? In particular, if $0<a<T$, is $\int_0^a YdF+\int_a^T YdF=\int_0^T YdF$? This is clearly true for the ordinary Stieltjes integral coming from F_n. Thus from the above estimate, let $\{F_n\}$ be a sequence of piecewise linear functions converging to F in $V^p([0,T])$. Then this convergence takes place in $V^p([0,a])$ and $V^p([a,T])$ also. Then

$$\begin{aligned}
\left|\int_0^a YdF - \int_0^a YdF_n\right| &\leq C_{pq}\left\|Y\right\|_{V^q([0,T])}\left\|F-F_n\right\|_{p,[0,a]} \\
\left|\int_a^T YdF - \int_a^T YdF_n\right| &\leq C_{pq}\left\|Y\right\|_{V^q([0,T])}\left\|F-F_n\right\|_{p,[a,T]} \\
\left|\int_0^T YdF_n - \int_0^T YdF\right| &\leq C_{pq}\left\|Y\right\|_{V^q([0,T])}\left\|F-F_n\right\|_{p,[0,T]}
\end{aligned}$$

Then letting n be large enough, the right sides of the above are all less than ε. Hence, from the triangle inequality,

$$\left|\begin{array}{c}\int_0^a YdF-\int_0^a YdF_n+\left(\int_a^T YdF-\int_a^T YdF_n\right) \\ +\left(\int_0^T YdF_n-\int_0^T YdF\right)\end{array}\right| < 3\varepsilon$$

but that on the inside equals $\left|\int_0^a YdF + \int_a^T YdF - \int_0^T YdF\right|$ because this property of the integral holds for Stieltjes integrals. Thus the claim holds. Other properties of the integral can also be inferred from this approximation with the Stieltjes integrals for piecewise linear F_n. It follows that

$$\left|\int_s^t YdF\right| = \left|\int_0^t YdF - \int_0^s YdF\right| \leq C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[s,t]}$$

What of continuity of $t \to \int_0^t YdF$? Letting F_n be the sequence of piecewise linear functions described above, $t \to \int_0^t YdF_n$ is continuous and letting $p' > p$

$$\left|\int_0^t YdF - \int_0^t YdF_n\right| = \left|\int_0^t Yd(F - F_n)\right| \leq C_{p'q} \|Y\|_{V^q([0,T])} \|F - F_n\|_{p',[0,T]}$$

Thus, as $n \to \infty$, one has uniform convergence of the continuous functions $t \to \int_0^t YdF_n$ to $t \to \int_0^t YdF$ thanks to Theorem B.2.3 which implies that the right side converges to 0 since $p' > p$.

Consider the other estimate 2.4. Let $\mathscr{P}$ be a dissection $0 = t_0 < t_1 < \cdots < t_n = T$. Let $\Psi(t) \equiv \int_0^t YdF$. From the definition of the integral in terms of a limit as $|\mathscr{P}| \to 0$, it follows that, since $p \geq 1$,

$$\sum_{i=0}^{n-1} |\Psi(t_{i+1}) - \Psi(t_i)|^p \leq \sum_{i=0}^{n-1} \left|\int_{t_i}^{t_{i+1}} YdF\right|^p \leq \sum_{i=0}^{n-1} C_{pq}^p \|Y\|_{V^q([0,T])}^p \|F\|_{p,[t_i,t_{i+1}]}^p$$

$$\begin{aligned}
&\leq C_{pq}^p \|Y\|_{V^q([0,T])}^p \sum_{i=0}^{n-1} \|F\|_{p,[t_i,t_{i+1}]}^p \leq C_{pq}^p \|Y\|_{V^q([0,T])}^p \left(\sum_{i=0}^{n-1} \|F\|_{p,[t_i,t_{i+1}]}\right)^p \\
&\leq C_{pq}^p \|Y\|_{V^q([0,T])}^p \|F\|_{p,[0,T]}^p
\end{aligned}$$

Recall that since $p \geq 1$, if each $a_i \geq 0, \sum_i a_i^p \leq (\sum_i a_i)^p$. Then taking the sup over all such dissections, it follows that $\left\|\int_0^{(\cdot)} YdF\right\|_{p,[0,T]} \leq C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]}$. Also, from the estimate,

$$\left|\int_0^t YdF\right| \leq C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,t]} \leq C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]}$$

Taking sup for all t, $\left\|\int_0^{(\cdot)} YdF\right\|_\infty \leq C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]}$ and so

$$\left\|\int_0^{(\cdot)} YdF\right\|_{V^p([0,T])} \leq 2C_{pq} \|Y\|_{V^q([0,T])} \|F\|_{p,[0,T]} \blacksquare$$

Bibliography

[1] **Apostol, T. M.,** *Calculus second edition,* Wiley, 1967.

[2] **Apostol T.M.** *Calculus Volume II Second edition,* Wiley *1969.*

[3] **Apostol, T. M.,** *Mathematical Analysis,* Addison Wesley Publishing Co., 1974.

[4] **Artin M.,** *Algebra,* Pearson *2011.*

[5] **Baker, A.** *Transcendental Number Theory*, Cambridge University Press 1975.

[6] **Baker, Roger**, *Linear Algebra*, Rinton Press 2001.

[7] **Bartle R.G.,** *A Modern Theory of Integration,* Grad. Studies in Math., Amer. Math. Society, Providence, RI, 2000.

[8] **Bartle R. G. and Sherbert D.R.** *Introduction to Real Analysis* third edition, Wiley 2000.

[9] **Buck, R. C.** *Advanced Calculus* 2 edition. McGraw-Hill, 1965.

[10] **Chahal J.S.,** *Historical Perspective of Mathematics 2000 B.C. - 2000 A.D. Kendrick Press, Inc. (2007)*

[11] **D'Angelo, J. and West D**. *Mathematical Thinking Problem Solving and Proofs*, Prentice Hall 1997.

[12] **Conway J. B.** *Functions of one Complex variable Second edition,* Springer Verlag 1978.

[13] **Fitzpatrick P. M.,** *Advanced Calculus a course in Mathematical Analysis,* PWS Publishing Company 1996.

[14] **Hardy G.,** *A Course Of Pure Mathematics, Tenth edition,* Cambridge University Press 1992.

[15] **Halmos P.,** *Naive Set Theory,* Van Nostrand 1960.

[16] **Henstock R**. *Lectures on the Theory of Integration,* World Scientific Publiching Co. 1988.

[17] **Hewitt E.** and **Stromberg K.** *Real and Abstract Analysis,* Springer-Verlag, New York, 1965.

[18] **Hobson E.W.,** The Theory of functions of a Real Variable and the Theory of Fourier's Series V. 1, Dover 1957.

[19] **Jones F.,** *Lebesgue Integration on Euclidean Space,* Jones and Bartlett 1993.

[20] **Kuttler K. L.**, *Basic Analysis*, Rinton

[21] **Kuttler K.L.,** *Modern Analysis* CRC Press 1998.

[22] **McLeod R.** *The Generalized Riemann Integral,* Mathematical Association of America, Carus Mathematical Monographs number 20 1980.

[23] **McShane E. J.** *Integration,* Princeton University Press, Princeton, N.J. 1944.

[24] **Rudin, W.,** *Principles of mathematical analysis,* McGraw Hill third edition 1976

[25] **Rudin W.,** *Real and Complex Analysis,* third edition, McGraw-Hill, 1987.

[26] **Widder, D**. *Advanced Calculus,* second edition, Prentice Hall 1961.

[27] **Zygmund A.**. *Trigonometric Series, Volume I and II*, Cambridge University Press 1968.

Index